Sarmila Banerjee
Anjan Chakrabarti *Editors*

Development and Sustainability

India in a Global Perspective

Springer

Development and Sustainability

Sarmila Banerjee · Anjan Chakrabarti
Editors

Development and Sustainability

India in a Global Perspective

Springer

Editors
Sarmila Banerjee
Anjan Chakrabarti
Economics
University of Calcutta
Kolkata
West Bengal
India

ISBN 978-81-322-1123-5 ISBN 978-81-322-1124-2 (eBook)
DOI 10.1007/978-81-322-1124-2
Springer New Delhi Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013936967

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Dedicated in fond memory of
Kalyan Kumar Sanyal
(May 12, 1951–February 18, 2012)
and
Pabitra Kumar Giri
(October 1, 1955–May 20, 2011)

Preface

This volume is the fall out of a capacity building endeavor initiated in 2007 in the Department of Economics of the University of Calcutta, India, under the aegis of Special Assistance Program of the University Grants Commission in the form of Departmental Research Support (DRS) over the Eleventh Five Year Plan period (2007–2012). The thrust area was "Globalization and Sustainable Development" and the objective was to promote research on sustainable development in face of global connectivity, develop curriculum for research scholars, impart training to the young researchers, and build up interactive platform for exchange of ideas. Over this phase, a number of academicians from different institutions of India came and spent time with our faculty and students as Visiting Scholars and quite a few research methodology workshops were held to provide hands-on training in data management with statistical software. The DRS also organized many seminars and conferences to provide opportunity to disseminate and share research findings for assessing our own progress as well as inadequacies. In the process, it has been realized that just like 'sustainability' the concept of 'globalization' is also a multi-faceted one and they together interrogate the existing paradigm of development by highlighting the need for accommodating new contexts, issues, organizations, and institutions through tailoring appropriate theory and methods to handle them with sufficient analytical rigor. As an illustration of this realization, this volume has been planned where most of the articles are jointly written by our faculty and research scholars in collaboration with our visitors and mentors during this phase.

It is felt that a collection of these articles as an edited volume would serve as a good reading material on the present standing of the economy of India for the advanced Masters level students and the beginners in the M. Phil./ Ph.D. courses. The research students need to comprehend the process underlying the conception and formulation of a research plan on economic development. The experience of India is their readily available subject. They may be familiar with Indian reality, observing unprecedented changes all around, but are not always aware of the cumulative influence of globalization on the economy, polity, and society. By raising questions from different angles and perspectives, the articles included in this volume may help them to start exploring the subject better. Other than the students, this volume should also be of interest to researchers of Indian economics

in general, not least because of the wide-ranging topics that are addressed and a conscious emphasis on methodology for examining Indian economic problems.

We are grateful to a number of institutions and individuals who extended their support in completing this project. Mention should be made of the UGC nominees, Professor C. Thangamuthu and Professor N. Lingamurthy, whose constant encouragement and guidance inspired us to remain committed to our goal. Support and cooperation from our visitors, mentors, in-house members, and the fraternity of Economists at large deserves special mention. Quite a few of them contributed in this volume and many others reviewed the papers, gave impartial assessment, and helped in improving the quality of the final output. The research scholars associated with the DRS extended their round-the-clock-cooperation in editing and formatting the final draft. Mention should be made of the enthusiasm and coordinated production plan of Ms. Sagarika Ghosh and Ms. Sahadi Sharma of Springer India and Ms. Sundari of Scientific Publishing Services (P) Limited which played the most crucial role in making the book available to the reader on time.

During the tenure of DRS-I we have lost two very dear colleagues of ours, Professor Pabitra K. Giri and Professor Kalyan K. Sanyal. Both of them were closely associated with different academic activities undertaken, used to provide wise suggestions and insightful advices, and had brilliant analytical mind. We dedicate this book in their fond memory with deep respect and appreciation.

Kolkata, December 2012

Sarmila Banerjee
Anjan Chakrabarti

Reviewers

Arup Kumar Mallik	Honorary Associate, DRS, University of Calcutta
Partha Ray	Indian Institute of Management, Calcutta
Brati Sankar Chakrabarti	Indian Statistical Institute, Kolkata
Manabendu Chattopadhyay	Indian Statistical Institute, Kolkata
Bhaskar Majumder	Central University of Bihar, Patna
Swati Ghosh	Rabindra Bharati University, Kolkata
Arpita Dhar	Jadavpur University, Kolkata
Bhaswar Moitra	Jadavpur University, Kolkata
Dipayana Dutta Chaudhuri	Indian Institute of Management, Indore
Sunanda Sen	Jamia Milia Islamia University, New Delhi
Kausik Chaudhuri	Madras School of Economics, Chennai
Amita Majumder	Indian Statistical Institute, Kolkata
Priyodarshi Banerjee	Indian Statistical Institute, Kolkata
Purnamita Dasgupta	Institute of Economic Growth, New Delhi
Biswajit Chatterjee	Jadavpur University, Kolkata
Anindita Chakrabarti	Madras School of Economics, Chennai
Ranjita Biswas	School of Women Studies, Jadavpur University, Kolkata

Contents

Contributors

Subhra Bagchi Vidyasagar College for Women, 39, Sankar Ghosh Lane, Kolkata, West Bengal 700006, India, e-mail: shubhrabagchi@gmail.com

Simanti Bandyopadhyay Indian Council for Research on International Economic Relations, Core-6 A, 4th Floor, India Habitat Centre, Lodi Road, New Delhi 110003, India, e-mail: simantib@gmail.com

Asis Kumar Banerjee Institute of Development Studies Kolkata, 27/D, Block DD, Salt Lake, Kolkata 700064, India, e-mail: asisbanerjee.cu@gmail.com

Pratiksha Banerjee Research Fellow Department of Economics, University of Calcutta, Kolkata 700050, India, e-mail: pratikshabanerjee@gmail.com

Sarmila Banerjee Department of Economics, University of Calcutta, Kolkata 700050, India, e-mail: sarmila_b@yahoo.co.in

Joysankar Bhattacharya Central University of Bihar, Patna, India, e-mail: joysankar.bhattacharya@gmail.com

Rabindranath Bhattacharya Honorary Associate UGC DRS-1, Department of Economics, University of Calcutta and Center for Studies in Social Sciences Calcutta (CSSSC), Kolkata, India, e-mail: rnbeco@gmail.com

Rajesh Bhattacharya South Asian University, New Delhi, India, e-mail: rajesh2466@gmail.com

Snehashish Bhattacharya South Asian University, New Delhi, India, e-mail: b.snehashish@gmail.com

Sukanta Bhattacharya Department of Economics, University of Calcutta, Kolkata 700050, India, e-mail: sukanta_70@yahoo.com

Ranajoy Bhattacharyya Indian Institute of Foreign Trade, Kolkata Campus, Kolkata 700091, India, e-mail: branajoy@gmail.com

Sankar Kumar Bhaumik Department of Economics, University of Calcutta, 56A, BT Road, Kolkata 700050, India, e-mail: bhaumiksk@yahoo.co.in

Montu Bose DRS Project Fellow, Department of Economics, University of Calcutta, Kolkata 700050, India, e-mail: monbose@gmail.com

Anjan Chakrabarti Department of Economics, University of Calcutta, Kolkata 700050, India, e-mail: chakanjana@yahoo.co.in

Gagari Chakrabarti Department of Economics, Presidency University, Kolkata 700073, India, e-mail: gagari_1974@yahoo.com

Bhaskar Chatterjee Larsen & Tubro Limited, Chennai, India, e-mail: bhasx68@gmail.com

Susmita Chatterjee Globsyn Business School, Kolkata, India, e-mail: susmita.cb@gmail.com

Saumen Chattopadhyay School of Social Sciences, Jawaharlal Nehru University, New Delhi 110067, India, e-mail: schatto@gmail.com

Sudeshna Chattopadhyay Bidhannagar College, Kolkata, India, e-mail: sudeshna.chattopadhyay@gmail.com

Debabrata Datta Institute of Management and Technology, Gaziabad, India, e-mail: d.debabrata@gmail.com

R. S. Deshpande Institute of Social and Economic Changes, Bangalore 560072, India, e-mail: director@isec.ac.in

Anup Dhar School of Human Studies, Ambedkar University, Kashmere Gate Campus, Lothian Road, New Delhi 110006, India, e-mail: anup@aud.ac.in

Arijita Dutta Department of Economics, University of Calcutta, 56A, BT Road, Kolkata 700050, India, e-mail: dutta.arijita@gmail.com

Mousumi Dutta Department of Economics, Presidency University, Kolkata 700073, India, e-mail: dmousumi1970@gmail.com

Sabitri Dutta DumDum Motijheel Rabindra Mahavidyalaya, 208/B/2 DumDum Road, Kolkata 700074, India, e-mail: sabitri.d@gmail.com

Nilabja Ghosh Institute of Economic Growth, Delhi 110007, India, e-mail: nilabjag@yahoo.com

Zakir Husain Population Research Centre, Institute of Economic Growth, Delhi 110007, India, e-mail: dzhusain@gmail.com

Sabyasachi Kar Institute of Economic Growth, Delhi 110007, India, e-mail: skar_ieg@yahoo.com

Subrata Majumder Commerce Evening Department, St. Xavier's College (Autonomous), Kolkata 700016, India, e-mail: subrata_eco@yahoo.co.in

Tiyas Mallick Tata Consultancy Services, Bengaluru, India, e-mail: tiyas_mallick@yahoo.co.in

Kumarjit Mandal Department of Economics, University of Calcutta, Kolkata 700050, India, e-mail: kumarjitm@hotmail.com

Rabindranath N. Mukhopadhyay Department of Economics, University of Calcutta, Kolkata 700050, India, e-mail: mukhopadhyayr@gmail.com

Sk. Abdul Rashid Raja Birendra Chandra College (affiliated to University of Kalyani), Kandi, Murshidabad, West Bengal, India, e-mail: rashidsa123@rediffmail.com

Biswajit Ray Department of Economics, University of Calcutta, Kolkata, West Bengal 700050, India, e-mail: biswajit_roy1002@yahoo.com

Kalyan Kumar Sanyal University of Calcutta, Kolkata, India

Amitava Sarkar Indian Statistical Institute, Kolkata 700203, India, e-mail: amitavasarkar@hotmail.com

Prasenjit Sarkhel Department of Economics, University of Kalyani, Kalyani, West Bengal 741235, India, e-mail: prasenjitsarkhel@gmail.com

Chitrakalpa Sen Auro University, Surat, Gujarat, India, e-mail: chitrakalpa@gmail.com

Soumyen Sikdar Indian Institute of Management Calcutta, Kolkata, India, e-mail: soumyen1956@gmail.com

About the Editors

Sarmila Banerjee is at present the Rajiv Gandhi Chair Professor of Eco-systems and Sustainable Development in the University of Calcutta. She has been in the Faculty of Economics of the same institution for more than three decades. She obtained her Ph.D. degree in Economics from the University of Florida, Gainesville, USA and carried out her post-doctoral research on Environmental Economics in the University of California, Santa Barbara, USA. Her teaching interests include Econometrics and Environmental Economics. Her research interest is mostly on Quantitative Aspects of Sustainable Development. She is the coordinator of UGC-SAP program in her Department and Convener of Ph.D. Program in Economics at the University of Calcutta. She is connected with different academic institutions and research networks in India and abroad through her active involvement in capacity building and curriculum development.

Anjan Chakrabarti did his Ph.D. in Economics from the University of California, Riverside and is currently Professor of Economics, University of Calcutta. He is the Chairperson, Undergraduate Board of Studies in Economics and Committee Member of the Ph.D. Program in Economics of University of Calcutta. His teaching interests include Political Economy, History of Economic Ideas and Macroeconomics while his research interests include Political Economy, Development Economics, Indian Economics and Political Philosophy. He has published five books and over 30 academic articles. His published books include "Transition and Development in India" and "Dislocation and Resettlement in Development: From Third World to World of the Third", both from Routledge and "World of the Third and Global Capitalism" from Worldview Press. He is the recipient of Dr. V. K. R. V. Rao Prize in Social Science Research in Economics for the year 2008.

Chapter 1
Introduction

Sarmila Banerjee and Anjan Chakrabarti

Context

This book is on the emerging Indian economic map materializing in the backdrop of the phenomena of globalization and sustainability following two decades of liberalization policies. Globalization and sustainability are our chosen entry points which allow us to organize our story in a manner that is able to highlight the essential ideas, events, breaks, problems, as also the rise of new kinds of institutions, relationships, and engagements that are connected to the epochal transition of the Indian economy. Our intervention is as much on method as it is on substance which is indicated by a variety of new methodological approaches, corpus of concepts, and style of reasoning that move beyond the conventional system and known territories of coverage. The story of the Indian economy that emerges from our foray is of a new Order of Things punctuated with novel problems, crises, and modes of intervention by the state that are not comparable to what existed during the planning era.

In the following sections, we discuss the points of departure, the motivation underlying our intervention as also its specific contribution, and the organization of the book.

S. Banerjee · A. Chakrabarti (✉)
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: chakanjana@yahoo.co.in

S. Banerjee
e-mail: sarmila_b@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_1,

The New Economic Map Through the Lens of Globalization and Sustainability

India's economic transition in the last two decades has been marked by growing expectations and mounting societal contrasts. Riding on the crest of a high growth regime, this period has witnessed a structural transformation that is unprecedented in post-independent India, with perhaps the exception of the 1950s when the process of large-scale industrialization unfolded. A list of such changes will lay credence to our claim. Globalization and an open economy, concerted drive toward industrialization (inclusive of expansion in both the manufacturing and service sectors), rise in corporate capitalism, rapid urbanization, and formation of global cities, transformation of many middle-sized towns into cities, with some villages donning increasingly town-like features, increased presence and spread of a privatized media, digital and IT revolution opening new mediums of communication and information flow as also creating a host of novel service products, delivery mechanism and space for social movements, unprecedented market penetration into various aspects of social life, rise of consumer society with new marketing and sales system, phenomenal growth of financial capital and services, shifts in agricultural practices, deep and growing penetration of market and commercialization in agriculture accompanied by shocking level of social destitutions, disturbing exclusions (social and income) and breakdown of agrarian life in many parts of the country, unparalleled migration and trafficking within and across the rural–urban landscape, breakdown of traditional family and community structures, rapidly growing access for women into the public space, public position, and public discourse, shifts in environmental landscape and quality as also in land use, fast changing values and norms, growing awareness of and opposition to extant gender and caste hierarchies are, on their own and in tandem, literally reshaping the cartography of India and its economy. Even a cursory look at the media output, stories and data, overflowing in plentiful, confirm this imagery of an Indian economy in a state of epochal transit.

The above transformation was preceded by four decades of a development plan in which state-sponsored industrialization through capital accumulation became the strategy for fulfilling the defined objective of economic growth; critically, the allocation of resources stayed firmly in the hands of the state and through that it exercised control over the market economy. While the initial development experience in the 1950s reveals an attempt to exploit the trickling-down relation between growth and poverty reduction by pursuing the objective of growth alone, by the late 1960s, problems of food crisis, destitution, and social instability brought to the fore the issue of sustainability, particularly economic and social ones. To partly internalize this problem, the goal of Indian economy accordingly shifted to growth with redistribution (the Fifth Five Year Plan). The crisis of the late 1980s saw the Indian state abandon its state-sponsored development strategy even as it reiterated the objective of growth as central to its development paradigm. As part of the neoliberal philosophy, the Indian economic reforms merged

with the Washington Consensus paradigm to usher in a competitive market economy in a global context. Unlike Eastern European nations, the Indian economy's integration into the global economy so as to inaugurate the formation of global capitalism was not part of the big bang strategy; it was instead a gradualist approach punctuated with gaps and pauses even though the overall tendency should not be doubted. This induced the kind of systemic transformation described in the previous paragraph.

Alongside globalization, the other dimension making its presence felt is that of contrasts and schisms with regard to Indian economy which invite attention to the growing problem of sustainability. The term sustainability encapsulates ecological, economic, social, and even political dimensions that are related to a variety of exclusions (structural, income, and social). That is, the exclusions and their impact are indicative of the dimensions of sustainability, which while arising in relation to global capitalism also impedes its growth or even threatens its stability in some instances. Rather than being independent, these dimensions of sustainability or the underlying forms of exclusions may intersect, reinforce, and compensate one another; this in turn invited the idea of inclusive development and state-sponsored intervention to tame the problems of sustainability. This multi-layered presence of sustainability informs the second node of our intervention. As part of a two-decade evolutionary process, restructuring of the Indian economy not only placed it into the triad of neoliberalism, global capitalism, and inclusive development, but in the process also relocated the state in a complex role where it is simultaneously passive and active (Chakrabarti and Dhar, Chap. 2).

Some of these schisms and contrasts, such as those pertaining to ethnicity, caste, and gender, are old and saw renewal under the new conditions, while some others are original having appeared as a result of the transformation of the Indian economic map. However, by all accounts, the extant and new divides are a result of, and in turn affect, the new economic map emerging out of globalization. As such, notwithstanding their relative autonomy, globalization and sustainability are not independent and autonomous of one another; each feeds into the other and hence configures their respective presence. Particularly revealing are the new kinds of schisms, which at times may even exacerbate old social divisions; these include the simultaneous presence of growing income prosperity and increasing income divide, the emergent divide between groups whose poverty is fast reducing and others (such as Dalits and Adivasis) who are trailing far behind, relatively prosperous urbanity in comparison to the lagging rural heartland, rapid industrialization-urbanization and fast changing/eroding ecological and environmental topography, expansion of industrial capitalism accompanied by augmented pace of displacement-dislocation, rapid globalization and increasingly growing vulnerability to shocks including financial and currency shocks. This list is only partial. The schisms are simultaneously a hope and a concern; hope because the new Order of Things has enabled India to transcend many fetters previously seen as problematical (examples are technological transformation, increased competitiveness, growing income prosperity, etc.) and a concern because the same Order of Things is producing a nation of stark contrasts that are raising serious sustainability

questions and threatening the gains in the process. This new Order of Things, with its conflicting pulls and pushes, can be seen as producing new institutions, opportunities, negotiations, adjustments, policy challenges, and interventions as also new languages of criticism, contestation, and resistance that, in their totality, are helping propel India's economic transition, albeit not always in the direction and with the outcomes that are expected, predicted, or foreseen.

Themes of Globalization and Sustainable Development

As we explained, the reshaping of India's economic map is also creating a nation of contrasts. The phenomena of reshaping and contrast, on their own and in their intersection, constitute the content of the papers. The reshaping had transpired due to multiple factors, not least India's integration into the global economy and partly global order. India's integration in a globalized economy is perhaps coterminous with its so-called liberalization policies that are hailed as responsible for the so-called miracle high growth rate regime, particularly in the last decade. The influences, effects, and even criticisms of globalization in their varied forms and capacities constitute one node of our intervention.

Globalization has multifaceted dimensions. While a few of its cultural and political features are discussed in some of our presentations, our focus is essentially on the economic dimension. One claim that runs through many of the papers is that neoliberalism is the philosophy underlying globalization, which has induced a transformative effect not only on the Indian economic structure but also on the policy paradigm. Another fallout of globalization is the event of global integration, particularly economic, which also opens up the Indian economy to the mechanism of competition, new kinds of markets, and also ensures that extant markets get transformed by it. Much has been talked about the financialization of the economy that has followed globalization; particularly potent has been the global integration and rapid expansion of the Indian stock and currency market. The claim of these financial markets as capturing the fundamentals of the real economy though must be put under scrutiny in the Indian context especially when counterclaims have forwarded the argument about the relative autonomy of the financial sector vis-a-vis the real sector. Moreover, are the Indian stock and currency markets governed by nonlinearity and chaos? Does the global integration make these markets invariant to policy interventions (Sarkar, Chakrabarti and Sen, Chap. 4)? The challenge of crises materializing from those global connections also demands interrogation so as to tease out the problems in the process of financialization. Unpacking the given theories of global financial crisis and its possible impact or non-impact on the Indian economy acquires special significance. How far can the cause of the current Indian economic crisis be put at the doorstep of global economic crisis or are internal structural forces more responsible for it (Mandal and Kar, Chap. 3)?

Other than financial integration, the other facet of economic form of globalization has been trade-related integration. One of the immediate and central policies following liberalization has been trade liberalization culminating ultimately in India's entry into WTO. The claim has been that trade liberalization will lead to increased competitiveness, technological improvement, and income increase. Attention has also been drawn to the question of the relation of trade with income inequality. Particularly, does integration through international trade have an effect on income inequality, say, through the conduit of wage differential? Do received theories explain this phenomenon and do these explanations fit the data (Bhattacharyya, Chap. 5)? Global integration may of course also take the conduit of production-related value chain that, though, raises some vexed questions in relation to labor and environment. Notwithstanding its other virtues, competition may take place under perverse situations such as weak regulations and norms which in turn may produce pseudo comparative advantage where the structural weakness of the economy is exploited to keep the final price down through transfer of cost enhancing processes to the unregulated informal sector; resultantly, in spite of being intensive in labor the traded good would fail to bring about improvement in the share of wages in the ultimate distribution of factor earnings. If this happens, instead of improving the standard of living and quality of life, this setup may represent a sheer profiteering by the business class without having any positive influence on factor productivity or technical efficiency. Is it then the case that globalization restructures the formal–informal linkage such that unskilled labor and pollution-related activities are exported to the Southern countries (Chattopadhyay and Banerjee, Chap. 15)? This insight may have far reaching consequences. While the mainstream thought would ideally want to see the formalization of the informal sector (ILO sets it as its goal), the mechanics of competitive capitalism working in a global context seem to be permanently positioning the informal sector so as to exploit any advantage from its presence. Therefore, instead of eradicating the informal sector, competitive global capitalism may be cultivating it. This is also the contention of Bhattacharya, Bhattacharyaa and Sanyal (Chap. 14) who, in making this point, are also arguing for a reconceptualization of the concept of an informal sector itself.

One of the most discussed areas in recent times has been the recalcitrant presence of the informal sector which, even with globalization, continues to grow in India. On the one hand, we are witness to a long-standing debate regarding whether this sector can be treated as continuist or dualist, whether the informal sector is independent or autonomous of capitalist formal economy, and whether any linkage has positive or negative effect. On the other hand, the advent of globalization has renewed the debate on the issue of linkages between the two sectors and whether and if so how far the informal economy can be considered as conceptually distinct from the formal economy (Bhattacharya, Bhattacharya and Sanyal, Chap. 14). Making a case for the conceptual autonomy of the informal sector paves the way to not only present it as a disaggregated space of capitalist and non-capitalist enterprises engaging in surplus producing activities, but also in the process examine the disaggregated conditions of existence, reproduction, and

survival of the informal sector. It is interesting in this regard to explore the linkage of the formal and informal sectors under conditions of globalization so as to specify the extent of the independent existence of the informal sector in India. That will throw light on the depth of penetration of globalization in India's industrial economy as also seek to explore whether (whatever maybe the extent of penetration) it is beneficial to the informal sector. The paper argues for a vast presence of an autonomous informal sector and a not so beneficial result for the informal sector following its linkage with the formal sector. In so far as the latter is concerned, this reconfirms Chattopadhyay and Banerjee's (Chap. 15) insight that we had earlier referred to.

The impact of globalization on market structure, of utilities and industry, is important enough to be detailed. Aspects like public utility, infrastructure, education, health, and environment were hitherto considered non-market public goods and thought to be under the purview of the state. An enduring facet of neoliberal globalization is its endeavor to transform public goods into private (Chakrabarti and Dhar, Chap. 2) which entails a new strategy of the state to commoditize these domains by way of public–private partnership. Specifically, a key arena of this partnership is the case for road, seaport, and airport construction which needs to be dissected not only to draw attention to the effects of globalization on these kinds of products, but also highlight the peculiarities and problems of such private–public partnership models. Here the regulatory challenges are multidimensional and the complexity of designing a balanced contract increases manifold due to its integration with the global financial market through the channel of joint venture partnership (Chatterjee and Banerjee, Chap. 9). One of the founding principles of globalization is the centrality of competition which in turn makes the following counterintuitive question interesting: does evidence, say for a globalized commodity such as steel, demonstrate this supposition (Mallick, Chap. 12)? Moreover, hasn't globalization ushered in a technological transformation producing in turn an alteration in the character of extant commodities and their market structures? For example, what about the appearance of wireless technology based on the scarce spectrum that fundamentally altered the character of wired line telecommunication? What does the resultant change in market structure from natural monopoly to oligopoly imply for the telecommunication industry and its regulatory structure (Datta, Sikdar and Chatterjee, Chap. 11)? If the rules of competition dominate the organization of production then the observed input–output relations are expected to express production efficiency. The methodology to be adopted for verification of any such claim has been extensively discussed in Bandyopadhyay and Majumder (Chap. 10) and is illustrated for the garment producing units of India where the presence of local and global market integration is strikingly evident through the chain of sub-contracting, outsourcing, and off-shoring.

Globalization often combines the economic with cultural aspects, which leads to the matter of transformation of the social. One of the standout features of a global competitive market economy is the oft made claim of feminization of the labor force, particularly in the formal sector. The ongoing gender transformation raises the question of the *perspective of woman* as they try to balance home and work.

Interesting is the mechanics of choice that they make and whether it is geared toward maximization of their income or focused more toward a satisficing approach (Dutta and Husain, Chap. 13). If the latter, as is argued in the paper, then it indicates a departure from the standard supposition of neoliberal *homoeconomicus* underlying neoclassical economics thereby calling for a revision and a rethinking of the conventional idea of optimization as the basis of decision making even on supposedly economic matters.

Given the vast population surviving on agriculture, its importance in reducing absolute poverty through agricultural growth, the looming food security problem, and changing nature of interconnectedness of industry with agriculture, it becomes vital to investigate the distinct affects of liberalization and globalization on Indian agriculture and its sustainability in terms of yield, cropping pattern, the influence of market integration on farmers' production decision, etc. This exercise would be particularly insightful as Indian agriculture itself is undergoing massive transformation following a paradigm shift toward a market-oriented approach that is steered by the decontrol of domestic barriers and opening up of external markets to domestic produce. This transformative process of commercialization of Indian agriculture following liberalization and globalization though is fraught with the possibility of systemic instability that not only threatens to leave the agrarian economy in a dire situation, but also may undercut the major objectives of development. The destabilization outlook comes into prominence in the backdrop of relatively weak performance of the agricultural sector, particularly concerning its production. Despite inter-state differences in production performance, Indian agriculture overall is showing signs of faltering growth rates in food grain, non-food grain, and all crop production, feeble expansion of crop intensity and productivity as also downward trend of growth rates of various inputs; problems of land and credit availability, high input prices and withdrawal of the state, besets agriculture (Bhaumik and Rashid, Chap. 6). Along with these, attention must also be drawn to aspects of pricing and marketing, more specifically, of imperfect agricultural markets and price discovery mechanism that reveals the unpreparedness of Indian agriculture to embrace the market centric approach; these distortions amidst a market centric approach is seen as responsible for the unsatisfactory post-liberalization performance of agriculture production and farmers' distress (Deshpande, Chap. 7). This comes against the backdrop of the embedded problems of federalism (agriculture being both union and state subject making uniform policy impossible), liberalization, and globalization impacting agriculture especially through the pricing mechanism (both product and cost) and the inept role of the state in implementing measures such as minimum support price. The importance of attending to issues of state would have to encompass, in addition to its increasingly restricted physical presence, the facet of quality of service as well; the search for alternative institution building also can no longer be discounted. All these need to be put under critical lens and new avenues found to put this sector back on track, failing which the systemic instability that is economic in nature may get coupled with political instability that, if it transpires, would slow down if not derail the proclaimed development objective.

One of the outstanding features of the transformation of the Indian economy and its policies has been the acknowledgement of the above described contrasts, the problems of sustainability inherent in them, and the need to account for these in state-sponsored policies. This maximalist position of the state in relation to the problems of sustainability comes against the background of a minimal understanding of state vis-a-vis competitive market economy (Chakrabarti and Dhar, Chap. 2). The burgeoning presence of the state in countering the new scars of sustainability has led to a slew of policies. This situation becomes particularly acute in the case of a perspective that sees growth as exclusionary leaving in its wake a landscape of exclusions—structural, economic, and social—that it cannot but now account for. An important aspect of our book is an interrogation of these exclusions and their underlying sustainability problems, the way the state deals with these issues and the limitations of this dealing, both theoretical as well as in relation to its implementation. Prominent among these interventions is the microcredit approach that seeks to tie credit (financial sector) with productive economy to fulfill the goal of poverty reduction; the analytics of this approach, its mechanisms, forms, and outcomes need serious attention. Moreover, how far the Indian scenario of microcredit is distinct from all other global experiments on it and how do these distinct features fare in fulfilling its defined objectives constitute important points of discussion. A vital question crops up in this regard: is the microcredit approach a medicine for alleviating poverty or one of coping with poverty (Banerjee, Chap. 17)? Two more interventions by the Indian state to address exclusions are the cases of MGNREGA (Sarkhel, Chap. 18) and women's empowerment (Ghosh, Chap. 8). As in case of microcredit, both are distinguished by their attempt to influence the production economies of the poor and marginalized, and that too in a manner which will bring an end to their poverty and eradicate their exclusionary existence. Similarly, the above mentioned question that Banerjee (Chap. 17) asked of the Indian microcredit approach acquires importance for the case of MGNREGA and woman's empowerment too. This, in turn, raises questions regarding the success of the interventions as also finding ways to overcome those problems, both important considerations in Sarkhel and Ghosh. Moreover, women's empowerment is a vexed issue that cannot be reduced to categories such as autonomy. This is compounded by the problem that if one is to get to the measurement of empowerment then working with broad indicators and units (including household) will not do; there must be some way to connect empowerment to the level of the individual which evidently is a methodological problem and requires knowledge of new techniques (with the challenge of modeling latent variables) (Bhattacharya, Banerjee and Bose, Chap. 16). Finally, beyond the known exclusionary indicators, the evolving capitalist market economy can throw up many seemingly invisible processes of exclusion of its own that in turn can have detrimental effect on the poor. Important in this context is the phenomenon of how the gainers of India's high growth rate regime, the new rich, use their financial power to create private institutions (call these provisionally as clubs). This leads to the privatization of the hitherto publicly provided merit goods in their favor and in doing so put pressure on the state to cut down on the public

provisioning of the good itself and negatively affect the quality of the good meant especially for the poor (Bhattacharya, Chap. 19). This somewhat devious phenomenon shows the paradox of how modern institutional building in the post-liberalization era, with all its paraphernalia of efficient services, can be the breeding ground for institutionalized social exclusion (these clubs rule out certain kinds of people), pro-rich influence on state policies, and cause of decline in the quantity and quality of public provisioning for the poor. In short, such exclusionary institutions may produce optimal public policy response that is harmful for the poor, thereby aggravating an already pressurized public service and in the process deepening public discontent. Indeed, if such kinds of exclusions are combined with known forms of exclusion, then we do have a potent problem on our hands.

Two areas that are integral to the neoliberal turn are health and education. As the competitive market economy penetrates the extant 'non-market' and 'non-economic' domains so as to bring these under its dictum, these two sectors are at the forefront of a transformative exercise that is unparalleled in scope and quality. In health, two such importantly new areas that are attracting attention are the phenomenon of mental health (Dhar, Chakrabarti and Banerjee, Chap. 25) and of HIV/AIDS (Dutta, Chap. 24), that is, the space of dis-eased psyche and body. Interestingly, discourses around these are part of a global knowledge, engagement, and policy of how to deal with these phenomena. Similarly, referring to the neoliberal take on higher education, of privatizing it, as well as transforming its content, questions emerge regarding the limitation of market determined education (Chattopadhyay and Bhattacharyya, Chap. 23). This in turn brings into focus once again the utility of the state as a provider of education. The interventions with regard to the sectors of health and higher education not only point to the neoliberal turn in viewing the mind, dis-eased body and psyche, but also bring to light the limitations and problems of such an approach. This is not to say that these studies do not have specificities. Mental health in India is a huge canvass connecting the agents with the state and with institutions which are private and public, modern and traditional, thereby forming a distinct political economy; in addition to the global rendition of mental health producing its own institutions and methods of 'treatment', Indian peculiarities also generate Indian ways of comprehending and dealing with these specificities and that includes many examples of institutional building that challenge the dominant global knowledge and method of dealing with the problem of the psyche. The HIV/AIDS epidemic is marked by the production of a global discourse that by what it foregrounds as essential for controlling the epidemic rules out and/or demotes certain other essential aspects tuned to care and treatment; particularly disturbing for India is the emphasis on behavioralism procured through the mechanics of awareness (huge funds are forthcoming in this regard) as part of a global discourse which in turn has demoted the endogenous determinants of HIV/AIDS targeting care and treatment that necessitates strong institutional intervention (to address problems of absence of access to health care, lack of nutrition, and so on). Finally, education in India is marked by the growing tilt toward neoliberal commodification that is not only

fraught with the dangers of worsening social exclusion, but also of delinking the valued relation between education and society; the almost unregulated privatization of higher education in India along with the privatizing of knowledge is also a matter of concern.

Even before initiation of the debate around socio-economic acceptance of global market integration, designing of market instruments to ensure efficient management of natural resources and control of environmental pollution was well documented in the literature. Thus, any volume on globalization and sustainability will remain incomplete without mentioning the location of these issues in the new economic cartography. Instead of identifying newer issues here, emphasis has been given on developing context-specific appropriate methodology for impact assessment. Three different types of methods have been illustrated in three papers (Chaps. 20, 21, and 22). For example, it is shown that if the common property natural resources are not efficiently and equitably managed then that would lead to asymmetric power relation among the local people and consequently social inequalities in the form of vulnerability of women and poor. That would culminate into a negative attitude toward resource conservation. Ray and Bhattacharyya (Chap. 20) addresses the issues centered on assessing the influence of attitude on behavior and highlighted the need to borrow methods from other related disciplines like psychology and sociology to come up with more comprehensive evaluation techniques. The importance of field experiments has been discussed in terms of specific studies on the joint forest management in the state of West Bengal. While the problem of resource management was discussed in a rural setup that of pollution management has been posited in an urban context. One example has been taken from the household sector and another from the industrial units. In case of the former the issue of exposure to indoor air pollution and its effect on morbidity has been extensively discussed by using unit level NSS data. The importance of data exploration in formulating and statistically testing any research hypothesis has been illustrated and in this process the method of construction of a pollution exposure index has been suggested to generate a credible exposure-morbidity mapping (Dutta and Banerjee, Chap. 21). For the efficient management of industrial effluent the relevance of all insights derived from the solution to cooperative game has been discussed and the shapely value solution has been worked out for the common effluent treatment plant of Calcutta Leather Complex (Bagchi and Banerjee, Chap. 22). It has been shown that along with technical feasibility importance should be given to economic viability of the arrangement to make it self-financed and to reduce the burden on public subsidy. Here, efficiency would supplement sustainability as well as ensure a better common future.

The Exceptionalities of the Book

Both in terms of what it covers and how it covers, this book stands out for its exceptionality. At least, four such exceptionalities must be mentioned.

First, one feature of our intervention is the diverse range of topics that are covered. While globalization and sustainability are the broad entry points that attend to the new situations, schisms and contrasts we were referring to, the topics we crisscross to highlight our theme are multi-faceted, encompassing aspects such as state and its policies, neoliberalism, capitalism, stock and currency market, industry, agriculture, services, labor, health and education, gender, and environment, to just name a few. The idea is to present the 'Indian economy' per se as a research problem and that too in all its richness.

Second, our collected essays for the book digs deep into this newly produced Indian economic map. It distills how and in what manner this map has been fundamentally redrawn paving the way from a new set of problems that cannot be compared to anything existing prior to 1990. This realization comes against the backdrop of a growing disquiet about the inability in the existing literature to fathom that an explanation of this new Indian economy and its problems presents a methodological challenge. That is, engagement on issues in the rise, reproduction, crises, and resolution mechanism (including policy matters) of the new Indian economy must involve a distinct lens of analysis as compared to previous historical instances. Many of the essays in this book take up that challenge at the level of epistemology, qualitative and quantitative; this paves the way for a new corpus of concepts as also methods of analysis. Moreover, an attempt is made to rethink the rationale and character of old and new institutions as also rules of engagement, many of which are dissected in this book. In other words, taken as a whole, this book is distinguished by its effort to produce a distinct set of concepts, methods, and focus of analysis so as to make sense of the evolution of the Indian economy in the last two decades. By acceding to a need for methodological self-introspection, it engages with and challenges not only the existing discourses of Indian economy, but in the process also offer new insights into comprehending much of the underlying nuances and contradictions pertaining to the Indian economy.

Third, along with theoretical analytics we have tried to develop an appropriate quantitative methodology to handle different issues related to the question of sustainability in the backdrop of global integration and as far as practicable large secondary data sets have been utilized to illustrate these methods. For example, in Chap. 10 the techniques of assessment of production efficiency have been illustrated with the unit level ASI (annual survey of industries) data, in Chap. 16 a methodology has been developed to assess woman empowerment at the individual level and illustrated in terms of NFHS-3 data base, in Chap. 21 the morbidity effect of indoor air pollution is studied by using NSSO data where the unit level data have been utilized to illustrate different methods of data exploration, and so on. In doing so one has to make use of different types of statistical software like STATA, SPSS, DEAP, LISREL, etc. So, some illustrations of package-outputs are also appended with the relevant chapters.

Fourth, there is much to value on how the topics are covered in this book. That is, the methods range from neoclassical, Marxism, to even institutional approaches. It follows from our self-conscious intent to take a heterodox approach. To begin with, we consider it apt to present the new economic map in all its

underlying complexities and diversities by way of approaches that are varied in theoretical orientation; our understanding of Indian economy, rather than getting constricted, is enriched as readers come across a variety of theories and insights, some of which may even be clashing. This is furthermore complemented by our belief in the rich possibility and value of the heterodox approach per se; rather than being avoided or cocooned into distinct schools, there is perhaps an intrinsic value in the process of sharing positions and debates. We sincerely hope that our book will emerge as an example of rigorous heterodoxy, an experiment of what a study on Indian economy can be if we allow ourselves to open our self-imposed closets.

The Structure

Given the themes of globalization and sustainability, we organize the book over a number of sub-themes. Our starting point is the location of the state (Chap. 2) followed by global integration (Chaps 3, 4 and 5). To discuss the issues related to growth, stability, and sustainability of the economic sector we have started with agriculture (Chaps 6, 7 and 8) and placed the discussion on infrastructure and industry in the subsequent sections (Chaps 9, 10, 11, 12, 13, 14 and 15). The problem of social security and social exclusion has been discussed next (Chaps. 16, 17, 18 and 19). The specific issues related to social sector, its integration with economic sector, and effect on sustainability has been placed in two following sub-themes under the heading of Resource and Environment (Chaps. 20, 21 and 22) and Education and Health (Chaps. 23, 24 and 25). We started by locating state within the triad of neo-liberalism, global capitalism, and inclusive development and ended the volume with a discussion of mental health to show how the comprehensive effect of this triad is not contained within the economic domain. It is influencing the psyche of the 'economic man' and therefore, the process of globalization is really invasive with deep and long lasting impact on all spheres: economic, social, cultural, and psychological; this in turn causes a further churning in the idea of sustainability that we believe is yet to be acknowledged at the level of the discourse of Indian economy. Knowing very well that no single project can be sufficiently comprehensive in handling such a complex polygonal sphere in totality, we are nevertheless making a case for situating and analyzing the Indian economy within this new Order of Things that globalization and sustainability has opened up. In doing so, our endeavor is to conceptually displace the body of Indian economy into a new terrain, thereby fundamentally relocating and reorienting the discourse of Indian economy as an object of analysis.

Chapter 2
Rethinking and Theorizing the Indian State in the Context of New Economic Map

Anjan Chakrabarti and Anup Dhar

Introduction

In this chapter, we argue that the rationale for the existence of the Indian state has undergone a fundamental displacement since the adoption of the New Economic Policy (NEP henceforth). This displacement is in alignment with the re-articulation of the Indian economic cartograph into the mutually constitutive triad of neo-liberalism, global capitalism and inclusive development that form the Order of Things (*a la* Foucault) at present. India's economic transition now must take this historical conjecture as its point of reference and departure. In this regard, the theory of the Indian state must contend with (i) why and how it helped create this triad, (ii) what this triad entailed for its own existence and (iii) how the state encounters and negotiates with new-fangled contradictions emanating from the triad and thereby transforms itself or gets transformed in the process. State as a transformative entity must thus be understood in relation to the new Order of Things in whose creation it plays a central role and which in turn affects it. There is one important thread that runs through the changes that has materialized with respect to Indian state in the last two decades: the rationale for the existence of Indian state has changed which implies in turn a different philosophy of governance.

For convenience sake, the analysis of Indian state in the post-liberalization era can be divided into three parts: 1991–2004, when the template of neoliberal globalization in tandem with global capitalism was rapidly gaining ground along with the belief that growth is sufficient to reduce absolute poverty and end social exclusion, 2004 to perhaps the continuing present, when it was realized that

A. Chakrabarti (✉)
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: chakanjana@yahoo.co.in

A. Dhar
School of Human Studies, Ambedkar University, Kashmere Gate Campus,
Lothian Road, New Delhi 110006, India
e-mail: anup@aud.ac.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_2,

growth in the presence of growing income inequality and persistent social exclusion was not sufficient to reduce poverty and marginalization entailing that neoliberal globalization and global capitalism must be blended with inclusive development, and 2007 to perhaps again the continuing present, when the combination of global economic meltdown and a frail Indian economy signaling a crisis in neoliberalism and global capitalism is producing new arenas of discord, not only in practice but also at the level of thought. Each stage is marked by a shift in the rationale of the state. When required, this paper will activate this periodization to highlight some aspects of the shifting rationale. This shifting rationale, in turn, points to one phenomenon that we want to underline: the continual attempt to balance the priority of high growth propelled by the expansion of global competitive market economy with the negative effect of un-sustainability (political, economic, social, and ecological) resulting from it. Political un-sustainability pertains to frictions taking the form of instability, resistance, violence, and repression arising from the expansion of global capitalism, economic un-sustainability refers to possibilities or events of breakdowns emanating internally from economic structure, to the divide arising from income inequality and not so satisfactory reduction in poverty rate,[1] social un-sustainability points to a growing social divide, encompassing class, ethnic, caste, gender, and age-based exclusions and contrasts, and ecological un-sustainability highlights the friction, displacement and breakdown of ecological systems and values. We shall examine the changing rationale to comprehend the continual effort of the Indian state to re-locate itself between the seemingly conflicting domains of growth and sustainability, so as to ensure the process of a relatively smooth transition toward capitalist development.

Our chapter is focused mainly on four aspects: (i) the nature of the re-location of Indian economy in terms of the triad: globalization under neoliberal conditions, global capitalism, and inclusive development, (ii) rationale for the continued presence of the Indian state in this changed scenario, (iii) elaboration on some of its policies based on the rationale and their changes over time, especially when confronted with the spectre of sustainability, and (iv) the limits of its rationale-policy regime marking the inauguration of coercion/repression/violence as a mode of engagement in those limiting instances. These four aspects constitute for us a *partial theorization of Indian state*.[2]

---

[1] Our treatment of financial sustainability (a subset of economic sustainability) here is sparse. It is particularly taken up in details by Mandal and Kar (Chap. 3) and Sarkar, Chakrabarti and Sen (Chap. 4). Our take on ecological sustainability is similarly muted.

[2] We desist from reading the evolution of Indian state as mirroring known approaches of state: classical, liberal, libertarian, Fabian, and Marxist. Nor do we attempt to read Indian state in relation to specific theories of state such as neoclassical, public choice, transaction cost, information theoretic, structuralist, postmodern, and so on. Rather, our strategy is to make sense of the practice of Indian state in relation to the context which arises from the new situations and through that encounter try to derive a theory of the Indian state. This is not to say that the known theoretical interventions played no role in influencing the rationale of the Indian state or even that they cannot be used to explain certain features of this rationale; when required we relate our analysis to them as well. However, the essential strategy is to intervene in the context of the new

Before delving into our problem, let us chart out the points of reference and departure pertaining to the pre-NEP period of centralized planning. This is to contextualize the passage which constitutes the transition of Indian state from its centralized planning role to that in the post-NEP period.

Some Features of Indian State in the Planning Era: Marking Points of Departure

In its effort to operationalize a developmental model of modernization based on capital accumulation, Indian state during the centralized plan era formed, among others, three associations with the economy: control of the market economy, welfarism through invocation of public goods, and active policy management (fiscal, monetary, and currency).

State control of market economy extended far beyond restrictions on land, labor, and money, so that for its survival, the market economy was ultimately made dependent on the state. Exercising control over market economy entailed the state occupying commanding heights of the economy, including acquiring virtual monopoly in heavy capital goods, minerals, and formal finance; through reservation policy, industrial licensing system and import substitution policy this control was further extended. This secured its effective control over allocation of resources and hence over the private sector. Both supply and demand management of the private market economy by the state were considered essential to fulfill the goal of economic growth through capital accumulation. The state did not govern *for* market economy. Literally, it governed market economy; any distinction between regulation and control blurs in this case. Not surprisingly, competition and free trade were non-starters under such a regime.

Stemming from the professed inability of the markets to form due to standard arguments of nonrivalry, non-excludability or even coordination failure, the rationale of public goods enabled the state to further entrench itself in the economy. Claiming itself as the exclusive legitimate player in sectors such as health, education, and infrastructure on account of these being public goods, the state took effective control over 'social sectors'. This, in turn, prevented the market economy from making any substantial inroads in those sectors.

Finally, both active fiscal, monetary and exchange rate policy founded on Keynesian arguments or otherwise were considered acceptable and indeed actively practiced. Whether it was the matter of growth, inflation or trade, the state not only

(Footnote 2 continued)
Order of Things in which the Indian state finds itself embedded (both as one of its producers as also being a site of their effects); our strategy will hopefully help the readers see the path-dependent nature of the transition of Indian state as well as the complexity which surrounds its particular existence. Indian state and its transformation is certainly no mirror of any theory, but must be seen/theorized from within the historical juncture and process of the evolution of Indian economy, its breaks, departures, and transformation.

fixed the objectives but was instrumental in driving these in ways it deemed necessary. Resultantly, the Indian state was seen as actively intervening in the market economy through its macroeconomic policy regime.

The Indian state then controlled the market economy, actively participated in it, manipulated it and had the authority to intervene in the market economy. What sustained this kind of state were three underlying beliefs. The first was that the state—as sovereign—can master the totality of the economy; assuming the economy as finite fueled the faith that state sponsored economic engineering is feasible. The second pertains to the denial of the self-regulating character of market thereby legitimizing the rationalization of state control of market economy; efficiency argument was demoted in favour of state sponsored allocation of resources which was argued as essential for capital accumulation induced growth strategy (2nd five-year plan) and later for redistributive policy (5th five-year plan) (Chakravarty 1987). Finally, underlying the rationale of the state was another belief that individuals/agents are akin to children, that is, incapable of forming independent decisions and undertaking actions that are beneficial for them and for the society at large. Instead, it is the state which must rule on behalf of the citizen subjects. This is in opposition to the liberal supposition of independent and autonomous individuals as capable of taking decisions and actions (Hayek 1944; Friedman 1962). How state paternalism and liberal autonomy of the citizen subject got reconciled in the imagination of the Indian state is not very clear. One argument could be that in the sphere of the economy the fundamental premise was one of paternalism and in the political sphere the fundamental premise was one of autonomy; after all, state sponsored development and democracy ran a parallel course representing the peculiarity of Indian transitional route. These three beliefs were fundamental to the very imagination of centralized planning system in India. Moreover, this entailed that the state's technology of power came to be based on the principle of 'more governance is good governance'.

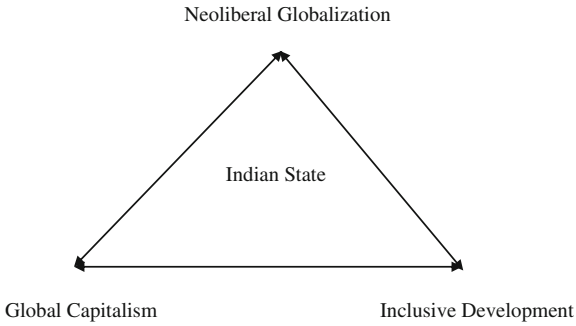
As we shall see, the new Order of Things and particularly neoliberalism fundamentally reconstituted the economic–political philosophy and challenged these three beliefs so as to displace the underlying rationale of the Indian state.[3]

The New Order of Things

What is new about India's economic transition? *Nationalist discourse based on self-sufficiency giving way to neo-liberal globalization, national capitalism*

[3] We do not offer any normative judgment on the rationale and role of Indian state. This is not to ignore normative issues that find entry into the state's evaluative space and practice, but here that never translates into a question regarding the veracity of those positions which, if it is to be undertaken, would involve a full scale critical analysis of them. Therefore, so far as normative issues are concerned, they find its presence only in the context of our analysis of the Indian state's rationalization of its practice.

increasingly becoming global and development going inclusive are unraveled as constituting a unique historic triad in the Indian economy. The Indian state must be re-located within this new Order of Things, both as a producer of those conditions and as being constituted by their effects.



Let us explore the role each played in reshaping the rationale and practice of Indian state. Since the philosophy of neoliberalism underlying the globalization process is a new phenomenon in the Indian economic and social life, it will be worthwhile to spare comparatively more space to it.

Neoliberalism: State and Counter Revolution

Harvey (2005) traces the advent of the Anglo–American version of neoliberalism to the interventions of Margaret Thatcher, Paul Volcker, Ronald Reagan, and Deng Xiaoping. Foucault (2008) studied various kinds of neoliberalism: Anglo–American, Ordoliberal, French, and so on. Because of the distinct historical conditions in which different countries found themselves at the moment of embracing neoliberalism, the entry point to neoliberalism differed and accordingly the resultant trajectories took varied path dependent routes.[4] Notwithstanding their diverse origins, it was the Anglo–American version that achieved relative dominance in the globalized post-Soviet era. With the help of international institutions such as IMF, World Bank, WTO, etc., this model was successfully advertised and exported to the rest of the world through a series of structural adjustment programs (SAP) designed to create conditions for a competitive market economy. India was no exception, as the appearance of the NEP in 1991 indicated. In addition to attacking planned economies, the neoliberal model also dismantled in the course of its journey so-called welfare capitalism and severely dented the political model of social democracy. In that regard, neoliberalism must not be merely seen as growth

[4] In elaborating upon neoliberalism as a historical juncture, we do not harp much on the issue of the distinction between liberalism and neoliberalism. Evidently, many features of neoliberalism are an extension of the broader liberal ideas while in some cases it adds something distinctly new. Even as the spotlight is on neoliberalism, we hope that both the similarities and differences with liberalism will be identified by the readers.

facilitating, but also as instrumental in changing the extant Order of Things. It is this connection of neoliberalism to systemic transformation that particularly animates us. Moreover, it is not just the change in structure, but the alteration in the process of subjection–subjectivation (or what Butler [1997, 83] calls *assujetissement*), that makes neoliberalism potent. Our intervention below on neoliberalism is restricted to what is of utility to our understanding of the transition of Indian economy, and specifically the changing rationale of the Indian state.

We start by arguing that one of the thrusts of neoliberalism has been its critical reflection on governmental practice. It professes and represents a formula and mentality of rule based on a critical look at the principle that "one always governs too much." (Foucault 2008, 318–319) To sharpen this element of doubt, we can go back to the planning era when in line with the embedded rationality that 'more governance is good governance', the dominant principle of the state was *governance for itself* and by default *governance of society*. Since its legitimacy and strength lay in its own expansion, the state was inclined to maximize its control of spaces and populations. In line with this rationalization, the state's domain and object of policing the economy was absolute, guided by the belief that it was possible to control the allocation of resources and rule the economy on behalf of the population. To a greater or lesser extent, the state under centralized planning and welfare capitalism adhered to this principle of governance.

Under neoliberalism, this rationality was rendered suspect against the backdrop of this question: "why must one govern?…What makes government necessary, and what ends must it pursue with regard to society in order to justify its own existence?" (Foucault 2008, 319) As we explain subsequently, neoliberalism seeks to overturn the rationale of the existence of the state, namely, that *the state must exist to govern for society and not for itself*. Rather than consider the efficiency and optimality of state intervention, the central issue is whether it is at all necessary, and hence legitimate for the state to arbitrate in society. The state, in turn, embodies a self-limiting principle: *self-regulation* of governmental practice and a *self-critical* approach toward governmental practice (driven by the question as to whether the state is governing *too much*). Under neoliberalism, this shift in the idea of state becomes a general formula for the *idea* of governance across society producing in turn the principle 'less governance is good governance'.

A question however persists. If state cannot directly govern society in the sense of running and controlling it, what institution and mechanism will constitute society? What does it then mean to say that the state will govern for, and on behalf of, that society? Answering these questions require emphasizing the link between society and market principle that is highlighted by neoliberalism. The linkage arises because neoliberalism is a "political rationality that tries to render the social domain economic" (Lemke 2001, 203), economic in the sense of society being modeled on market principles, that is, to use the Smithian phrase, on 'truck, barter and exchange' and that too operationalized through the mechanism of competition. The independence of competitive market economy and of human agents interacting in it (the homoeconomicus) in shaping society and its outcomes must be, at least in principle, secured independent of the state as far as possible. Both

competitive market economy and mass of independent and autonomous homo-economicus (economic man), each constituting the other, must be taken as capable of performing not merely economic function (so as to generate efficient outcome), but also social function (generate social harmony/order which previously was the prerogative of the sovereign state). This combined existence of liberty principle epitomized by free individuals and efficient economic outcome symbolized by maximum possible output has long been defended as the best form of society (Hayek 1944; Friedman 1962). In searching for minimum politics in a market order, James Buchanan avers:

> The logic of the structure fully confirms that only a market economy is capable of allowing for a minimal politicization of the national economy…And only through such minimization of politicization-bureaucratization (or at least through some reduction) could meaningful individually based social objectives be secured, whatever these objectives might be…Such an economy produces a larger bundle of goods and services, as measured by the evaluations of persons who consume them. The economy organized on market principles produces more value than an economy organized on nonmarket principles (Buchanan 1991, 1–2).

To have this society though, one requires a new conception of state whose nature and scope of intervention is now vastly different, since its goal is altered to enable the agents to compete and trade freely and fairly in the market economy and protect the resultant outcomes (gains, losses) from the spectre of expropriation or threat; private property rights and legal structures underlying these become fundamental. Moreover, it is the claim of neoliberal macroeconomics that if such an economy is allowed to be operational, then the process of maximization of individual welfares will also generate greater growth as compared to other social arrangements (such as an active state controlling and directing the economy as in case of the planning period in India). Putting the same somewhat differently, if the state's objective is to guarantee maximum economic growth and welfare of individuals (that constitutes governing *for* the economic good of society), then the state should accept the idea of individual liberty and by default give up on its sovereign character. Rather than trying to the good society qua the competitive market economy, the sovereign character should be reserved for creating and then preserving this good society from internal and external threats. Individual liberty and economic gain combine to produce a powerful trope that is neoliberalism; this trope has in the last two decades fundamentally altered the rationale of the Indian state. How far the Indian state has traveled in the last two decades in rethinking the basis of its existence can be gleaned from a recent quote of Economic Survey (the mouthpiece of Indian state):

> In many poor nations the Government takes the stance that, when in doubt about the goodness or badness of two or more adults voluntarily conducting an exchange, stop them. An enabling state, on the contrary, takes the view that, when in doubt, do not interfere. There are, of course, many actions of individuals and groups that will need to be stopped for the welfare of society at large. But the default option of an enabling state is to allow rather than stop, to permit instead of prevent. This altered conception of the state can have

> dramatic effect on the functioning of an economy, in general by promoting greater efficiency and higher productivity (Economic Survey 2009–2010, 24).

The above is a recognition of the displacement in the technology of power from the principle of 'more governance is good governance' to 'less governance is good governance.' Having explained one of the rationales of the Indian state as being neoliberal, at this juncture, it will not be too out of place to discuss the potent outcomes of a competitive market economy and free individual agents, ask how do these combine to produce the phenomenon of 'economization of the social'. In the course of the discussion, it will become clear that the changed rationale of state is inalienably tied with this alteration in the Order of Things. Placing India's economic reforms in this background will enable us to sharply distinguish the exact role of the Indian state, its statecraft and make sense of the changing policy paradigm.

Let us first examine the structural changes sought in neoliberalism, which, with few exceptions, have been integrated in the post-planning Indian policy paradigm. At the base is the belief that markets are optimal and self-regulating social structures, which, if allowed to work freely, would "optimally serve all economic needs, efficiently utilize all economic resources, and automatically generate full employment for all persons who truly wish to work" (Shaikh 2004).

Second, any constraints put before the functioning of the mechanism of competition identified typically as state, trade unions, 'third world' social structures, and trade barriers should be dismantled, curtailed or controlled, so as to allow markets to function in a competitive environment. In other words, the supremacy of competition must be internalized in the functioning of market; competitive market regime is thus the key. Naturally, the economic rationale of neoliberalism integrates the central claim of classical trade theory that argues that free trade will automatically benefit all nations by channeling the engine of competition in order to exploit comparative advantage, produce economic development, and eradicate poverty (Bhagwati 2002; Shaikh 2004).[5] The formation and growing importance of WTO and regional trading blocs as facilitators of global integration lends credence to this conviction. This, almost universal faith in the mechanism of competition to pick out winners and losers, is supposed to be a historic departure from previous instances when market-related conflicts were resolved through wars (colonialism, world wars, coups in countries, occupation, etc.). In the context of India's economic reforms, the supremacy of market and the medium of competition to guide it are evident in the transformation of the economy in the last two decades. This conviction is marked in the liberalization of financial markets (stocks, insurance, banking, pension, etc.), in partial labor market reforms (encouraging contract labor, curtailing severely the power of the unions and

[5] In a contrary position, Bardhan (2007) suggests that the sharpest ever reduction in absolute poverty happened in China which was due to neither industrialization nor globalization, but fundamentally a result of agricultural growth in the 1980s following de-collectivization and equitable land reform.

allowing capitalists greater leeway to hire and fire as also have more say in relocation, downsizing and closure; although seen as necessary, full implementation of labor law reform is still pending), in ensuring capital and current account convertibility as also in removal of trade barriers (allowing the financial funds along with goods and services to flow freely across the globe) and so on; many such changes have already transpired and some are in the avenue. The deep-seated skepticism about state's own role in the competitive market economy is encapsulated in the spate of privatization and disinvestment since 1991, the self-critical gaze on fiscal policy (captured by the effort to put in place the Fiscal Responsibility Budget Management Act) and monetary policy (with the focus on inflation targeting, de-emphasis on its role in growth management and increasing decontrol of interest rates that were previously administered).

Finally, neo-liberalism privileges the freedom of decision makers—producers and consumers—to choose-act without interference, so as to fulfill their respective objectives (utility, profit and so on); the earlier quote in Economic Survey is indicative of state's commitment to this principle. Modern macroeconomics has not only highlighted the aspect of maximum possible growth and individual welfare from such an arrangement, but also argued that it will help create a depression free economy (Lucas 2003); this once again put to sword the possibility of an active state as an agent of demand management, monetary intervention, and stabilization. If the state does intervene with respect to these which it has historically done by treating individuals as passive (the famous invariant principle of government policy criticized in Lucas [1976]), then unpredictable and inefficient outcomes will transpire since in reality individuals are not passive but active decision makers who will internalize any changes in government policies, and therefore whose resultant responses may not tally with what is forecasted by the state (Lucas 1976; Sargent and Wallace 1976); moreover, if the state violates its self-proclaimed policy rule so as to intervene, still inferior outcomes will materialize (Kydland and Prescott 1977). If individual behaves in a systematic manner (that is, rationally) which is the basic precept of homoeconomicus then direct state intervention in a private economy, so as to change output and employment is futile and hence needs to be abandoned. This critique acquires significance when a contrasting claim is made that the combination of individual liberty (personified by homoeconomicus) and facilitating structures in which this liberty could be exercised (competitive market economy in a global scenario) produce the following alternative scenario that we have already described: maximum possible economic growth for society and best possible individual welfare. Under this latter arrangement, independence of enterprises including the high value generating big corporations which, in the era of globalization, are essentially global capitalist enterprises is considered essential; these are deemed as legitimate market players—free economic agents—as long they adhere to the neoliberal principles, including its laws and regulations. Correspondingly, neoliberalism has been associated with centralization of the creation, protection, and expansion of global capital—productive and financial; particularly potent has been the rise of financial capital and its influence over productive capital (shareholder capitalism as well

financial control of projects), although the extent of the influence of such an order in India is debatable. Moreover, the expansion of the economy centralized around the global capitalist enterprises are supposed to bring in its wake superior product quality, lower price, higher investment, better innovation, continual technological development, and increased employment. Expansion of global capitalist enterprises is thus deemed as good for developing countries. Finally, crucial for the success of this neoliberal arrangement is the global reach of the Rule of Law (invoking Hayek) to legalize foreign income, profits, assets, innovations, property in general, all seen as fruits of entrepreneurial activities, and ensuring that owners are protected from threats of (national) expropriation, unfair practices and social instability; in other words, the Rule of Law must be fundamentally geared to secure the liberty of economic agents, and consequently must expel any possibility of arbitrary behavior which would unmask the state's sovereign character (Hayek 1944). Correspondingly, shifts in the role of state, and that includes India, have been marked by creation of a new national legal structure to facilitate these features and in certain cases subsuming it under the global rule of Law (say, through the WTO) if that is required.

It is evident that the above described structural transformation would involve a dramatic alteration in the conception of individual. Let us for the moment reflect on it. The thesis that a competitive market economy would demand a homoeconomicus is an old liberal adage; what stands out under neoliberalism is the transformation in the concept of homoeconomicus itself. This homoeconomicus is no longer a mere man of exchange (that is, not a mere *laisse faire* personality), but also one who (i) personifies human capital a la ability machines; who is in turn an "entrepreneur of himself, being for himself his own capital, being for himself his producer, being for himself the source of (his) earnings," (Foucault 2008, 226), and (ii) capable of rational conduct guided by the capability to weigh any situations through cost-benefit and maximize welfare in a dynamic field of competing ends (Schultz 1960, 1961, 1962; Becker 1973, 1993). Accepting that such an innate quality exists in humans and hence is a pre-given potential encased in them, Becker and co would also argue for the need to cultivate, produce, and make visible this homoeconomicus or rational ability machines. Appropriate education and training, norms and values, etc., are considered critical, both for creating a rational agent and for shaping the person's future productivity and income; to this end, the role of family too can no longer be kept outside of economic discourse. In short, they argue for both subjection–subjectivation if the structural changes are to deliver the good results. If the individuals are not to be treated as docile, that is, not under the control of the state, then they must be created as capable of functioning in a competitive market economy in a manner that would help reproduce and transform the latter without the direct intervention of the state. The redirection of the concept of homoeconomicus, in combination with the structure of competitive market economy, has had a seismic transformative effect.

To begin with, along with the acknowledged importance of family as an economic institution, it is evident that this will require a wholesale transformation of the education system. The goal of education must transit into the production of

neoliberal homoeconomicus, capable of self-governance through rational conduct, and producing economic value through cultivation of human capital, two features considered indispensable to fit the individual into a competitive market economy. In other words, the goal of education must not be *'atman nirman' or* even what Radhakrishnan identified as "training for Citizenship." Rather the objective here is to create rational subjects that is, self-governing, saleable competitive selves for the market, and not 'developmental subjects' under the disposal of the state as was the case during the planning period. The entire education system needs drastic reorganization to make such subjects possible. A classic appreciation of this objective is the Birla–Ambani Report (2000), where it was recommended that ideally the state should only be restricted to primary schools and the rest of education be transferred to the private sector (also see Nilekani 2006; for a critique, see Kumar 2008). Calling for a fundamental change in Indian mindset of seeing education as a component of social development, it argues that the goal of education should be to shape adaptable, competitive workers with technical, managerial skills, and as deliverer of services, i.e., "ability machines" who can govern themselves and are also governable. This is a paradigmatic shift, as education is no longer to have any "intrinsic merit"; the only relevant measure is that of "economic value." Put bluntly, Indian needs a "… a revolution in education that embraces information technology, fosters freedom and innovation and induces a market oriented competitive environment" (Birla 2000; preface). This is supposed to facilitate a change in the constitution of the idea of behavior itself whereby not only will the market economy impose the rule of competition on the individual, but equally potently now the individual would *apriori* personify competition as they enter the market economy and literally become, along with other elements, the creator and driver of competition in it. Indeed, in the last two decades and despite uneven applications across various regional states[6], one can

[6] The problem of legislating on issues such as education (agriculture, land acquisition, and so on) highlights the difficulties in weaving through the complexities imposed by India's federal structure that divides legislative power across one central state and many regional states; power to legislate is distributed in line with union list (pertaining to what is the domain of central government—100 items), state list (capturing the domain of regional states—61 items), and concurrent list (the intersection domain of the two—52 items). This distribution of sovereign power impacts the transition of Indian economy unevenly across regional states; this is because the latter may adopt policies in their domains that are contrary to what the central government adopts, some that may even be contrary to the precepts of neoliberalism. However, notwithstanding this complexity tension and accepting any transition in this scenario as uneven, it is also true that, with financial power distributed in its favor, the central government has been able to increase its influence over the regional states and impose its policy paradigm; moreover, competition among states for access to capital has also made it difficult for regional states to play truant. With the new Order of Things favoring a unified stance across India so far as economic policies are concerned (see Nilekani 2006), federalism is often not in great vogue. Given this clarification, we have been concentrating on the relation between this new Order of Things and the Indian state. Not surprisingly our attention is inordinately fixed on the central government (its policies, documents, practices it favors), even though as in case of agriculture, land acquisition and education, the conflict between it and regional states that appears material to the discussion

clearly discern along this line the tendency of a somewhat unregulated evolution of private education in Indian education, including in higher education.

It is now claimed that the success or failure in achieving welfare of society qua economic growth depends less on the return on physical capital, labor and land than, along with technological development, upon investment on and accumulation of human capital and its rate of return, a point emphasized forcefully by endogenous growth theory (Mankiw et al. 1992; Lucas 1988, 2004, 2009; Becker and Tamura 1990; Rebelo 1991). The emphasis should also be on removing constraints that incapacitate the decision-making process of individuals to maximize their welfare; this, of course, demands a favorable shift in state's presence and policy in relation to growth (Barro and Sala-i-Martin 2003). That is, the success and failure of countries to grow rests upon the ability or inability to create an entrepreneurial society entailing the production, augmentation, and spread of human capital of individuals in a competitive market economic structure. Therefore, if the state's objective is to maximize economic growth and welfare of individuals, then it must strive for the creation of the neoliberal homoeconomicus within a competitive market economy; hence the "motto: privatise education"; privatize not just by bringing education service under the competitive market economy, but make the *purpose* of education submit to the competitive market economy. It is being claimed that if Indian economy wants to be located within an entrepreneurial society and if innovation is to become central to business practice, then such a redirection of education is mandated (Economic Survey 2011–2012, Chap. 3). Without the creation of these rational personalities, Indian competitive market economy will continue to lag behind the rest of the world.[7]

Another big change facilitated by neoliberalism is what we call the commodification of social life; this is one way to understand the phenomenon of economization of social life. Starting from commodities defined "as objects produced for sale on the market," Karl Polanyi (2001) explains that the birth of commodification as a general form of use value was an integral component of a self-regulating market system, and hence must be seen as an indispensable condition for the subordination of society to market economy. In this context, some of the commodities without which this transformation toward market economic system is not possible are fictitious; labor, land, and money are such 'fictitious' commodities, which historically have never been, and indeed and cannot be, produced for sale in any social organization of life.

> Labor is only another name for a human activity which goes with life itself, which in its turn is not produced for sale, but for entirely different reasons, nor can that activity be

(Footnote 6 continued)
also need highlighting. It is our contention that internalizing federalism in the theorization of Indian state will considerably enrich the complexity of Indian state, bringing further into sharp focus the limitations of neoliberalism (in terms of practice) as also aspects like policy gridlocks and corruption.

[7] For a critical analysis of neoliberal education approach in India, see Bhowmick and Bhattacharyaa, (Chap. 23).

> detached from the rest of life, be stored or mobilized; Land is another name of nature, which is not produced by man; actual money, finally, is merely a token of purchasing power which, as a rule, is not produced at all, but comes into being through the mechanism of banking of state finance. None of them is produced for sale. The commodity description of labor, land, and money is entirely *fictitious* (147–148).

However, without creating these 'fictitious' commodities as a general form of commodity, there can be no private property and no impersonalized transaction, no surplus value and no appropriation of surplus value and hence no capitalist and capitalism. Creation of a market economy (and indeed of capitalism) thus involves a dis-embedding of land, labor, and money from society and their re-articulation into the market system which necessarily implies the subordination of society to the latter. Witness to the devastation brought about by capitalism by way of the great depression, Polanyi criticized the reduction of the otherwise complex existence of these three elements to that of market and avers that treating them like other commodities is simply unsustainable; this is why the creation of market economy with genuine commodities has to be reversibly associated with attempts to restrict (in which state plays an important role) the action of market relative to land, labor, and money; this constituted the now famous Polanyi *double movement*. Indeed, Polanyi's double movement paralleled the rise of welfare capitalism riding on Keynesianism which tried to create a constrained market economy and constrained capitalism with state exercising some degree of control over labor, land, and financial market by preventing the general reduction of society to market system.

Against this backdrop, we contend that neoliberalism constitutes an insurrection against the Polanyian double movement (alongside against Keynesianism) as it attempts to shred to pieces the restrictions against labor, land, and money so as to make them increasing amenable for sale in the market. Turning land/nature into natural *resource* (epitomized by pervasiveness of primitive accumulation and the bourgeoning real estate boom in a private economy), human into human *resource* (characterized by the rise of human capital being pioneered by the new education system and corporate management) and money into financial *resource* (typified by the rise of quick profit driven financial system and speculation focused on non-core practices such as investment banking, derivatives, etc.) constitute a historical shift in the way land, labor, and money have and are getting re-located; whose management now in a market economy becomes the means for maximizing profit. In India's case, while the reforms regarding the latter two aspects have been relatively smooth, the question of land acquisition for the purpose of capitalist expansion has led to fierce resistance from below by social movements which in turn has opened the question regarding whether land acquisition is to be seen as an exclusive market principle or requires some form of state and social control over the process (as epitomized in proposed land acquisition bill to be tabled before parliament). The aspect of land acquisition will be dealt with later. The connection of neoliberalism to commodification is however deeper.

By dismantling the distinction between 'economic' and non-economic' domain, neoliberalism has made the scope of competitive market system infinite and absolute in organizing our forms of life. Two such changes can be detected. The first is the

displacement of state intervention in the field of education, health, infrastructure, and so on into a domain where these hitherto 'social' goods and services could be unbundled and parcelled out as commodities produced and sold by enterprises in a competitive environment to potential consumers. The reason for commoditization is fundamentally not so much an argument of deficiency concerning the inability of state (or any collective) to deal with these (say, due to inadequate resources), but stems more from a change in outlook in which what previously was a 'social good' is now seen as better produced and delivered in a competitive market regime. Both Coase (1960) and Buchanan (1959), in different ways, were among those who cast doubt on the veracity of a natural association between state and 'social goods' by arguing that, under certain conditions, markets, if given a chance to be formed, may perform, and perform better than state; neoliberalism would argue that the job of state is to ensure that those conditions thrive (such as reducing transaction cost and frictions in bargaining between agents). For example, if there is a problem of market formation, the state may coordinate in bringing the market about through, say, the private–public partnership (PPP), whereby, notwithstanding the variability in the composition of ownership, financial, and risk sharing arrangements, the responsibility of production of goods and services tend to fall on the private sector; this again reflects the state's negative assessment of itself, its self-critical disposition.[8] Resultantly, with increased privatization and commoditization of previously delivered social insurance pertaining to health care, education and infrastructure, the risk of failure which was hitherto insured by the state (and managed by experts tied to the state) is now being displaced into the individual domain. Consequently, what appears now is the rationalization pertaining to shifting of responsibility of health, education, and other goods from the domain of the state to the individual, both as supplier and demander.[9] Coupled with this structural adjustment, the parallel inauguration of the homoeconomicus as self-governing ability machines capacitated with risk assessment and risk management abilities sets the stage for the demotion and then gradual withering away of social insurance. A complementary industry (part of the booming service sector in India) concerning individual risk related technologies (of information collection, calculation, and assessment) has risen in the last 20 years as facilitators to enable the rational agents to become (increasingly better) risk managing entrepreneurs as well.

The second transformation which is a civilization game changer concerns the human body-being. We are witnessing its dismantling into parts, that is, the segmentation of this body-being into commodities for sale; the mind is for sale, the bios is for sale, and even the emotion and passion is for sale (psychiatrics, counsellors, genetics, human resource management and stress managers in corporations, management schools, yoga gurus, etc., are some of the experts in the

[8] This facet of PPP model in infrastructure is discussed in details in Chatterjee and Banerjee (Chap. 9).

[9] Problems of re-location of agriculture in a market economy under liberalization–globalization and the role of state in this transition are discussed in Deshpande (Chap. 7). This paper also engages with the issue of federalism.

new market place dealing with these 'commodities'). The commodity character of this segmented bazaar of body-being is of course entirely fictitious, because these cannot be produced for sale; for example, emotions (that includes, among others, pain and anxiety) are not produced for sale. By turning attention to the individual, neoliberalism brings along with it this commodification of individual's body-being as part of its philosophy in propelling what could be called the *new* 'care of the self' (as against Foucault's care of the self); rather than being self-introspective, this new care of the self in effect turns into an introspection of the outside—the market—onto oneself; an introspection based on the exteriorization of the interior. What this exteriorization of the interior ala the gaze of market does to the claim of 'free' choice and to the ethical fabric of individuals is of course another matter, though no less important. Nevertheless, this fundamentally alters the idea of homoeconomicus. The homoeconomicus is thus not just one who uses the market place to organize his value, resource, and career path; he is also able to internalize the market into the organization of his body, thought, and emotion. In short, homoeconomicus as the rational ability machine is one who in totality is able to embed himself, as far as possible, into the market system with the intent to succeed in out-competing others in the process. To make that possible arises in this bazaar of body-beings a new entrepreneurial group of sellers of problem solving—of body, of mind, and of emotion. The policy of privatizing health is partly a logical corollary of this shift in the conceptual location of body-being.

What is being consequently produced are a set of new values and norms that are influenced by these contextual encounters (for values and norms are also influenced by reflection on ones' body-being) and have to be carefully delineated; but at the least, it seems to us that to claim that such values and norms are as if outside the market economy or of its signals is naïve. This is not to reduce the entire canvass of values and norms to the competitive market economy, but to argue that with the growing subjugation of the social—including the institutions and individuals—to competitive market economy, the location, meaning and role of values, and norms should also be seen in relation to the economic approach, processes and outcomes. If values and norms are relocated as a mirror of conducts which are 'economic' then the classical albeit somewhat strict division between 'values-norms' as cultural and market as economic may no longer be true, at least in the world of neoliberalism or even otherwise. Perhaps, following neoliberalism, one of the great transformations can also be seen as the instituting of a fundamental shift from cultural economy to economic culture; notwithstanding the presence and challenge of other cultural ethos, India is also witness to this changing culture.

To sum up. As a result of the shift in the rationale for the existence of state that followed its embracement of neoliberalism, two broad kinds of change can be ascertained. The first layer of change—the initial transition under new economic policy—aimed at creating and organizing the space, institution and rules for market and competition (the first phase of reform), and the second layer of change (the second phase of reform) is allowing the formed markets to operate freely in a competitive environment as also authorize the creation of new markets in areas which till now have been dealt with in non-market ways or have not been

articulated globally. In the course of this slow but steady unraveling of the structural adjustment program unfolding steadily from one generation of reform to another, India has more or less successfully shifted towards the second layer of change.

The above shift in the rationale had a telling effect on governmental practice of Indian state. If the Indian state exists for the 'good' of society, then it must follow the norms, principles, and rules which help sustain an entrepreneurial society namely those relating to competitive market economy. Neoliberalism avers that

> …not only the individual body, but also collective bodies and institutions (public administration, universities, etc.), corporations, states have to be 'lean', 'fit', 'flexible', 'autonomous': it is a technique of power (Lemke 2001, 203).

The principle founding the new art of governmentality pervades the inner veins of the state apparatuses bringing in its wake the template of 'economizing' to inflect the technique of governance of state as well. This, in turn, demands a convergence among ministers, bureaucrats, managers, and workers regarding the way statecraft is contemplated and activated.

The fact that state's performance comes to be judged on the basis of economic markers such as efficiency, productivity, profitability, etc., means that it now becomes a site of measuring and verifying 'good governance' in a similar vein as other sites of collective bodies and institutions. The correlationality and partnership of governmentality between the state and homoeconomicus as if now turns its gaze back on the domain of the state. In the context of India, actions like reducing fiscal deficit (backed by Fiscal Responsibility Budget Management act), disinvestment and privatization, merger of administrative departments, shedding off activities completely (like management of pension funds or hotel business), delegating and outsourcing previous state activities (such as in utilities like water, education, health, and infrastructure) to private agencies and individuals (say, through public–private partnership), etc., are some of the ways in which the principle of economizing is being activated. The remaining governmental functions pertaining to production and distribution of state corporations, public utilities such as infrastructure, water, health, or education (which though, as mentioned, are also fast being privatized and/or made marketable through PPP), poverty management and public distribution system are all ever more being increasingly subjected to efficiency criterion and individualization; the transition in the latter two are signified by the move from universal access to targeting symbolized by the division between above poverty line (APL) and below poverty line (BPL) and now the replacement by cash/voucher of the existing in-kind distribution system. An expectation shift is transpiring whereby the state must not be simply seen as being a service deliverer but also enterprising, attending to the individual's freedom and welfare. The state must be seen as being equal if not better than private providers of service (in terms of speed, quantity, and quality of service) if it wants to justify its existence. Instead of merely subjecting others to its rules and objectives, the state brings its own existence and operation into the realm of critical assessment and that too subjected to those modes of assessment and rules which it asks or

desires others to follow. The state in the process opens itself up for critical public scrutiny to charges of ineptness, bureaucratic red-tape and corruption like never before. While by no means homogenously true across all departments, organs or regions, the transition of Indian state bears testimony to these modifications that signals a change in its rationalization, its mode of functioning and criterion in terms of which it judges itself and is judged.

Global Capitalism and World of the Third: The New Structure of Indian Economy

What is the relation of capitalism with competitive market economy? What kind of economic map appears as a result? Taking off from a class-focused approach producing a de-centred and disaggregated economy (Resnick and Wolff 1987; Chakrabarti and Cullenberg 2003, Gibson-Graham 1996, 2006), Chakrabarti et al. (2012) theorized this new economic map as a structural re-articulation of Indian economy that takes the form of circuits-camp of global capital and world of the third; global capitalism is the conceptual shorthand for circuits-camp of global capital. We present a brief summary here. First, circuits of global capital.

With class processes, that is, processes of performance, appropriation, distribution, and receipt of surplus value fragmented across nations,[10] global capitalist enterprises are taken as the primary unit in the reorganization of the structure of Indian economy; it entails the effort to initiate the transition from capitalism circumscribed by the nation–state (the hallmark of capitalism in the era of centralized planning) to global capitalism. This is typified by the transformation of 'Indian' enterprises such as the TATA group of companies or Reliance into global capitalist enterprises (headquartered in India); accompanying these 'Indian' global capitalist enterprises are other 'foreign' global capitalist enterprises (Microsoft, for example) with activities in India pertaining to one or many class processes we mentioned earlier. Global capital or the value flows emanating from global capitalist enterprises is the centrifugal force of the circuits of global capital. Moreover, alongside changes in economic processes required to protract a competitive market economy, political process (say, provided by the state or the managers inside the global enterprise; say laws, regulations, supervision), cultural process (say, provided by schools, colleges, and universities or even enterprises advertising for the concerned global enterprise) and natural process (say, provided by access to land, forest, gas, water, etc.) constitute the *conditions of existence* of global capitalist enterprise. It is not surprising that changes in Indian laws pertaining to competition, property, disclosure, risks, accounting practice, education, forest, water, mining, land acquisition, labour, etc., are some of major reforms that have already been undertaken or are pending; much of these reforms have been undertaken

[10] Ruccio, Resnick and Wolff (1990), Chakrabarti, Dhar and Cullenberg (2012, Chap. 5).

keeping in mind the centrality of global capitalist enterprises . We refer to only a few of the economic conditions to elucidate the definition of circuits of global capital.

Local capitalist enterprises and non-capitalist enterprises are intertwined within the global production system through outsourcing, subcontracting, and so on (for example, by supplying the global capitalist enterprises with means of production or required services); thus a chain evolves through the connection of these enterprises with the global capitalist enterprises. Moreover, in addition to merchanting and providing credit to both local and global enterprises in the chain, the trading and financial enterprises (commercial banks, investment banks, mutual funds, and brokerage enterprises) also invest and take part in activities, practices and relationships pertaining to stocks, bonds, insurance, etc., setting in motion economic processes and conditions that allow capitalist enterprises to function globally; in the process, the above-mentioned enterprises increasingly take on global dispositions and become global merchanting or financial enterprises. We define the "*circuits of global capital*" as comprised all those processes that are directly or indirectly connected with the global capitalist enterprises, that is, all those processes that relate to processes pertaining to global capital. Evidently, with global capital at the center of the structure, global capitalism qua circuits-camp of global capital is bigger than the mere combination of global capitalist enterprises. In this context, capitalist development comes to be signified by the expansion of the surface and depth of the circuits-camp of global capital. In the circuits, there are some enterprises (in production and in circulation) which are directly and immediately connected to the global capitalist enterprises and together these cluster into what we provisionally name as the *hub* of circuits of global capital; while the rest in the chain which are indirectly connected and are some distance away from the global capitalist enterprises comprise the *margins* of the circuits of global capital.

As already argued, the neoliberal turn encapsulates the relocation of subjectivities as also of cultural space, constituting in the process new notions of success, of ways of judging performance and conduct, entrepreneurship and consumerism, of changing gender and caste relations, customs and mores, etc. What appears is a social cluster of practices, activities, and relationships connected to the circuits of global capital; we name it the '*camp of global capital'*. They are instrumental no doubt in the reconfiguration of the spatial contour of India's old and new cities typified by the ongoing creation of highways, wide roads, golf courses, shopping malls, seven star hotels, swanky housing projects, costly educational institutions, multi-facility hospitals, entertainment including gambling and sleaze, artifacts of native/ethnic wilderness for tourism, and so on. Reshaping of urban space is transpiring rapidly along with a shifting idea of urbanity that is concomitant with neoliberal values. A new culture that is global and market oriented is appearing around this camp, especially its hub; it is striving to dismantle extant meanings of 'good' life in India and exchange it with the tooth and claw model that emphasizes competition, possession and accumulation, and that too in a global context.

The demarcating boundary between the circuits of global capital and what we have designated as 'world of the third' crucially depends on the *local–global market*. Let us explain with the help of an example.

TISCO's interaction with the outside world is part of the global market. Similarly, ancillary enterprises, say local capitalist in nature, interacting with TISCO is strictly speaking a local market interaction, that is, an interaction materializing within the national border. However, since TISCO is a global capitalist enterprise, the interaction of ancillary enterprises with TISCO through local market constitutes an important element of the "circuits" of global capital. Moreover, ancillary enterprises may subcontract specific tasks to still other enterprises (which may even be non-capitalist in character) deep into the heartland of informal sector; this is an example of a local–local exchange that, through a chain, becomes ultimately connected to the circuits of global capital and hence is part of it. When we refer to the local–global market with reference to a commodity, we mean the chain of local–local… to… global–global exchanges that make up the entire value sequence of a commodity. Generalizing from the sequence structure, we can say that local markets and global markets are part of the same market chain as long as they connect to the circuits of global capital; outsourcing, subcontracting and offshoring are now well-known conduits deployed in creating this chain. Markets that connect to the circuits of global capital are part of *local–global markets* because the exchanges whether in local or global settings constitute the circuits of global capital.[11] In contrast, there are local markets (say, in vast segments of informal sector[12] , agricultural sector[13] and home based sector[14]) that do not form part of the "circuits of global capital" in the sense that none of the concerned enterprises engaged in these markets are connected to the circuits of global capital. We define such markets as *world of the third markets* (WOT markets).

---

[11] An exemplification of the local–global market linkage in this book is the chapter by Chattopadhayay and Banerjee (Chap. 15).

[12] A segment of informal sector is connected to the circuits of global capital (Chakrabarti et al. 2009) and the rest procreating outside its domain constitute the world of the third (Chakrabarti et al. 2007).

[13] The retail sector reform allowing control of enterprises to 'foreign' capital is supposed to create intricate linkages (production-wholesale-retail) with the farmers and draw them into the circuits of global capital which the state argues would be beneficial for them. From contract farming to such local–global market chain, there is no doubt about the structural changes transpiring within a section of agriculture too, especially pertaining to agricultural practices and institutions. Beyond the circuits of global capital is a vast enclave of agriculture sector making up the terrain of world of the third (Chakrabarti et al. 2012, Chap. 8). It is another matter that despite the dispersed influence of liberalization policies on agricultural sector which in no small part is guided by the neoliberal turn given to state (with somewhat truncated state's extant role, such as in capital formation) its performance in the last two decades leaves much to be desired (Bhaumik and Rashid, Chap. 6). The Indian state though sees the retail sector reform as one of the cornerstones of turning the performance of agriculture around.

[14] Part of the home based sector too is drawn, through the local–global markets, into the circuits of global capital while the rest is outside it, procreating in world of the third (Chakrabarti et al. 2012, Chap. 8).

WOT is a space that is conceptually never part of global negotiations; it is *outside* the, if we may borrow a term from Gayatri Chakravorty Spivak, Empire–Nation exchange. From our vantage point, the domain of Empire–Nation exchange is coterminous with the local–global market connected to the circuits of global capital. WOT is outside local–global market exchanges. A huge segment of the enterprises procreating here would fall under non-capitalist forms ranging from independent, feudal, communitic, slave to communist (Chaudhury and Chakrabarti 2000; Chakrabarti et al. 2012, Chap. 2). We will encounter in all probability capitalist class enterprises as well, including a large number of capitalist enterprises of simple reproduction type; these enterprises would have its own interlinkages forming a contrasting chain. This chain, in turn, is constituted by a cluster of economic, political, cultural, and natural processes; together, they help constitute the circuits-camp of world of the third. Notably, the circuits of WOT are knotted into WOT markets as well as into non-market exchanges; that is, much of the exchanges within WOT may not be market oriented at all but could flow simply through informal or formal arrangements not dictated by market. Further, for those driven by market considerations, it would typically not be determined by the considerations of the networks of the circuits of global capital or the rules and norms that operate within it.

Whenever and wherever we locate an economic space and associated forms of life procreating outside the circuits-camp of global capital, we conceptually come face to face with the circuits-camp of WOT. In elaborating this space, no claim is made regarding the economic standing of WOT society (it could be rich or poor, exploitative or non-exploitative), the cultural ethos (it could fundamentalist in some axis or more than liberal in others), the political regime (it could be authoritarian or open ended with regard to rules of authority) and the relation with nature (it could be friendly or non-friendly toward its surrounding environment). Importantly, then, no value judgment (good, bad, useful) can be attributed to WOT. The heterogeneity and complexity of WOT space implies that it must be seen as an *ex ante* empty space, a "hollowed out void" with(in) global capitalism, that cannot be fixed a priori by some normative presumption generated from elsewhere. If the goal is modernization through capitalist development, then world of the third stands for *structural exclusion* vis a vis the circuits-camp of global capital.

This changing map of the Indian economy was procured by, among other things, the policy primacy accorded to global capitalist performance, appropriation and distribution of surplus which, with greater rate of capital accumulation,[15] resulted in the expansion of the circuits-camp of global capital. The objective of economic growth no doubt helped fuel this expansionary logic; high economic growth and material prosperity, especially surrounding the hub of global capital (typified by the rise of the so-called middle class), followed. However, this

[15] Rate of capital formation which was 22.1 % of GDP in 1991 increased to 36.5 in 2009 (Source: Economic Survey 2009–2010, 10 and A11, Economic Survey 2010–2011).

process, not surprisingly, was associated with what Marx referred to as 'primitive accumulation'. A rapid process of expropriation of land, forest, rivers, etc., followed (especially in the last two decades), that was justified no doubt by the idealized trope of 'progress' of third world from its so-called decrepitude state.[16] However, what appears as progressive from the perspective of global capital comes to be seen as violence from the perspective of world of the third, which in turn has produced country-wide resistance against what is deemed as unjust intrusion and plunder; such has been the fierceness of resistance that land acquisition, at least on a big scale, has considerably slowed down and in some places such as West Bengal entirely stopped. The rapid land acquisition in India to expand the surface and depth of the circuits-camp of global capital and the recalcitrant opposition to it in many places across the country (Kalinganagar, Nandigram, Singur, Raigarh, Niyamgiri, Jaitapur, Chindwara, Noida, Agra-Mathura belt, Barmer, to name a few) stand testimony to these colliding perspectives (Basu 2008, Chakrabarti and Dhar 2009). In many instances, this expansion has raised ecological concerns as well. Other than the known ecological problems (for instance, global warming and climate change or extinction of vital species such as ants), India has been witness to a series of social movements that have drawn attention to the critical connection of nature with livelihood existence of people in world of the third, and on how the process of expansion of global capitalism by driving a wedge between the two (through deforestation, water pollution or absorption for industrialization, air pollution, change in land use, etc.) are threatening the existence of circuits-camp of world of the third itself. In many such instances, the combination of economic and ecological un-sustainability is highlighted in critiquing the process of primitive accumulation of capital. Parallel to the problem of balancing globalization and sustainability, India has to contend with the specific challenge of harmonizing global capitalism and world of the third. Today, India not only has a materially divided society (in terms of income, and social indicators), but also a nation with divided perspectives and understanding of what is 'progress' and 'just'. In the new economic map, the state stands as dislocated–relocated in this fractured and antagonistic field even as its commitment to capitalist development remains steadfast.

With the state going all out to ensure the expansion of the circuits-camp of global capital, the social movements pointed to the complicity of state with capital and accordingly targeted the state-capital nexus; evidently, in a democratic polity, this is bound to, and it has, put the state in stress.[17] One fall out of this conflict is the apparent securitization of an economic issue; let us briefly discuss.

---

[16] Chakrabarti and Dhar (2009).

[17] Our analysis also highlights the sharp distinction, at the level of principle, between the origin of capitalism, its birthing (which itself is a continual process) and where it has acquired a functional form as in competitive market economy. The relatively passive existence of state pertaining to the latter (professed by neoliberalism) melts away in the face of its complicity with capital in the process of the birthing of capitalism and whose appearance resembles what Marx described as an event soaked in 'blood and dirt'.

As already argued, claims over lands, forests, rivers—essentially for minerals and other raw materials—have led to a process of enclosure and the privatization of commons to 'companies' (inclusive of the phenomenon of rapid urbanization as well). Because the Dalits and particularly Adivasis reside in these areas, they have been subjected to the process of land alienation, forced evictions, and displacements; as recent evidences suggest, the state organs—national/regional/local—played an instrumental role in facilitating this process of primitive accumulation by using, bending or even violating the existing laws which anyway were inimical to the interests of the local. Instead of reducing poverty, economic growth and modernization created fresh poverty, and in this case turned out to be exclusionary; an exclusionary process that was not socially neutral. This process of dispossession was particularly potent for the Adivasis who happened to come in the way by virtue of their social life being embedded in forests and the associated surroundings (where minerals are located). Not surprisingly, these groups (alongside farmers, agricultural laborers (containing Dalits in vast numbers), and other stake holders) are in the forefront of both non-violent resistance organized by social movements against this process of primitive accumulation as also forming the main support base for the more violent resistance shaped by the Maoists. The response of the state has been manifold.

Recognizing that mere repression will not do, the state had to initiate a rethinking on the issue of land acquisition and forest rights. A union bill on land acquisition law is pending and under scrutiny. It is supposed to guarantee compensation and rehabilitation, and requires a ground level approval of stakeholders (two-third owners as in the proposed bill) which in turn is a reflection of the state's recognition of this problem as a political and not merely an economic one. Similarly, the Forest Rights Act 2006 strives to give back command of forests to the Adivasis; pressure to implement the existing environmental laws more strictly has intensified too. It is another matter that Indian regional states have their own take on land acquisition (and initiate their own laws) which has hardly prevented the growing clash and divides on this issue; this policy gridlock on the question of creating a uniform land acquisition law is indicative of positional differences within the state-capital nexus complex regarding the question of how to deal with the issue of land acquisition. Similarly, the violation of the Forest Rights Act 2006 and bending of environmental laws at the local level too is rampant (with the support of government bodies). Besides, what none of these reforms however address is what if, despite offers of compensation and resettlement, the aggrieved parties say 'no' to primitive accumulation; that is, plain and simple 'no' to the expansion of the circuits-camp of global capital. Such opposition comes under special scrutiny if they appear under the Maoist influence that intends to subsume these social struggles within a broader strategy to capture state power thereby putting them into a direct confrontation with the Indian state.

In case of the latter, the state has even chalked out a holistic strategy of an all-out assault (operation Green Hunt) accompanied by a poverty reduction program

designed for the Adivasis (whom the state openly acknowledges as the main support base for the Maoists) so as to separate them from the Maoists; the violent and benevolent side of the state is now part of an open strategy to combat and control this space-people so as to renew the combined process of the expansion of circuits-camp of global capital and inclusive development. Interestingly, here, poverty eradication is not considered a problem in itself. Rather, it is instead used as a war instrument for some other purpose, that is, in defeating the Maoists; such a reduction of the problem of poverty to a different problem was criticized heavily by Amartya Sen (2008). Overall, an interesting point emerges regarding the shape of Indian state. The changing structure of India's economic map accompanied by the logic of capitalist expansion has not merely produced an 'economization of the social', but also a 'securitization of the economic'. The economy is a security problem in two ways: (i) if economic state of exception arises via the fragility of a competitive market economy (as we shall discuss later), and (ii) if political state of exception arises via a fetter to the process of primitive accumulation. If the former entails state intervention (say, through fiscal and monetary policies) to rescue the competitive market economy from itself (but only momentarily so as to put it back on track), the latter paves the way for a securitization of the scenario where an entire area-people could be declared a 'state of exception', normal laws suspended and the sovereign takes control of the area-people (Agamben 2005). In case of the latter, the opposition to 'economic' development is consequently being displaced into a particular 'political' terrain, not for engagement or debate but to be positioned as and declared the 'enemy of state'; thus the scope of defining, branding, and punishing the 'terrorists and secessionists' is being gradually broadened to encompass the 'development dissenters'. Troublingly, what this often entails is the reductionist tendency to view even non-violent social movements through the lens of the security state, literally as eventual ruptures that are threatening to open antagonistic (political) fields beyond the confine and control of state (Badiou 2005; Laclau and Mouffe 1985; Mouffe 2005). At this point, *sanctioned politics* thrives not on expanding the scope and scale of democratization within Indian society, but on its truncation through the suppression of dissent to the so-called nationalist consensus on progress. It signals perhaps the effort to appropriate and evacuate the contesting space of the political by eliminating any antagonistic challenges such that the only reference point for the political would remain the state. This suggests that the state's passive existence courtesy neoliberalism is not totally defiant of its repressive existence derived from securitization; or should we say that the latter is a condition for the birth and functioning of the former. The concomitant process of birthing and functioning of global capitalism and neoliberalism in India is not without, to use Marx, blood and dirt. Against this backdrop, a question emerges: Is the *new* Order of Things driven by the high growth rate regime politically sustainable?

Inclusive Development and Sustainability

To begin with, it is important to conceptually differentiate between inclusive growth and inclusive development as embodied in the practice of Indian state.[18] This comes against the backdrop of some confusion regarding the contour of, and difference between, the two. While the term inclusive growth is more in vogue, the trope of inclusion in the Indian context far exceeds its meaning and mandate; to capture the term inclusion in its more holistic sense so as to make sense of the state's somewhat complex thought and practice, we need to invoke the term inclusive development.

From our vantage point, inclusive development combines the existence of structural exclusion, social exclusion and income exclusion; these in turn are acknowledged as pointing to a landscape of unsustainability which appears in relation to the existence and expansion of the circuits-camp of global capital. The challenge for Indian state is to secure this expansion of global capitalism symbolized by high growth rate (which it identities as the passport to progress and also the number one recipe for overcoming mass structural poverty) in the event of the undeniable problem of unsustainability that follows this expansion and threatens the transition process in more than one ways. The point is not to ignore or demote economic growth via global capitalist expansion, but to find a way to preserve it which in turn involves the following project: harmonize global capitalism with world of the third so as to internalize the problems of these multi-faceted unsustainability conditions. To this end, the Indian state envisages a model of inclusion that would account for the scourge of structural exclusion, social exclusion and income exclusion. 'Development' thus acquires a specific meaning in the Indian context, encompassing not only the neoliberal churning that is constitutive of the expansion of the circuits-camp of global capital, but also an internalization of these exclusions through a process of inclusion of world of the third in the growth story, either as a partner or as beneficiaries of it.

To put it succinctly, inclusive development unpacks the benevolent side of state whereby its intervention is projected in terms of a re-presentation of world of the third as a *victim other*. This is in contrast to the unsaid moment of inclusive development, namely the process of primitive accumulation that represents, as we saw, the unrelenting violence of capitalist development over world of the third whereby, due to the range of resistance and possible social instability, the latter is re-presented as the evil other, the dystopia.[19] One cannot ignore this dystopia for it

[18] Similar to neoliberalism, we make no effort to examine the existing theoretical literature on inclusive development/growth. Our focus is on how the Indian state perceives inclusive development; the theoretical literature on inclusive development, notwithstanding its importance, is a somewhat unnecessary detour in the context of our intervention. For our vantage point, these ideas—neoliberalism or inclusive development/growth—are discussed and analyzed in relation to the evolving rationale and practices of Indian state.

[19] It is another matter, one we do not pursue here, that development cannot but work by displacing the world of the third into an abnormal presence, either as dystopia or victim or

is a threat to the modernization project of capitalist development, a security threat which has to be accounted for and nullified as the former CEO of World Bank makes clear:

> A balanced and holistic understanding of the causes and effects of poverty can lead to reforms that promote inclusion, economic growth that reaches the poor, and social development—these are key to *sustainable peace*…Our job will be to help countries harness the trends… to promote growth, poverty reduction and *social harmony* (Wolfensohn 2000, 7–8).

Securitization of the economy cannot be merely a matter of repression. If the project of capitalist development is to be secured then sustainable peace or social harmony must be guaranteed. The importance of this issue of political sustainability for capitalist development becomes palpable in the case of land acquisition induced unrest that has stalled/slowed down the process of expansion of the circuits-camp of global capital and has even prompted the capitalist class to seek state intervention through new land acquisition law so as to stabilize the situation. It is a recognition of the political nature of the land acquisition problem, a point that in our opinion is not quite fathomed in general by economists; as we argued earlier, this political dimension is, however, recognized by the Indian state (and even the capitalist class) as it attempts to include voting rights as also standardized compensation and rehabilitation provisions into the new law. This, though, is only one kind of intervention. In what follows, we shall explore numerous other non-repressive inclusive approaches (addressing structural exclusion, income exclusion and social exclusion) through which the Indian state attempts to exercise control over the management of world of the third. Inclusive development is a trope that, keeping intact the expansionary impulse of global capitalism, strives to ensure the assimilation of world of the third amidst the antagonistic friction arising from the dismantling–dislocation–disruption of world of the third. Resultantly, in addition to its relatively passive existence in relation to competitive market economy, the Indian state comes to occupy a complex position vis a vis world of the third, one that is simultaneously repressive and benevolent.

In contrast, inclusive growth is more focused and pertains to the relation between growth, income inequality, and poverty in terms of what accrues to the bottom quintile of the poor as a result of growth; it acknowledges the limitation of the distributive capacity of global capitalism and seeks to internalize the need to address the poorest of the poor through the medium of the state. In so far as it addresses fundamentally income exclusion, inclusive growth is a subset of inclusive development. Having laid down the conceptual differences between the two meanings of inclusion, we now explain the change in the character of the

(Footnote 19 continued)
unknowable; sometimes as third world, as community or as social capital, all devalued re-representation (weak, fledgling, innocent, threatening, irrational, etc.) in relation to that of modern capitalist economy (Chakrabarti and Dhar 2009, Chakrabarti et al. 2012).

Indian state as it tries to deal with the multifaceted exclusions that became particularly evident in the first decade of the twenty-first century.

Despite conceptually distinguishing the various exclusions, the set of exclusions intersect, compensate, and reinforce one another which implies that the project of inclusion, more often than not, crisscross between these. Together, they form the state's intervening landscape, especially in world of the third and to some extent the distant margins of the circuits-camp of global capital. First, structural exclusion.

Structural exclusion pertains to the exclusion of world of the third from the circuits-camp of global capital as already explained by us. Addressing this exclusion means an intervention in world of the third so as to transform the conditions (say, by creating market, introducing competition, profit mentality, efficiency criteria, providing subsidized credit, techniques of production, infrastructural facilities, and so on) that in turn will alter the production structure of its economy. Instead of giving handouts, the idea here is to provide social protection, so that the people can develop human capital and productive capacity in a manner that help them internalize and manage risks in a better way. In the World Bank paradigm, social protection via risk management strategies include (i) prevention strategies (to reduce the probability of a down-side risk before a shock occurs ensuring preventive social protection), (ii) mitigation strategies (decrease the potential impact of a future down-side risk if the shock was to occur anyway), and (iii) coping strategies (to relieve the impact of the shock once it has occurred); all these strategies are unique for embedding the programs, and hence the state within the production economy of world of the third (World Bank 2001, Chakrabarti and Dhar 2013). Not only does the idea of inclusive development incorporate this aspect of structural inclusion, but in doing so, the Indian state view this process of assimilation as encompassing the possibility of endogenously derived structural transformation that will re-locate the world of the third as partners in the growth story. The objective is not merely to enable the poor to survive but to facilitate their growth out of poverty. In short, along with the survival goal, structural inclusion in the Indian context incorporates the possibility of poverty reduction. An example of structural inclusion is the MGNREGA in which the objective is to prop up the production economy of world of the third by facilitating its conditions of existence (such as infrastructure) and providing social protection to the poor (say, by way of secured employment); its commitment to prevent rural to urban migration makes perfect sense in this context.[20] Similarly, the microcredit approach through the self-help group is a state sponsored intervention to influence and transform the production economy in world of the third.[21] Wittingly or unwittingly, by securing a control of the livelihood by making it dependent on state, the latter seeks to acquire a control over world of the third and its subsequent transition.

[20] Sarkhel (Chap. 18).

[21] Banerjee (Chap. 17).

Accompanying structural exclusion is social exclusion and income exclusion to which we turn now. Let us begin by placing the historical context in which both social and income exclusion became pertinent. The disaster of 'Shining India' in the 2004 elections symbolized by the defeat of the BJP led NDA alliance led to a rethinking among policy-makers. A realization dawned that high growth has been exclusionary not only merely in the above sense of primitive accumulation (considered as inevitable in the event of capitalist induced industrialization) which exposes the phenomenon of structural exclusion to antagonistic contestation. Another facet of growth led exclusion was indicated in the argument that trickle down is not working sufficiently well, leading to growing income divide. Specifically, the benefits of high growth appear as highly unequal, thereby manufacturing, magnifying and deepening extant, and new forms of exclusions. Even as official indications, no matter how controversial, suggest that high-growth rate regime has reduced head count income poverty (Tendulkar Committee estimate, Planning Commission estimate), attention has turned toward two kinds of exclusions—social and income based—which many maintain have continued to fester under the new economic map thereby producing, in combination with the already discussed structural exclusion, the looming problem of sustainability. The phenomenon of social exclusion is sharply brought into focus in the following quote:

> In the first few years of this decade, there was a feeling that 'India was shining'. It was, however, realized that the 'feel good factor' was only in some indicators…There was an increasing feeling that only few sections of the population such as rich and middle class particularly in the urban areas, corporate sector, foreign institutional investors, IT sector have benefited from the economic reforms…On the other hand, rural India and social sector have not been shining. Social exclusion is taking place in terms of regions, social and marginal groups, women, minorities and children (Dev 2007, 14).

Distilling some figures pertaining to a few of these groups would suffice to highlight the aspect of social exclusion. Take the instance of Dalits and Adivasis who constitute close to one-fourth of India's population (around 250 million people); Dalits make up 16 % while Adivasis 8 %. (Report of an Expert Group to Planning Commission, Government of India 2008, 3) While head-count poverty rates have declined for both (the speed of decline greater for Dalits), as of 2005–2005, Adivasis make up 43.8 % (44.7 % rural and 34.3 % urban) of the population below the poverty line and Dalits 37.9 % (37.1 % rural and 40.9 % urban) (World Bank 2011, 11). The scenario of social deprivation mirrors that of income poverty, so far as the Adivasis and Dalits are concerned. For example, as of 2005/2006, infant mortality (per 1, 000 live births) and under five mortality for Adivasis and Dalits respectively were {44, 96} and {51, 88}; this is far worse than the same for other categories of population. (Expert Committee Report, Government of India 2008, 3). A similar trend can be discerned concerning literacy rate, underweight children and malnutrition among woman (BMI < 18.5); for all these cases the scenario is worst for Adivasis and Dalits (Institute of Applied Manpower Research, Planning Commission 2011). This report and others also draw attention to problems of social exclusion of the Muslims, woman, and children in India. Moreover, the quantitative figures do not capture the extent and depth of ethnic,

caste, religious, sex-based discriminations and social segregations, including those permeating through modern institutions, and state organs that continue to put up barriers against the growing aspirations of, and possible opportunities for, these groups. In short, the problem of social exclusion is undeniable.

One school of thought believes that, driven by the neoliberal order, establishment and expansion of global capitalism by assimilating world of the third into the conduits of competitive market economy would demolish these barriers (Nilekani 2006); this presumes that competitive market driven processes and outcomes are socially neutral. Through a process of leveling the social field in so far as access and opportunities are concerned, economic transformation would produce social transformation, one of whose major results (which is also the medium for producing further change) would be income growth of the marginalized sections of the population. In other words, growing economy will not only reduce poverty, but will induce social transformation in groups hitherto marginalized. Evidences so far suggest otherwise. The Indian state now confesses that the lowest quintile of the population has received minor benefits of growth; Economic Survey of India (henceforth ES) highlights that while, from 1977–1978 to 2004–2005, the average income of the poorest quintile of the Indian population showed marginal increase in comparison to its own income, it got less than 0.2 of the aggregate increase in income (ES 2010 Box 2.1, 23). Another way in which the social divide has been highlighted is the rural–urban divide in poverty rate; decline in rural poverty is far less than that for urban poverty. The breakdown figures for Adivasis and Dalits (who live overwhelmingly in the rural areas), as indicated earlier, are particularly revealing. Mal-distribution of income benefit seems to have reinforced extant structures of social divisions thereby, instead of solving, magnified an already existing problem. In other words, after two decades of reforms, the reinforcing cocktail of income inequality and social inequities entails that those in the lower bracket of income growth are also ones who typically are candidates of continuing social exclusions. This seems to suggest that the 'economic' outcome of growing income divide is not socially neutral, and that the patterns of income distribution and those of social division are mutually constitutive. By itself, the thesis of global capitalism driven economic transformation eradicating social exclusion seems moot or even spurious.[22]

If exclusion, income and social, is to be considered as a desirable goal then the limitation of competitive market economy becomes palpably clear, a point now readily acknowledged by the Indian state:

> The truth…is that, while markets have a natural propensity to deliver on efficiency, they do not have any innate propensity for equity or equality. Hence it is true that for

[22] Bhattacharya (Chap. 19) suggests that neoliberal private institutions were created by the new rich (call these clubs) to self-generate services for themselves from the market, services that were hitherto public goods. Such institutions are by construction personification of social exclusion since they constitutionally leave out the poor; they are part and parcel of the production of the camp of global capital. This highlights the point that global capitalism may not only reinforce existing social exclusions, but may even be vehicles for creating new ones.

> eradicating poverty and creating a more equitable and inclusive society, there is need for purposive action by Government—Central, State, and local. The view we take is that Government should play an enabling role vis-a-vis the market, facilitating trade, exchange, and enterprise. On the other hand, when it comes to distribution and the mitigation of poverty, Government has to be more proactive with policy interventions (ES 2011, 36).

Note the emphasis on both social inclusion and income poverty. Because competitive market economy is not capable of addressing the problem of social and income divide, it falls upon the state to intervene and enact a control over distribution, on behalf of the segment in the margins of the circuits-camp of global capital and particularly world of the third; this state sponsored distribution to improve social inequities and income inequality is beyond what is the usual ('efficient') distribution staged through market outcomes.

The state accordingly is now to be seen as combining two rationales, one that implies its somewhat passive existence with respect to competitive market economy and the other all throughout interventionist; it is to be simultaneously neoliberal and dirigiste, but now working in a coherent but divided manner.

> We need a government that, when it comes to the market, sets effective, incentive-compatible rules and remains on the sidelines with minimal interference, and at the same time, plays an important role in directly helping the poor by ensuring that they get basic education and health services and receive adequate nutrition and food. This roll back of the government in the former (circuits-camp of global capital—emphasis ours) will enable it to devote more energy and resources to and be more effective in the latter (world of the third—emphasis ours) (Economic Survey 2010, 23).

It becomes evident then that the problem of distribution of the benefits of growth to the excluded populace, whether in the form of social programs addressing socially excluded groups (with the objective to reduce social inequities) or redistributive programs targeting the bottom quintile of the income bracket (with the objective to reduce head count income poverty), further re-organized the character of the Indian state. Taken together with the intervention pertaining to structural exclusion, we come across a rationale of an active state that goes against the rationale for the absence-ing, at least rollback of the state in the context of circuits-camp of global capital.

However, recognizing the constitutive relation between income inequality and social inequities is one thing, but addressing them together is a different issue altogether. This is because the two, along with structural inequality, are distinct, albeit related, problems in the Indian context. In case of social inequities, one influential school of development has identified the problem in terms of the backward institutions, low human capital investment, and group dynamics, which prevents these marginalized groups from breaking out producing, resultantly, mass poverty in the world of the third. Discussing the situations of Adivasis and Dalits in India, the World Bank is explicit about the relation between poverty and the trap of structurally imposed social inequities epitomizing social exclusion:

> Social exclusion…has its roots in historical divisions along lines of caste, tribe, and the excluded sex, that is, women. These inequalities are more structural in nature and have kept entire groups trapped, unable to take advantage of opportunities that economic

> growth offers. Culturally rooted systems perpetuate inequality, and, rather than a culture of poverty that afflicts disadvantaged groups, it is, in fact, these traps that prevent these groups from breaking out (The World Bank 2011c: 2).

In the context of India, it is claimed that structurally imposed social inequities produce endogenously derived social exclusions that perpetuate poverty. To circumvent this problem, redistribution demands a two pronged approach: to combine structural inclusion with social inclusion. For example, development programs should seek to expand the scope of market economy in relation to its production economy, so that the excluded groups can have access to market and income opportunities (the aspect of addressing structural exclusion that we already discussed), and hence bypass the cultural traps which are holding them back. This is in addition to various kinds of transfer payment type of redistribution which pertains to the objective of social inclusion per se. What is being sought is the transformation of world of the third that will enable both structural and social assimilation; inclusive development thus telescopes political, economic, and social sustainability. As we argued, the Indian state is not totally averse to this route even though democratic politics also forces accommodation with existing 'reactionary' forces at times which may complicate and even undercut the said purpose.

However, what if the problem is additionally located *at* the centre and is *not* that of the margins. Specifically, what happens if, as the Economic Survey claims, the benefits of growth have spread unequally and did so in a manner where certain groups such as those in the hub of the circuits-camp of global capital garner a bulk of the increase in income? Economic wisdom (Kuznet's hypothesis) considers normal the phenomenon of growing inequality with increased income growth at this stage of development. What is though also acknowledged now is that structurally imposed unequal endowment may lead to a poverty trap that is irreversible (Galor and Ziera 1993); others contest this feature, but argue that some steady-state solutions could have the feature of a poverty trap (Piketty 1997). In other words, the presence of income inequality may complicate and indeed impede the classical relation between high growth and poverty reduction. Such traps may not respond to the market-based logic of trickle down, producing in turn the necessity of state intervention in redistributing resources in favor of the poor.

In India's case, the problem is consequently seen as one of somewhat weak connection between growth and poverty reduction; markets cannot be relied upon to solve a problem of this magnitude. Hence, the rationale of the existence of state as a re-distributive agency for poverty eradication as distinct from that of (local–global) market within circuits-camp of global capital and from markets that are being implanted in world of the third as part of institutional intervention aimed at subverting the extant backward structures by transforming its production economy and mitigating social exclusions.

Resultantly, other than the programs pertaining to structural and social inclusion, there was a felt need to address the aspect of growth induced income inequality, that is, address the dimension of quality of growth. That brings us to inclusive growth.

> A nation interested in inclusive growth views the same growth differently depending upon on whether the gains of the growth are heaped primarily on a small segment or shared widely by the population. The latter is a cause for celebration but not the former. In other words, growth must not be treated as an end in itself but as an instrument for spreading prosperity to all. India's own past experience and the experience of other nations suggests that growth is necessary for eradicating poverty but it is not a sufficient condition. In other words, policies for promoting growth need to be complemented with policies to ensure that more and more people join in the growth process, and further that there are mechanisms in place to redistribute some of the gains to those who are unable to partake in the market process, and hence get left behind (Economic Survey 2010, 22).

This again acknowledges the strategy of both market induced activities that would mitigate/eliminate poverty and of handouts via transfer payments to enable people to survive. Having situated inclusive growth in the context of growth, income inequality and poverty, the Indian state strives to give a sharper focus to this idea.

> A simple way of giving this idea of inclusive growth a sharper form is to measure a nation's progress in terms of the progress of its poorest segment, for instance the bottom 20 % of the population. One could measure the per capita income of the bottom quintile of the population and also calculate the growth rate of its income; and evaluate our economic success in terms of these measures that pertain to the poorest segment. This approach is attractive, because it does not ignore growth like some of the older heterodox criteria did. It simply looks at the growth of income of the poorest sections of the population. It also ensures that those who are outside of the bottom quintile do not get ignored. If that were done, then those people would in all likelihood drop down into the bottom quintile and so would automatically become a direct target of our policies. Hence, the criterion being suggested here is a statistical summing up of the idea of inclusive growth. The policy discussions that follow do not explicitly refer to this but are inspired by this idea of inclusive growth, which, in turn, leads to two corollaries, to wit that India must strive to achieve high growth and that we must work to ensure that the weakest segments benefit from the growth (Economic Survey 2009–2010, 22).

In the specific context of inclusive growth, the target of economic growth and income equality are re-located as complementary to one another; via the expansion of the circuits-camp of global capital, higher rate of growth will expand the pie which in turn will allow the state to enlarge and deepen its redistributive programs pertaining to poverty reduction.

While the idea of inclusive growth helped the state to sharply focus on the phenomenon of poverty reduction, the difference between it and inclusive development often blurs. Indeed, redistributive programs such as MGNREGA try to combine structural, social, and income exclusion by targeting the *rural poor* through *productive activities* that would typically include large numbers of Adivasis and Dalits; similarly, the poverty programs targeting Adivasis in Maoist infested districts or those that target woman have an objective of inclusion exceeding the mandate of poverty reduction; so, for example, microcredit programs through self-help group targets empowerment of women[23] (social inclusion),

[23] See the chapter by Ghosh (Chap. 8).

expansion of productive economy of world of the third (structural inclusion) and mitigation of poverty (income inclusion). Such programs are, in turn, complemented by other schemes of transfer payments (handouts), including the Mid May meal scheme for children and impending Food Security Bill; another example of transfer payments is that of *Rashtriya Swasthya Bima Yojana* (National Health Insurance Scheme) under the *Unorganised Workers' Social Security Act* (UWSSA) which covers close to 90 million people. Therefore, one comes across a complexity of state sponsored inclusive programs pertaining to production, distribution, and consumption in world of the third; this state led intervention is achieved through a redistribution of funds generated in the circuits-camp of global capital which again shows the necessity of global capitalism as a component of inclusive development.

The rethinking on part of Indian state points to an acknowledgement of the existence of two territorially distinct societies that in turn requires a bifurcation of the very rationale of the state itself. It is now conjectured that neoliberalism propelling the creation of an entrepreneurial society based on competitive market economy will drive the circuits-camp of global capital while a different dose of policies, based on redistribution of a portion of economic growth, will be targeted at world of the third; the latter will be conducted directly by, or under the guidance of, the Indian state.

Not surprisingly, inclusive development amended the contour of Indian politics, especially at the margins of the circuits of global capital and particularly world of the third. Identity, regional and even class politics are often getting reduced to cries for inclusion, represented by competition to get access to state-related distributional networks; regional state governments are being held accountable by the electorate for their performance with respect to various social programs. Put slight differently, politics including issue-based local politics is becoming state-centric, a struggle of how and whether groups can get absorbed into the conduits of inclusive development. This competitive politics over economic resources, a far cry from the criteria of efficient distribution that animates the competitive market economy, is reshaping India's democratic polity as well; particularly, identity politics (based on region, ethnicity, caste, religion, etc.) have acquired growing significance which in turn, by expanding the demand for social programs, is straining the given policy paradigm for it must account for any additional funds from the value generated in the circuits-camp of global capital. Not to lose sight of the complexity, the growing clamor for inclusive projects (such as the impending Food Security Act) runs against the voice of the capitalist class and the elite (especially those connected to the hub of the circuits of global capital, productive, and financial) calling for action to reign in fiscal deficit. In 2012–2013, government's attempt to reduce fiscal deficit, combined with other neoliberal reforms, has in turn led to a sharp increase in price of diesel, fertilizers, and seeds thereby putting the production economy in world of the third, particularly agriculture, under severe strain. This shows how, placed amidst contrasting rationales, the state could be caught in a contradictory bind vis a vis world of the third.

Finally, the relation between neoliberalism and inclusive development is not entirely about conflict. Regarding its interventionist role within world of the third,

the state is now shaped by a different menu of techniques and strategies imbibing neoliberal modes of governance (targeting, cost-benefit approach, modern management techniques, etc.) and this often requires public–private partnership whereby the state plans the schemes while the private players are put in charge of implementing or overseeing the implementation of the schemes; the rise of NGOs, microfinance institutions and private pressure groups are testimony to this changed scenario. It is another matter that the resultant expansion of NGOs has put this sector in a complex relation with state bodies and political establishment which, while mostly facilitating, is also sometimes tense. Overall, the benevolent intervention of the state into world of the third under the guise of third worldism now carries the neoliberal techniques of governance into the conduct of these economies itself even though the rationale for its intervention is hardly neoliberal. Tropes of efficiency, competition, etc., are playing a seminal role in the way the programs of inclusive development are to be fathomed; this aspect of rationalizing inclusive programs sits somewhat uncomfortably with the competitive identity politics over economic resources, a complexity that is part and parcel of India's new economic formation.

At one level, the discourse of inclusion foregrounds identity groups (based on ethnicity, caste, gender, region, religion, etc.), but in terms of some of its practice, especially pertaining to inclusive growth, the focus, no doubt made possible by the targeted categorization of income 'poor', is on the individuals or at best the households. To this effect, we are now witness to a growing monetization and individualization of the inclusion directed programs through new-fangled mechanisms such as voucher or cash payment which are now being forwarded as the conduit of distributing state benefits; these are being defended on grounds of targeting (avoid giving money to the well-off) and efficiency (avoid wastage). New technologies of identifying individuals to enable targeting such as UID are already in full swing. This neoliberal influence on the governance of inclusive development programs suggests a transformation of the philosophy of 'welfare' conduct of state, an aspect of governance we had already discussed. However, this raises one vexed question. The only way such target-based entitlement program can be absolutely successfully (that is, everyone receiving or not receiving what they deserve) is if the state can know all the relevant information concerning all the citizens of India (hence the importance of UID technology). This phenomenon of absolute knowing though is contrary to the very principle of neoliberalism that denies any possibility of state or a centralized entity knowing the 'economy' in totality, a point that paradoxically was one of the central focus of attack on India's centralized planning system (Chakravarty 1987). It is though believed that this time the silver bullet of UID technology will make the state omnipresent, and hence once again enable it to conduct social engineering. Notwithstanding the claims and counterclaims, the debate between universal and targeted welfare, between cash verses kind distribution is bound to continue in the future.

Global Economic Meltdown, Indian Economic Crisis, and The State

Placed in the triad of neoliberalism, global capitalism and inclusive development, the Indian state has been shown to be pulled and pushed by dissimilar rationales and contradictory roles. These contradictions, compounded by external events, led to a crisis like situation in the Indian economy since 2007 which, in turn, forced a further bout of self-introspection by the Indian state. This crisis, concentrated in the circuits-camp of global capital, paralleled the appearance of global economic meltdown. While debates do exist concerning the extent to which the global economic meltdown can be held responsible for India's present economic crisis, there is no denying the fact that, by virtue of India's integration into the global economic order, the phenomenon of unfolding global economic meltdown is implicated in the Indian economic crisis (thereby putting an end to the thesis of decoupling). Let us examine the resultant alteration once again in the rationale and practice of Indian state that transpired following the advent of India's economic crisis.

At the level of principle, neoliberalism demands a 'thin' state that would govern for the competitive market economy. It will do so by (i) creating and then leaving that economy to self-regulate itself, and (ii) providing conditions for homoeconomicus to cultivate themselves and consequently exercise their freedom in an unbounded economy that respects no boundary between 'economic' and 'non-economic' domains. It is claimed that this unbounded competitive market economy ruled not by the state but freedom of the mass of homoeconomicus will generate optimum outcome and produce a self-regulation of the economic structure; in totality, such a harmonious system will produce the maximum possible growth and individual welfare for the society. Once such a standard is fixed, any deviation can be identified (which may happen due to high transaction cost, bargaining or coordination failure and so on), and the state will accordingly take policies, keeping the interest of society in mind.

This imagery led to a defense of the competitive market economy as against state intervention/control, and of global capitalism as against capitalism circumscribed by the nation–state. India's high-growth rate regime is supposed to be an example of such an economy. However, as global financial meltdown soon gave way to a global economic meltdown, since 2007 India's growth story is in serious peril which in turn forces our attention to areas taken thus far as immune to discussion. In fact, by 2012, Indian economy's trouble has expanded well beyond the faltering growth story. A cursory glance at some of the basic indicators tells a sorry story: falling growth rate, resiliently high inflation rate (transpiring, says RBI, mostly from supply side factors), falling private saving and investment rate, growing current account deficit driven mainly by worsening trade deficit, pressure on capital account, declining rupee value and at times volatile exchange rate movement, and uncomfortably high fiscal deficit. This is indeed a case of fundamentals gone haywire. This time though, with a drastically truncated role of state

vis a vis market and weak trade unions, the blame for this crisis can only fall on global capitalist production, distribution, and consumption of goods and services functioning through the conduit of competitive market economy. Is it then the case that the source of systemic crisis resides in the ills of competitive market economy and the structure of global capitalism it has helped shape?[24]

Putting aside this question for the moment, what interests us here is the response of the Indian state to this sudden interregnum to its dominant imagery. Faced with this near systemic crisis, it seemed to have accepted the exceptional nature of the situation and the need to reorganize the rationale of state to internalize the possibility of this and any other comparable exceptional situation. Because the economy is perceived as passing through an abnormal phase, 'normal' economic policies are suspended and replaced by state sponsored policies pointing to a full scale war to stabilize the competitive market economy so as to revert back to a high growth rate regime. In the last 5 years, fiscal, monetary, and exchange rate policies, to name a few, have subsequently unfolded in a manner that reflects this intervening characteristic of the state; this includes temporary suspension of the FRBM act, expansionary followed by tight monetary policy and fairly active role in trying to stabilize the currency market in favor of rupee. Two aspects stand out in this new formulation of the state of exception. To begin with, while appearing to be all encompassing and exhaustive, the state intervention rationalized by the announced presence of the state of exception is selective in its application and tries best to not alter the basic template of neoliberalism. The art of governmentality exercised by free agents (producers and consumers) is being left fundamentally intact or even if momentarily arrested in some situations never denied or usurped in any permanent fashion; the sanctity of market and competition is reiterated time and again, the latest proof being the reforms in the retail sector and the proposed banking, pension and insurance market reforms. There is a

[24] This global economic crisis has led many to question the neoliberal belief that the framework of modern macroeconomics has found the key to overcome economic depression and hence systemic collapse (see the articulation of this belief in Lucas [2003]). However, individuals (parts) being free decision makers, competitive markets self-regulating, and sufficient for the system evidently rules out any autonomy to the economic structure which is specified in neoclassical economics by general equilibrium economy (whole); the underlying assumption being that the whole is an addition of the parts. If autonomy of structure is granted then that could carry the possibility of effects and outcomes that cannot be reduced exclusively to the optimizing behavior of the agents interacting through market. But then, unless we agree that this autonomy exists, it becomes difficult to locate and explain the appearance of the current economic crisis that is now global. In explaining the periodic breakdowns of capitalism, Keynes and Marx, in different ways, insisted on the relative autonomy of the structure, a relative autonomy that can be traced to the peculiarity of the structure, and at times the non-optimizing behavior of the agents. This throws up the methodological challenge that parts do not add up to the whole; that the whole exceeds the parts and that the whole also needs to be specified and analyzed in terms of its unique features and effects. This is not to say that individual decisions and market do not influence the structure, but that structure cannot be seen as mere sum total of individual decisions. This unresolved tussle between the parts and whole presents a methodological conundrum from which the field of macroeconomics has not yet found a way out in its relatively young history.

thin but critical line between governing for society and governing society and, unlike classic Keynesianism, state interventions in this crisis seems to be carefully crafted in ensuring that the state does not bring society totally under its sovereign control. Nevertheless, by contending that the competitive market economy literally needs the state to survive and function, a new meaning of state of exception gets defined. This state of exception is economic in nature. Second, particularly glaring is the attempt to incarcerate state interventions within a definite time span and a commitment to quit once the economy is 'back on track'. The sovereign feature concerning the 'economic state of exception' is contemplated as limited and specific. Nowhere is this better exemplified than the policy emphasis that problematizes high fiscal deficit and fast paced monetary policy change (which has been followed in the last 5 years) which it seeks to roll back once the recovery becomes visible. The 'recovery' of course is indicated by the growth rate, manufacturing figures, corporate profit and investment, and private consumption demand under conditions of price stability. Once positive signals on these indicators emerge, the Indian state is committed to roll back to its old macroeconomic paradigm of tight fiscal and monetary policy which it perceives as correct/normal; its commitment to neoliberalism henceforth continues to remain unquestioned. Or, perhaps, because of the resilience of the crisis even after five years, the return once again to neoliberal macroeconomic policies (say, tight fiscal policy) and renewed structural adjustment is now being considered as a panacea to rectify the faltering growth rate and reduce inflation rate, that is, as a medicine to end the crisis; though the mechanism to achieve this is not clear, especially when the same set of conditions were at least partly responsible for the crisis in the first place. Such a reversal is already underway in countries of the European Union with their austerity program and in the USA (where the austerity drive is being driven by the regional states). Notwithstanding this confusion between pursuing temporary Keynesianism and renewal of neoliberalism, this short episode has also demonstrated that the state may have to internalize unintended consequences which transpires under conditions of neoliberal forms of globalization and global capitalism. The idea though is not to rule society but to save the competitive market economy from itself.

It is another matter that, as we are witness to, the economic structure (indicated by growth rate, inflation rate and rupee value) seem to be somewhat invariant to policy changes (pertaining to fiscal, monetary and exchange rate regimes) which suggests vulnerability of the structure of competitive market economy itself. This trouble is in fact finding further fodder through the global inter-linkages that is exporting global problems into India in plentiful forms (the deleterious effects from Europe being the latest addition) thereby aggravating an already difficult situation. The trouble is not merely that the economy is faltering and seemingly stuck in a trap, but that the process is transpiring in a dynamic environment that is private (competitive market economy) and global, in which many processes are not under the control of the policy-makers if they are at all known to them or can be reasonably forecasted in the first place. So much has been talked about the benefits India has garnered from its integration into the global economy. Yet, the last 5 years have shown that there is a cost of this integration too which now can

hardly be left unquestioned. A lesson: there is no win–win harvesting from globalization. Like all other entities, the process of globalization is beset with overdetermination and contradiction, throwing up unpredictable outcomes and harboring unknown possibilities; the presence of competitive market economy integrated with global capitalism suggests the possible existence and the need to not only accept the possibility of business cycles (fluctuations around trend lines), but also of breakdowns. However, dealing with a scenario of systemic crisis in a centralized planned economy is one thing and dealing with systemic crisis in a competitive market economy integrated with(in) global capitalism is quite another. As of now, knowledge regarding how to deal with the latter situation seems limited; this confusion is reflected in the apparent conflict of recovery strategy between Ministry of Finance and the RBI concerning the goal and nature of monetary and fiscal policy. That the former wants the RBI to take a proactive monetary policy regime which it refuses to undertake is an example). The current strategy (2013) of reverting once again from government expenditure induced growth with high fiscal deficit to that of private investment induced growth with lower fiscal deficit without quite laying down the mechanism of how such a strategy will actually transpire is another example of this confusion. The crisis in the interstices of neoliberalism and global capitalism seems to spill over into a crisis of policy-making as well.

Conclusion

Having examined the reorganization of the rationale and role of the Indian state in the event of a remapping of the Indian economy, we have not just explicated the changing relation between state, economy and society, but in doing so unpacked many of the new challenges facing it in the altered scenario. The new economic map and the problems therein indicate, among others, the potent challenges of sustainability (political, economic, social, and ecological), challenges that the state recognizes and strives to internalize at best without putting to risk the created Order of Things. This forces the state into many contradictory directions with contrasting underlying rationales; internalizing these ambiguities under the given conditions of neoliberalism, global capitalism and inclusive development is the challenge which the Indian state increasingly encounters today. Given its commitment to the existing Order of Things that it helped create, whether this maneuvering of the Indian state will succeed in harmonizing India's transition process is another matter. This difficulty is best appreciated in the context of the dual objective of inclusion and economic growth, an issue with which we end the discussion.

Redistribution for inclusion demands a certain flow of funds; it is held that this flow of funds would essentially come from economic growth transpiring in the circuits-camp of global capital. In other words, there must be a match across time between funds from economic growth and funds needed for re-distribution. The

trope of inclusive development thus must come to terms with a knife edge problem of financial sustainability. The trouble is that the relation between growth and inclusion is not supposed to transpire in a static but dynamic plane across time making it an inter-generational problem.

Already, the trope of inclusive development is under severe stress following the economic crisis symbolized by the falling rate of growth that is particularly affecting the hub of the circuits of global capital (the primary value creating sector) that encompasses manufacturing and services. This is putting severe pressure on the margin of the circuits of global capital because, as it is connected by the chain of local–global markets, any weakness in the hub puts the margin under acute stress, including that of livelihood. On the other hand, the high fiscal deficit (in part due to managing inclusive development) along with the pressure to cut fiscal deficit (because of its supposed growth detrimental effects) by reducing subsidy to segments in the margins of circuits of global capital and world of the third (the recent decontrol of oil prices and subsidy cap on LPG exemplifies this tendency) puts pressure on the feasibility of inclusive development itself. That is, expansion of inclusive development programs, especially in the face of somewhat stubborn slowdown in growth rate, points to the problem of sustainability of the whole paradigm of the new order itself. This anxiety over the question of financial sustainability is captured in the unfolding debate over the proposed policy of 'Food Security Act' whereby the 'radicals' seek the establishment of *right to food* as an unchallenged right with zero tolerance for its violation while the 'conservatives' led by neoliberal economists are challenging it on grounds of budgetary consideration and distortionary effects on the real economy arguing that the quantum of created, appropriated, and distributed wealth could even diminish. This tension between growth and inclusion encapsulates a problem and stress that is novel as it arises in the context of the new-fangled triad of neoliberal globalization, global capitalism and inclusive development that has altered the contour of India's economic map.

References

Agamben G (2005) State of exception. Chicago University Press, Chicago

Badiou A (2005) *Infinite Thought*. Translated and introduced by Oliver Feltham and Justin Clemens. Continuum, London and New York

Bardhan P (2007) Poverty and Inequality in China and India: Elusive Link with Globalisation. Economic and Political Weekly 42:38

Barro RJ, Sala-i-Martin XI (2003) (2nd edn). Economic growth. MIT Press, London

Basu PK (2008) Globalisation an anti-text: a local view. Aakar Books, New Delhi

Becker G (1973) The economic approach to human behaviour. University of Chicago Press, Chicago

Becker G (1993) (originally 1963) (3rd edn) Human capital: a theoretical and empirical analysis with special reference to education. University of Chicago Press, Chicago and London

Becker G, Murphy KM, Tamura R (1990) Human capital, fertility, and economic growth. J Political Econ 98 (October: Part 2):S12–S37

Bhagwati J (2002) Free trade today. Princeton University Press, Princeton
Birla K, Ambani D (2000) A policy framework for reforms in education. PMCTI
Buchanan JM (1959) Positive economics, welfare economics, and political economy. J Law Econ 2:124–138
Buchanan JM (1991) Minimal Politics of Market Order. Cato J 11(2):215–216
Butler J (1997) The psychic life of power: theories in subjection. Stanford University Press, Stanford
Chakrabarti A, Cullenberg S (2003) Transition and development in India. Routledge, New York
Chakrabarti A, Cullenberg S, Dhar A (2007) Orientalism and Transition in India in the Era of Globalization. In: Elliott E, Payne J, Ploesch P (eds.) Immigration and migration, social change, and cultural transformation
Chakrabarti A, Dhar A, Cullenberg S (2012) World of the third and global capitalism. Worldview Press, New Delhi
Chakrabarti A, Dhar A (2009) Dislocation and resettlement in development: from third world to world of the third. Routledge, London
Chakrabarti A, Dhar A (2013) Social funds, poverty management and subjectification: beyond the world bank approach.Camb J Econ Oxf J. Advanced Access, Published online doi:10.1093/cje/bes077
Chakrabarti A, Chaudhury A, Cullenberg S (2009) Global order and the new economic policy in india: the (Post) colonial formation of the small scale sector in India. Camb J Econ 33(6):1169–1186
Chakravarty S (1987) Development planning: the Indian experience. Clarendon Press, Oxford
Chaudhury A, Chakrabarti A (2000) The market economy and marxist economists: through the lens of a housewife. Rethinking Marxism 12(2):
Coase Ronald H (1960) The problem of social cost. J Law Econ 3(1):1–44
Dev MS (2007) Inclusive growth in India: agriculture, poverty, and human development. Oxford University Press, New Delhi
Economic Survey 2009-2010 (2010) Ministry of finance: GOI. Oxford University Press
Economic Survey 2010–2011 (2011) Ministry of finance: GOI. Oxford University Press. 21 March 2011
Foucault M (2008) The Birth of Biopolitics: lectures at the college de france 1978–79. Translated by Graham Burchell, Palgrave Macmillan
Friedman M (1962) Capitalism and freedom. University of Chicago Press, Chicago
Galor O, Ziera J (1993) Income distribution and macroeconomics. Rev Econ Stud 60(1):35–52
Gibson-Graham JK (1996) The end of capitalism (as we knew it): a feminist critique of political economy. Blackwell Publishers Ltd, Oxford
Harvey D (2005) A brief history of neo-liberalism. Oxford University Press, London
Hayek FA (1944) The Road to Serfdom. Routledge, London
Institute of Applied Manpower Research, Planning Commission (2011) India Human Development Report 2011. Towards Social Inclusion. Government of India, New Delhi
Kumar A (2008) Against neoliberal assault on education in India: a counternarrative of resistance. J Crit Educ Policy Stud 6(1) http://www.jceps.com/?
Kydland FE, Prescott EC (1977) Rules rather than discretion: the inconsistency of optimal plans. J Political Econ 85(3):473–492
Laclau E, Mouffe C (1985) Hegemony and Socialist Strategy: Towards a Radical Democratic Politics, Trans. W. Moore and P. Cammack, London: Verso Press
Lemke T (2001) The birth of bio-politics—Michel Foucault's Lecture at the Collège de France on Neo-Liberal Governmentality. Econ Soc 30(2):190–207
Lucas R Jr (1976) Econometric policy evaluation: a critique. In: Carnegie-Rochester conference series on public policy 1(1):19–46
Lucas R Jr (1988) On the mechanics of economic development. J Monetary Econ 22(1):3–42
Lucas R Jr (2003) Macroeconomics Priorities. Am Econ Rev 93(1):1–14
Lucas R Jr (2004) Lectures on economic growth. Harvard University Press, Cambridge
Lucas R Jr (2009) Ideas and growth. Economica 76(301):1–19

Mankiw GN, Romer D, Weil DN (1992) A contribution to the empirics of economic growth. Q J Econ 107(2):407–437
Mouffe C (2005) On the political. Routledge (Indian Reprint), London
Piketty T (1997) The dynamics of the wealth distribution and the interest rate with credit rationing. Rev Econ Stud 64(2):173–189
Polanyi K (2001) [1944] (2nd edn) The great transformation: the political and economic origins of our time. Beacon Press, Boston
Rebelo S (1991) Long-run policy analysis and long-run growth. University of Chicago Press, J Political Econ 99(3):500–521
Resnick SA, Wolff RD (1987) Knowledge and class: a marxist critique of political economy. University of Chicago Press, Chicago
Report of an Expert Group to Planning Commission (2008) Development challenges in extremist affected areas. Government of India, New Delhi
Sargent T, Wallace N (1976) Rational expectations and the theory of economic policy. J Monetary Econ 2(2):169–183
Shaikh A (2004) The economic mythology of neoliberalism. In: Saad-Filho A (ed.) Neoliberalism: a critical reader. Pluto Press, London
Schultz TW (1960) Capital formation by education. J Political Econ 68(6):571–583
Schultz TW (1961) Investment in human capital. Am Econ Rev 51:1–17
Schultz TW (1962) Reflections on investment in man. J Political Econ 70(5):65–94
Sen A (2008) Violence, identity and poverty. J Peace Res 45(1):5–15
Nilekani N (2006) Imagining India: the idea of a renewed nation. Penguin Press, New York
Ruccio D, Resnick S, Wolff R (1990) Class beyond the nation-state. Rev Radic Political Econ 22(1):14–27
Wolfensohn JD (2000) Interview. *Partnerships for Development*. World Bank Report 2000. Oxford University Press for the World Banks
World Bank (2001) Social protection sector strategy: from safety net to springboard. The World Bank Group
World Bank (2011) Poverty and social exclusion in India. Oxford University Press for the World Bank, New Delhi

Chapter 3
Global Financial Crisis: What Did We Know, What Have We Learnt?

Kumarjit Mandal and Sabyasachi Kar

Introduction

Financial crises, in one form or the other, have been ravaging the world economy from time to time. The great depression of 1929 was perhaps the most significant financial crisis in the economic history of the world in terms of its geographical spread and deep impact. Commentators are unanimous at crediting the global financial crisis of 2008 as the second-most significant crisis (Hall 2010). In between several crises of smaller dimension-like Mexican crisis, Latin American debt crisis and Asian crisis have frequented us. All these crises had played havoc with the rhythm of economic lives of the people of the affected countries. But at the same time, they made the people understand the importance of exercising prudence and caution in economic policy-making.

Like meteorologists, economists do not enjoy the luxury of conducting controlled experiments with their study objects. A meteorologist is incapable of generating a *sunami*-like disaster in the controlled environment of his laboratory in order to collecting data on that natural phenomenon. Similarly, an economist would not be able to simulate a financial crisis like situation in any laboratory or in the reality. The only route before him for enriching his knowledge-base about the financial crises is to analyze and scrutinize the events surrounding the crises that already happened. The science of economics has been discreetly building up its reservoir of knowledge about financial crises through this method.

The great depression of 1929 made the economists and policy-makers wiser and the domain knowledge of economics much richer. That crisis was the cradle of great many new thinking and ideas that dominated the minds of the next few

K. Mandal (✉)
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: kumarjitm@hotmail.com

S. Kar
Institute of Economic Growth, Delhi 110007, India
e-mail: skar_ieg@yahoo.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_3, © Springer India 2013

generations of economists. Similarly, the post-Global-Financial-Crisis 2008 period has been a witness to the germination of a lot of brainstorming debates and discussions about the causes and consequences of the crisis. Surely, some more time would elapse before new rays of light come out of the heat of those debates and discussions. However, the worth and appropriateness of an assessment of these discourses need not be overstated. An attempt has been made in the present survey to collate different streams of thinking around the global financial crisis of 2008.

Whenever the epicenter of the financial crisis has been a developed country its impact had been huge. The less developed countries could not escape the negative effects emanating from it. On the contrary, if the epicenter of the crisis happens to be a less developed country, its ripples does not reach the shores of those developed countries. Therefore, Mexican peso crisis or Asian crisis did not rock the world economy in the same way as the global financial crisis of 2008 did. Therefore, the emerging market economies like India need to study the financial crises with greater sincerity and adroitness. The fact that *thankfully* India has so far not been an epicenter of the financial crisis should not serve as an excuse for avoiding an analytical study of the same from Indian perspective.

The present survey starts by collating the different stances of the financial crises literature. It helps us understand the present state of knowledge about the financial crises. Then, it proceeds to examine the fundamental economic phenomena that have been sighted as the possible triggers for the crises. It also briefly assesses the policies undertaken by the different countries in combating the crisis of 2008 in a bid to understand the crisis better. This also provides an opportunity to present India's take from the crisis.

Keynesian Models After the Great Depression

As the students of economics are aware Keynes' (1933) famous treatise *General Theory* was set against the backdrop of the great depression. This contribution posited the unforgettable episode as the problem of effective demand management. Subsequently, a lot of formal theorizing took place which revolved around the challenge of specifying the appropriate *multiplier* that would boost up the sagging components of effective demand equation. Keynes visualized the effective demand equation[1] as the fundamental relationship amongst the macroeconomic variables in the form:

$$Y = C + I + G + X - M \tag{3.1}$$

[1] According to national income accounting, this is an identity reflecting a statistical relationship. Therefore, it always holds good *ex post*. However, Keynes treated it as an equation by making the variables function of *ex ante* (or planned) income (Y). Thus, relationship (3.1) is an equation *ex ante*, which becomes an identity *ex post* at the equilibrium value of Y.

That is, to say, that the domestic income (Y) consists of the consumption (C), investment (I), government expenditure (G), and exports (X) net of imports (M). Since the variables on the right-hand side of the Eq. (3.1) represent the components of effective demand any policy capable of pushing these components up would offset the cyclical downturn of the economy. Keynes proposed that the government expenditure—being somewhat independent to income—is the fittest candidate as the core component of any such policy. Hence, the dominance of fiscal policy in crisis management has been established.

Therefore, Keynes considered the economic crisis as fallout of effective demand mismanagement and the malady is curable through efficient policy maneuvering. This Keynesian position held its way for a long time. Even today, the St. Louis type models used by the central banks for regular forecasting purposes have drawn extensively from the Keynesian paradigm (Mankiw 2006).[2]

However, the dominance of the Keynesian 'effective demand' paradigm in the crisis management was first challenged by the Latin American currency crises of the 1970s. These crises convinced the policy-makers that the erstwhile practice of viewing macroeconomic crises as variant of the effective demand management problem would no longer be tenable. The foreign exchange markets were seen to be playing havoc in the economy, even though apparently there were no sign of effective demand failure. So the necessity was felt by the economic professionals of formulating analytical models focusing on the workings of the foreign exchange markets. That gave rise to a slew of crisis theories in the subsequent period.

The Bretton Woods System and Thereafter

The rising importance of foreign exchange markets in macroeconomic analysis is interwoven with the development of exchange rate regime since the World War II. The dispensation that followed the Word War II is known as the Bretton Woods[3] system. This was a kind of restricted gold standard. Under this system, each country was to peg its home currency in terms of gold and bilateral exchange rates would have been worked out on the basis of the concerned countries' pegs. Since the dollar occupied a dominant position in international payments and settlement, it was the only currency convertible into gold at the fixed price of $35 per ounce of gold.

---

[2] Mankiw (2006) said that even if one's knowledge of economics dates back to the undergraduate textbooks of the 1960s one would be able to sit in the FOMC meetings comfortably. Since the macroeconomics textbooks of the 1960s were loaded with Keynsian flavor the above comment of Mankiw basically points to the relevance of Keynesianism at the parlance of policy-making even today.

[3] Bretton Woods is a quiet town in New Hampshire, USA. Here negotiations took place to find a suitable alternative to gold standard as modes of international payments and settlement. Keynes also participated in the deliberations and took an active role. The negotiations were finalized in 1944.

The other currencies were convertible easily into dollars. The countries under the system used to maintain the announced parity between their home currency vis-à-vis the dollar by buying and selling dollars in the international foreign exchange market. However, the countries had the liberty of adjusting the peg in case of difficulty after due negotiation with the International Monetary Fund.

The declaration of non-convertibility of dollar into gold on August 15, 1971 forced the system to collapse. However, the seeds of this eventuality were sown much earlier with the fact that the amount of dollars officially held by non-US central banks was much greater than the official US gold reserve and the non-US central banks were not interested in converting their dollar reserve into gold reserve. This kind of situation could be reckoned as a signifier of dollar reserve system replacing gold reserve altogether.

The collapse of Bretton Woods system witnessed the emergence of 'nonsystem'[4]—to use the terminology coined by Williamson (1976). Under this new dispensation, there was no coherent set of rules and precisely measured parity calculations between the currencies worthy to be called an exchange rate regime. Gandolfo (2002) describes the situation as:

> In fact, the situation at the moment of going to the press is that each country can choose the exchange-rate regime that it prefers and notify its choice to the IMF, so that various regimes coexist. Some countries peg their exchange rate to a reference currency (usually the dollar, but also the French franc and other currencies) with zero or very narrow margins; naturally, they will follow the reference currency's regime with respect to the other countries. Then there are other countries which peg their currency to a composite currency such as, for example, the IMF's Special Drawing Rights (SDR). Groups of countries enter into monetary agreements to form currency areas, by maintaining fixed exchange rates among themselves, or monetary unions with a common currency, such as European Monetary Union. Further, the situation is continually changing, hence the reader had better to consult the monthly International Financial Statistics published by the International Monetary Fund where a table is present showing the exchange rate arrangements existing at the moment (pp 37–38).

The advent of such an arrangement in the international currency system has gradually become a favorable breeding ground for speculation. Whether speculation is good or bad for the economy has tormented the minds of economists for a long time—but the answer still remains inconclusive. Traditional wisdom would argue that the speculators' strategy of buying when price is low and selling when price is high might smooth the price fluctuations around its *fundamental*[5] value. Hence speculation might be beneficial.[6] On the other hand, some economists have reasoned that the possibility of speculators buying when prices are high in order to

---

[4] This denouement has been continuing since then.

[5] Strictly speaking, 'fundamental' refers to some equilibrium value. However, it sometimes refers to normal value in loose sense of the term.

[6] Friedman's (1953) remark in this context is worth mentioning: "People who argue that speculation is generally destabilizing seldom realize that this is largely equivalent to saying that speculators lose money, since speculation can be destabilizing in general only if speculators sell when the currency is low in price and buy it when it is high". McKinnon (1983) even showed that

pushing up the prices further and making profit out of such trading would render the above presumption untenable. For example, the noise traders are supposed to react to 'news' and uninformed demands in securities markets and their movement in the market are not rational. Many times their operations force the market prices away from their fundamental (Shleifer and Summers 1990). The destabilizing forces unfolded by these kinds of operations would make speculation harmful for the economy.[7] Therefore, the role of speculation remains an unsettled proposition in international finance, particularly so under flexible exchange rate arrangement.

The State of Crisis Theories

However, whenever the currency crises happened the speculation has been the usual suspect. The economists have formalized the role of speculation in those crises. This branch of the literature has been categorized as *first-generation*, *second-generation* and *third-generation* models.[8]

The motivation for the first-generation models came from Mexican crisis of 1973–1982 and Argentinean crisis of 1978–1981. Under this modeling strategy, the country is engaged in a domestic credit-financed fiscal profligacy under a fixed exchange rate. Given a fixed money stock, such a policy would deplete the foreign exchange reserve of the country. This process would push the foreign exchange reserve of the country to the unsustainable minimum waiting only to be vanished by the final speculative attack. This would prompt the country to abandon fixed exchange rate regime. Krugman (1979) and Flood and Garber (1984) presented their models in this line.

The European crisis in the early 1990s and the Mexican peso crisis of 1994 begot the second-generation models of economic crises. Under this modeling scheme, speculative attacks occur not in response of any maladjustment in economic fundamentals of the economy, rather to self-fulfilling panics. These models introduce non-linearities and endogenous response of government policies to changes in private behavior. The government faces a trade-off between various targets like exchange rates, employment, etc. Now, since the commitment to fixed exchange rate is state-dependent rather than state invariant as in the first-generation models 'the government can always exercise an escape clause, that is, devalue, revalue or float'. The presence of non-linearities results in multiple

(Footnote 6 continued)
in some cases speculation is capable of turning an otherwise unstable flexible exchange rate regime into a stable one.

[7] The cases of profitable and destabilizing speculation were cited by Kemp (1963), Ljungqvist (1992) and others.

[8] An excellent textbook exposition of these models would be found in Gandolfo (2002) and Agenor and Montiel (1999). This section is based on these two sources.

equilibria in these models, some of which are stable and some of which are unstable. Sachs et al. (1996) presented the pioneering work in this field.

The Asian crisis in the late 1990s laid bare the inadequacy of the first- and second-generation models. The first-generation models would fail to explain the Asian crisis because all the troubled Asian economies were enjoying comfortable budget surpluses prior to the crisis. The relatively strong position of these economies in terms of macroeconomic indicators like income growth, employment, and inflation bears evidence against the relevancy of second-generation models. Hence, a new breed of models have started coming up (Krugman 1999; Chang and Velasco 1998; Corsetti et al. 1999). As a common thread among these models, the link between banking crisis and currency crisis has been stressed; however, the chain of causation is not clear. These models attempt to highlight the interplay of a host of factors (both financial and real) that might lead to sudden conversion of foreign capital inflows into outflows, thus precipitating the banking and the currency collapse. Anzuini and Gandolfo (2000) traced the following three main causes from the literature of third-generation models setting forth the crisis:

(i) *Moral Hazard*[9]: The crisis is initiated, in general, by the sudden fall in the value of external investment by the corporate sector due to some external or internal shock. This begs the question why the corporate sector takes the risk of over exposure in its investment portfolio. The plausible answer lies in the intrinsic belief by the domestic firms that the government would bail them out in case they are threatened with bankruptcy. The announcement of any non-interventionist policy by the government is 'never fully credible ex-ante because the agents know that policy intervention will be decided ex-post via a cost-benefit analysis'.

(ii) *Financial Fragility*: The liquidity squeeze played a significant role in the crisis. The loss of confidence by domestic and foreign investors leads to premature liquidation of assets with banks and financial intermediaries. This has the real effect as the premature liquidation entails some loss in value of those assets. Moreover, the investors hold back those liquid funds in their portfolios while dryness prevails in the market.

(iii) *Balance Sheet*: The devaluation by the country in trouble might blow up the foreign debt of the corporate sector. This would jeopardize the balance of payment position of the country further.

Apart from the generations models, the other theory that attempts to provide an explanation to the financial crises is known as contagion. According to this literature, the financial crises are like infectious diseases tend to spread contagiously affecting both the weak (crisis-prone) and the economically sound countries. Eichengreen et al. (1996) found evidence of contagion in a panel data from 20 industrialized countries over a period of 30 years. They also found that contagion spreads more easily to countries that are tied by close international trade linkages than to countries in similar macroeconomic circumstances. The analytical model

[9] The use of moral hazard in modeling finance has become very popular. Interested readers may see Tirole (2005).

of Gerlach and Smets (1995) also lends support to the empirical findings of Eichengreen et al. (1996). The contagion effect is relevant in explaining the European Monetary System turmoil in 1992–1993. The contagion theory has very powerful policy implications. It means that if any country is in financial trouble that country should be bailed out lest it affects other countries.

The above survey of theories of the financial crises points to the fact that each major crisis has contributed a new modeling strategy. The global crisis of 2008 is yet to throw up a structure of analytical modeling. Though we are not sure about the kind of modeling the global financial crisis of 2008 would bring forth, perhaps it would be a mixture of third-generation models with contagion. Like the Asian crisis of late 1990s, this crisis also witnessed the merging of banking crisis and currency crisis threatening the financial stability and along with that the most dreaded element of it was the rapid spread of the crisis in global economy like a wild forest fire.[10] Therefore, the economists in the coming years would face the challenge of encapsulating these elements in a unified analytical structure.

Problems in Global Economy Prior to Global Crisis 2008

The global crisis of 2008 was the end result of interplay amongst a host of factors. Different commentators would pick up difference factors and would be inclined to ascribe different ordering to them in terms of importance. But perhaps the common toppers in those lists are two problems, viz., global saving-investment imbalance and the US financial friction.

Global Saving-Investment Imbalance

The persistence of global saving-investment imbalance since the early 1990s has been cited as a proximate cause for the global crisis of 2008 by some international financial institutions (BIS 2009). The imbalance is primarily attributed to the fact that the saving-investment gap (or, excess of saving over investment) is positive in the emerging market economies like China, while it is negative for the developed countries like USA and the Western Europe. The negative saving-investment gap is reflected as current account deficit and the positive saving investment gap is reflected as current account surplus. The equivalence is apparent from the following accounting relationship, which is based on Eq. (3.1):

[10] So much, so that International Monetary Fund was more focused this time with the task of containing the bandwagon effects of the financial crisis within a limited geographical spread rather than relieving the particular countries out of the balance of payments problems.

$$M - X = C + I + G - Y = I - (Y - T - C) - (T - G) \qquad (3.2)$$

where T represents tax, $(Y - T - C)$ represents disposable income $(Y - T)$ minus consumption or private saving and $(T - G)$ represents public saving.[11] Thus, the current account deficit[12] $(M - X)$ measures the excess of aggregate expenditure[13] over income, or investment over saving. A country having a current account deficit signifies the fact that the country concerned has been attracting net foreign investment, while a current account surplus country would be investing abroad in net terms. The accumulation of current account surplus and/or deficits over the years by a country gives its net international investment position.[14]

The continuing current account deficit by USA and other developed countries would mean that the emerging market economies have been investing or exporting their excess saving to capital-rich USA and other developed countries (Bernanke 2005).[15] This is in contrast to the neo-classical position. The neo-classical position would be consistent with a contrary picture where capital-rich countries would have been sending their excess saving to the capital-poor emerging market economies.[16] This prolonged reverse flow of saving from the capital-poor countries to capital-rich countries has been bothering the economists and policy-makers (Cooper 2008; Feldstein 2008). To cite a simple quantitative measurement of the problem, it might be noted that if one sums up the current account deficits of all countries that are running deficits in the world economy, the US deficit accounts for about 70 % of the total.

Current account surpluses, as can be seen from Eq. (3.2), basically mean an excess of national saving (domestic plus foreign saving) over domestic capital investment. A host of factors, *namely*, imperfect arrangements for consumer credit for large purchases, corporate management incentives for retaining rather than distributing earnings, the prospect of lower earnings after retirement, memories of past periods adversity—all contributed toward high levels of saving in current account surplus economies. However, Cooper (2008) added a new dimension to this discourse—dramatic demographic transformation in many countries—marked by increasing longevity and declining natality. The increasing average life

[11] Aggregate saving of the economy is split between private saving and public saving. Thus, the right hand side of Eq. (3.2) represents excess of aggregate investment over aggregate saving, in short, investment over saving.

[12] For $(M - X)$ to qualify as current account deficit both M and X need to be much more loaded variables like M standing for the excess of payments to the rest of the world for goods, services, investment income, and X for unilateral transfers over receipts from the rest of the world for similar items.

[13] Aggregate expenditure is measured by the sum $C + I + G$ in the middle equivalence relation of Eq. (3.2).

[14] Of course, the accumulated current account position needs to be adjusted for valuation changes for this purpose.

[15] Bernanke has coined the phrase 'saving glut' hypothesis to describe the situation.

[16] This might remind one of convergence hypothesis of Baaro and Sala-i-Matin (2003) in growth literature.

expectancy without a corresponding increase in working years has been forcing the household to save more for retirement.[17] Declining natality, on the other hand, tends to reduce investment as less new capital is required to equip the new addition in labor force[18], and relatively scarce labor at home would induce capital-labor substitution in the production process in such a way that domestic returns to capital would fall thus enhancing the attractiveness of investment abroad. Therefore, it is observed that the countries that have the largest trade surpluses like China, Japan, Germany, and the newly rich Asian countries[19] are also countries that are quite far advanced in the demographic transition.

Now, there remains the crucial question: why these savings from the rest of the world flow to the US financial markets. According to Cooper (2008), the answer must lie in the enormous size of the US financial markets as the latter is even larger relative to the rest of the world, accounting for half of the world's marketable securities in 2006. Apart from this, other contributing factors might be—the liquid nature of marketable securities in US markets, secure property rights and speedy and impartial dispute settlement system in US and highly innovative and flexible nature of US economy. Caballero et al. (2006) stressed the fact that shortage of adequate financial assets outside US, especially in emerging market economies, has been pushing the high foreign demand for US securities.[20] However, by heavily staying put in US markets foreign investors have been incurring a risk that a weaker economy would hurt their return measured in home currency, but these investors must be hedged against this exchange rate risk by lucrative returns on their portfolios in US markets (Cooper 2008).

However, the economists are far from unanimous on the question of sustainability of US trade deficit and saving-investment imbalance. Cooper (2008) believes that the high US trade deficit would continue for many more years to come. He reasoned that contrary to the calculations of pessimists, the foreign investments in US markets till now are well below the mark of "no home bias" level of investment.[21] The US economy has been characterized by the unique feature that the total value of its financial assets has risen at a much faster rate than its underlying economy, reflecting massive innovations by its financial sector. This

[17] Precautionary savings should be on the rise as well. Not only the lives of the people are longer, uncertainty prevails over how much longer, given the continuous medical advancement. Moreover, many countries, including India, are facing the uncertainty over future financial viability of their public pension scheme. This is expected to drive people toward more precautionary saving (Cooper 2008).

[18] This will ultimately reduce the demand for schools and housing.

[19] This category includes Hong Kong, Korea, Singapore, Taiwan, etc.

[20] This shortage of market securities outside US was a major factor behind emergence of 'shadow banking' in US (Gorton and Metrick 2012). It will be pointed out later at the appropriate place.

[21] "No home bias" is synonym for perfect financial globalization. "No home bias" level would accommodate foreigners holding 30 % of their financial assets in US, whereas the actual figure, at the end of 2006, is to the tune of 12 % (McKinsey Global Institute 2008).

phenomenon has been whetting the appetite of the investors from around the world for investment in US securities. There is no sign of this appetite to diminish in the near future. Therefore, the vulnerability of the rest of the world to the turmoil in US market would continue.

Feldstein (2008), on the other hand, believes that the current account deficit of US and the global imbalance is not sustainable. A large part of the foreign investment in US markets is by foreign governments, rather than by private individuals. So in a sense, the foreign governments have been financing the trade deficit of US economy. The foreign governments have been doing that because that might help them to *artificially* maintain current account surplus. But the expected future decline in dollar relative to foreign currencies would unleash forces of competition that would turn current account deficits of US into a surplus through a rise in US exports and a decline in its imports. Feldstein (2008) has further argued that the Chinese have been able to mobilize an enormous and growing trade surplus, largely by keeping their currency, the yuan, *artificially* undervalued. So this kind of artificial maneuvering cannot be sustainable. People notice the strain of mercantilism in such a stance by the Chinese (Chinn et al. 2011). However, this viewpoint holds financial factors, rather than real factors like demographic profile as pointed by Cooper (2008), responsible for the global imbalance.

US Financial Friction

In the financial market, the friction arises when there is a cost to one side of a transaction that is not a benefit to the other side (Hall 2010). The difference between an equity market and a debt market from the friction point of view is that an equity investor has to monitor an investment in all states of the world, while a debt lender only has to monitor if the borrower fails to repay in full—this proposition is known as "costly state verification" model of Townsend (1979). Thus, the pricing of a loan includes the charge for the likelihood of a default—the expected verification cost—constituting the financial friction. Now, the borrower's own wealth and the random distribution of events like low profit rendering the repayment of loan in full impossible determine the likelihood of borrower default. So the events leading to borrower default and low profit result in widening of financial friction. In the financial crisis of 2008 all institutions and businesses that held real estate and real-estate-linked financial assets experienced severe loss in wealth when real-estate prices declined. This resulted in a large decrease in the values of the asset holdings of financial institutions and aggravated the financial friction in the economy, particularly in the real estate-based financial companies relative to non-financial companies. Consequently, the credit spreads[22] widened and credit rationing became norm of the day. The rising credit spreads in the crisis

[22] Credit spreads refer to the difference between private borrowing rates and the rates the federal government paid.

involves a "flight to quality"—investors suddenly prefer high-quality debt, notably that of the US Government. This resulted in the diminished ability of the business firms to finance all types of investment. This resulted in the drastic reduction in the investment component (I) in Eq. (3.1), while the other components of GDP remained roughly constant (Hall 2010).

Now, a question might be asked why the fall in the real estate values in 2001 in US economy did not trigger the credit spreads of this proportion. The possible answer might be that the assets that decline in value during that episode were chiefly business assets, concentrated mainly in high tech that uses little debt finance and thus has little leverage. The financial friction did not bite much as those assets were held mostly in equity, not in debt. Business equities generally stay put in large portfolios of rich families, in mutual funds, in endowments and these entities rarely borrow against their holdings. As against this, the real estate has high leverage. Most homeowners borrow almost up to the entire price of the house when buying a house and they become unlevered only when they live in the house and pay down the mortgage. Added to this, the phenomenon of the multiple added levels of leverage among financial institutions holding real estate assets made the economy many times more vulnerable to risk compared to 2001 episode (Hall 2010).

Though the declining housing prices have been blamed for generating financial friction, Buiter (2008) has pointed out that there is no aggregate wealth effect of a decline in housing prices. The household sector in aggregate is both the owner of the housing stock and the consumer of services rendered by it. A fall in the housing price reduces the value of house as an asset, but at the same time also reduces the cost of buying the flow of housing services by exactly the same amount—canceling the two opposite effects. Thus, this thesis asserts that the effect of declining housing prices on the aggregate demand of the economy should be very small, if at all.

However, larger effects are possible if it is reckoned that the losses resulting from the collapse of housing prices were disproportionately concentrated in certain financial institutions 'which play a role in the allocation of resources that cannot easily be replaced by those to whom wealth was redistributed' (Woodford 2010).

Run-Up to the Crisis

The financial crisis had been building up since the first half of 2007. It began with the popping of housing bubble in 2006—whereby the house prices flattened and then started declining. Refinancing a mortgage was becoming difficult and this was resulting in increasing mortgage delinquency in the subprime market (Lo 2012). The problems in the subprime market became increasingly visible and included the failure of several subprime originators. The financial crisis was a bank run, but in sectors of the money markets where financial institutions provided bank-like debt products to institutional investors. These financial institutions were mostly shadow banks[23] (Gorton and Metrick 2012). Bernanke (2010) said:

> Before the crisis, the shadow banking system had come to play a major role in global finance; with hindsight, we can see that shadow banking was also the source of key vulnerabilities.

The main vulnerability was short-term debt, mostly repurchase agreements or repo[24] and commercial papers. The markets for these instruments were largely unregulated (Bernanke 2010).

In the run-up to the crisis the US economy witnessed the credit boom, fueled by the increase in the issuance of asset-backed securities, particularly mortgage-backed securities.[25] This led to the mushrooming of the shadow banking system. The traditional banking became less profitable in the face of money market mutual funds and junk bonds. Securitization[26] experienced tremendous growth during this phase as a response to this situation. Alongside this, institutional cash pools, who pool cash balances from all subsidiaries worldwide in case of global corporations or all funds in case of asset manager, stormed the markets. These institutional cash pools had an appetite for investing in safe assets. But there were not enough safe assets, *namely*, US treasuries, in the market because the foreign institutional investors held a large amount of US treasuries. This is the direct consequence of global imbalance on US financial markets.[27] The shadow banking system rose to fill this gap. They rampantly issued repo and commercial papers. This led to the widening of the credit spread. The subprime failure and declining housing price put housing under pressure on the asset back securities markets. When this pressure ultimately claimed Lehman Brothers as a victim, the overstressed interbank market and credit market collapsed, pushing the crisis to its peak level (Gorton and Metrick 2012).

From the above analysis, it is clear that anomalies in the US financial market alone would not suffice to trigger the crisis of 2008 dimension had it not been in collision with the global imbalance factor.

---

[23] Subprime organizations also acted as shadow banks.

[24] A repo transaction is a collateralized deposit in a bank. The depositor or lender deposits money in the bank for short-term, usually overnight. The bank promises to pay the overnight repo rate on the deposited money. To ensure the safety of the deposit, the bank provides collateral that the depositor takes possession of. If the bank fails, then the depositor can sell the collateral to recover the value of the deposit (Gorton and Metrick 2012).

[25] The structured investment vehicles in US market were basically portfolios based on these kinds of mortgage-backed securities. But asymmetric information problem was lurking there in great measure. At each step in the chain of such structured investment vehicles one side knew significantly less than the other about the underlying structure of securities. At the top layer of the chain, an investor knows absolutely nothing about the hundreds of thousands of mortgages several layers below the derivative being traded (Lo 2012).

[26] Securitization means the sale of loan pools to special purpose vehicles that finance the purchase of loan pools via issuance of asset-backed securities in the capital market. Securitization is off-balance sheet financing for banks and other financial intermediaries. Securitization practice began in 1990s, but its growth just before the crisis was phenomenal.

[27] The nature and causes of global imbalance has already been discussed.

Sustainability in the Face of Crisis

Each crisis offers some valuable lessons to learn. Many economic arrangements and policy paradigms are topsy-turvied by the dramatic turn of the events. When Mexico announced its policy of debt moratorium in August, 1982, the obsession of development policy-makers with 'import substitution' began to be questioned. The craze about pegging the domestic currencies to US dollar met an untimely fatality in December, 2000 with the Argentinean peso crisis. Thus, one might be legitimately wondering about the lessons that could be learnt from the financial crisis of 2008—what ideas might be at stake, what ideas might be shinning tomorrow?

US Federal Reserve tried to battle out the crisis with a host of policy tools—both traditional and unconventional. They tried to increase the liquidity in the market by cutting the interest rates[28] on a massive scale.[29] But this traditional measure failed to generate the expected results. This forced the Federal Reserve to undertake a number of unconventional measures like lending to the 'weak' financial institutions directly in terms of cash and securities on the basis of even doubtful collateral. No doubt, this kind of measures was in utter desperation to bring some order in the chaotic financial markets. But the costs of these measures were enormous.[30] The size of the costs has been forcing the Federal Reserve to ponder what would happen if Federal Reserve runs out of capacity through its engagement in such short-term transactions. The question remains unanswered like many other questions so as to what would be the response of Federal Reserve when the prices of risky assets decline sharply and the associated risk premia increase (Cecchetti 2008). Even the conservative legislators in US have raised the issue of jeopardizing the inter-generational equity through the above measures. Surely, the massive costs of these rescue operations would spill over to the next few generations. So the question of sustainability still bothers US.

For the policy-makers in the European Union the task of fighting the financial crisis was relatively easier. They primarily depended on the traditional monetary policy instruments like controlling the growth of monetary aggregates. Prior to the onset of the crisis, the European Union was beset with the problem of rapid monetary expansion. So the financial crisis provided the policy-makers an opportunity to dampen the excess 'bubbles' in the economy through a rather natural means (Gali 2010). However, some countries in the European Union messed up on the fiscal front. These governments accumulated huge sovereign debts with low capital to debt ratios,[31] e.g., Greece (Kim 2012). This has raised serious question mark over the sustainability of such debts.

---

[28] Both federal funds rate and prime lending rate were used.

[29] Federal Reserve cut interest rate by 325 basis points in a span of 8 months during the crisis (Cecchetti 2008).

[30] By one estimate, by the end of May, 2008 Federal Reserve had committed almost two-third of its $900 billion balance sheet (Cecchetti 2008).

The East Asian countries like South Korea suffered from the ripple-effects of the financial crisis. The negative effects originated mainly from the problem of loss of confidence of foreign investors. The high loan-to-deposit ratios of the Korean banks implied larger portion of those loans had to be financed by capital from abroad. But given the financial turmoil it would be difficult for the Korean banks to raise such huge funds. The deteriorating current account deficit of the Korean economy had been another source of concern for the investors. There was a sharp deceleration in the growth rate of the Korean economy during the period of crisis. However, the economy bounced back to normalcy from 2010 onwards (Kim 2012). Therefore, the issue of sustainability of growth was not that much striking for Korea.

India also came under the spell of global financial crisis like other emerging market economies. This proved the 'decoupling hypothesis'[32] wrong. However, it would be misleading to blame the global financial crisis entirely responsible for India's plight. The deceleration in India's growth rate was set in much before the onset of the global financial crisis as is evident in the macroeconomic data. The reasons for the slowdown were mainly (i) sharp reduction in the private capital formation, and (ii) drastic fall in exports (Rakshit 2009). Indian banks were relatively less affected by the crisis owing to their relatively smaller exposure to the US asset-backed securities and structured derivative products like collateralized debt obligations. However, India's capital inflows, particularly external commercial borrowings and foreign institutional investments, were severely impacted as fallout of the crisis. Many domestic producers use the external commercial borrowing channel to finance their purchase of machinery and equipments. The withdrawal of foreign institutional investments dampened the propensity to invest by the producers. Coupled with the above, the bearish trend among the Indian banks to extend credit to the domestic traders has contributed to the dryness on the investment front. The Reserve Bank of India has followed a 'dear' money policy to combat the spiraling domestic price level throughout the period of crisis. All these factors acting together have been prolonging the recovery of the Indian economy in the post-crisis period (Rakshit 2009). Therefore, the sustainability of growth process in India was not at stake due to the crisis.

It is apparent from the above discussion that policy responses to the crises varied between the developed western countries and the emerging market economies. The developed countries had been in a better situation than their

[31] However, the vulnerability of these economies has so far been contained within their national boundaries. If the sovereign debt crises of these economies explode to become a banking crisis, it might have the potential of triggering an even greater crisis than the global financial crisis of 2008. In the European debt crisis what is at stake is the confidence of the public in the sovereign debt of the governments (Kim 2012).

[32] 'Decoupling hypothesis' was put forward by the financial market experts to point out that the correlation between business cycles of the developed countries like USA and Western Europe and those of emerging market economies, especially the larger ones like China and India have become weaker over the years despite the increased globalization of the latter economies.

underdeveloped counterparts in terms of institutional set-up and the governance structure. The development economics prior to the crisis of 2008 were focusing on the 'institutions rule' as the predominant ethos (Acemoglu et. al. 2000), However, the crisis of 2008 has taught us that 'good institutions' are not enough. The syndrome of "elite capture of institutions" was prevalent in US economy prior to the crisis (Pritchett 2011). Therefore, there is a need for ensuring the robustness of the institutions under stress.

India is widely known to be enjoying the 'demographic dividends'[33] from its relatively young population. As has been mentioned earlier, the demographic profiles of these Asian economies have been behind the global saving-investment imbalance, which was a proximate cause behind the global financial crisis of 2008. But the demographic dividends are not amenable to policy maneuvering. The countries should try to adjust other macroeconomic variables to shield against the ill effects of such phenomenon. The institutions in India are well developed (Subramanium 2009). Though they have not yet been tested against the financial crisis, it might reasonably be expected that they would rise up to the mark in the hours of need. However, it should be kept in mind that financial crises have proved again and again that no amount of preparation is sufficient.

Conclusion

The impact of the financial crises has always pushed the economists to introspection and reckoning. It is no exception this time.[34] However, several narratives and perspectives are now available on the 2008 crisis. Lo (2012) has made an excellent survey of the different strains of thoughts on the crisis.[35] He has categorized the literature under academic accounts and journalistic accounts. The academic accounts of the crisis exhibit the most heterogeneous streams of thoughts by the authors, but at the same time they throw direction for future research. On the other hand, though journalistic accounts are complementary in many ways to their academic counterparts, they were more like the accounts by war correspondents as they stressed more on 'campaigns, battles, and exceptional acts of courage and cowardice'.

The present state of academic research is not adequate to explain the crisis of 2008. Only time may tell us whether the birth of a fourth generation (in sequel to

[33] There are more people in India in the age group of 15–60 than in the age groups like 0–14 and above 60 years. This has made the dependency ratio (number of people in the age groups 0–14 and above 60 years over the total population) relatively lower for India compared to many, particularly the western developed countries.

[34] The financial crisis of 2008 has left a deep imprint on the economics teaching also. It is pronounced in the inclusion of the separate chapters on crisis 2008 in the two popular graduate level macroeconomic textbooks (Romer 2011; Wickens 2012).

[35] Lo (2012) has surveyed 21 books on the crisis published in the recent years.

the third generation) crisis theories is in the offing. Macroeconomists are critical about the currently fashionable dynamic stochastic general equilibrium models and value at risks models. But alternative models are yet to come up. Woodford (2010) and Hall (2010) tried to provide the sketches of full-fledged macroeconomic models that would be robust in the face of crisis like that happened in 2008. Hall (2010) has tried to address the question why the real macroeconomic variables take so much time to recover in his model. Woodford (2010) modeled the financial friction aspect of the crisis. He stressed on the financial stability being the major plunk of the future monetary policy so as to control future financial crises.

The school of 'institutional rule' has been the dominant stance in the recent development thinking. It was presumed that good institutions were sufficient in themselves so much, so that "they would either muddle along even without particularly good policies or that these institutions themselves would eventually produce the equivalent of good policies in practice" (Pritchett 2011, 137). However, the global financial crises have taught us that it is dangerous to put the institutions under stress, particularly in the development context of premature load bearing (Pritchett 2011).

The financial crises generally end up in raising more questions than answers. More frustrating might be that the economists disagree not only on the causes and consequences of the crises; they are far from unanimity even on the facts surrounding the crises (Lo 2012). Many economists used to take pride in the fact that financial economics had been developing into a faultless scientific discipline, "but complex events like the financial crisis suggest that this conceit may be more wishful thinking than reality". It is as if we are in a world of Kurosawa's famous film *Rashomon*, where we never agree on a single narrative that explains all the facts. Probably, such a grand narrative is non-existent in the post-modern world where we learn to live with financial crises of global dimension.

References

Acemoglu D, Johnson S, Robinson J (2000) The colonial origins of comparative development: an empirical investigation. Am Econ Rev 91(5):1369–1401

Agenor P-R, Montiel PJ (1999) Development macroeconomics, 2nd edn. Princeton University Press, New Jersey

Anzuini A, Gandolfo G (2000) Currency crises and speculative attacks: a comparison between two different forecasting approaches. CIDEI working paper no. 60, University of Rome

Bank of International Settlement (BIS) (2009) 79th Annual Report. Basle, Switzerland

Barro R, Sala-i-Martin X (2003) Economic growth. MIT Press, Massachusetts

Bernanke BS (2005) The global saving glut and the US current account deficit. The Sandridge Lecture, 14 April, 2005

Bernanke BS (2010) Causes of the recent financial and economic crisis. Testimony before the Financial Crisis Inquiry Commission, Washington DC, September 2, 2010

Buiter WH (2008) Housing wealth isn't wealth. NBER working paper 14204

Caballero RJ, Farhi E, Gourinchas P-O (2006) An equilibrium model of global imbalances and low interest rates. NBER working paper 11996

Cecchetti SG (2008) Crisis and responses: the Federal Reserve and the financial crisis of 2007–2008. NBER working paper 14134
Chang R, Velasco A (1998) Financial crisis in emerging markets: a canonical model. NBER working paper 6606
Chinn MD, Eichengreen BJ, Ito H (2011) A forensic analysis of global imbalances. NBER working paper 17513
Cooper RN (2008) Global imbalances: globalisation, demography, and sustainability. J Econ Perspect 22(3):93–112
Corsetti G, Pesenti P, Roubini N (1999) Paper tigers? A model of the Asian crisis. Eur Econ Rev 43:1211–1236
Eichengreen B, Rose AK, Wyplosz C (1996) Contagious currency crises. Scand J Econ 98:463–484
Feldstein M (2008) Resolving the global imbalance: the dollar and the U.S. saving rate. J Econ Perspect 22(3):113–126
Flood RP, Garber PM (1984) Collapsing exchange-rate regimes: some linear examples. J Int Econ 17:1–13
Friedman M (1953) The case for flexible exchange rates. In: Friedman M (ed) Essays in positive economics. University of Chicago Press, Chicago
Gali R (2010) The monetary pillar and the great financial crisis. Unpublished, CREI
Gandolfo G (2002) International finance and open-economy macro-economics. Springer-Verlag, Heidelberg
Gerlach S, Smets F (1995) Contagious speculative attacks. Eur J Polit Econ 11:45–63
Gorton G, Metrick A (2012) Getting up to speed on the financial crisis: a one-weekend_reader's guide. J Econ Lit 50(1):125–150
Hall RE (2010) Why does the economy fall to pieces after a financial crisis? J Econ Perspect 24(4):3–20
Kemp MC (1963) Profitability and price stability. Rev Econ Stat 45:185–189
Keynes JM (1933) The general theory of employment, interest and money. Macmillan & Co., London
Kim IJ (2012) The global financial crisis and the challenges of the Korean economy. Seoul J Econ 25(3):339–366
Krugman P (1979) A model of balance-of-payments crises. J Money Credit Banking 11:311–325
Krugman P (1999) Balance sheets, the transfer problem and financial crises. In: Isard P, Razin A, Rose AK (eds) International finance and financial crises: essays in honor of Robert P. Flood, Jr. Kluwer, Massachusetts
Ljungqvist L (1992) Destabilizing exchange rate speculation: a counterexample to Milton Friedman. Seminar Paper 125, Institute for International Economic Studies, Stockholm
Lo AW (2012) Reading about the financial crisis: a twenty-one-book review. J Econ Lit 50(1):151–178
Mankiw G (2006) The macroeconomist as scientist and engineer. J Econ Perspect 20(4):29–46
McKinsey Global Institute (2008) Mapping global capital markets. San Francisco, USA
McKinnon RI (1983) The J curve, stabilizing speculation and capital constraints on foreign exchange dealers. In: Bigman D, Taya T (eds) Exchange rate and trade instability: causes, consequences and remedies. Ballinger, Massachusetts
Pritchett L (2011) The problem with paper tigers: development lessons from the financial crisis of 2008. In: Birdsall N, Fukuyama F (eds) New ideas on development after the financial crisis. John Hopkins University Press, Baltimore
Rakshit M (2009) India amidst the global crisis. Econ Political Wkly XLIV(13):94–106
Romer D (2011) Advanced macroeconomics, 4th edn. McGrow Hills/Irwin, New York
Sachs J, Tornell A, Velasco A (1996) The Mexican peso crisis: sudden death or death foretold? J Int Econ 41:265–283
Shleifer A, Summers L (1990) The noise trader approach to finance. J Econ Perspect 4(2):19–33
Subramanium A (2009) India's turn: understanding the economic transformation. Oxford University Press, Oxford

Tirole J (2005) The theory of corporate finance. Princeton University Press, Princeton
Townsend RM (1979) Optimal contracts and competitive markets with costly state verification. J Econ Theor 21(2):265–293
Wickens M (2012) Macroeconomic theory: a dynamic general equilibrium approach, 2nd edn. Princeton University Press, Princeton
Williamson J (1976) The benefits and costs of an international nonsystem. In: Bernstein EM et al (eds) Reflections on Jamaica, essays in international finance no. 115. International Finance Section, Princeton University
Woodford M (2010) Financial intermediation and macroeconomic analysis. J Econ Perspect 24(4):21–44

Chapter 4
Volatility, Long Memory, and Chaos: A Discussion on some "Stylized Facts" in Financial Markets with a Focus on High Frequency Data

Amitava Sarkar, Gagari Chakrabarti and Chitrakalpa Sen

Introduction

Starting from the "Tulip Mania' in the seventeenth century, financial sector crises have come in waves and in many different guises. While some of these remained confined to the regional boundaries, some acquired global dimension with the ultimate devastating impact on the real economy. The fact that the global economy has collapsed many a time following a financial panic has instigated the researchers, particularly after the recent global financial melt down, to explore the dynamics of global financial markets: the old issue of 'finance-growth nexus' is resurrected once again. The traditional school of the literature, however, is bifurcated on the issue. While one school perceives finance to be a 'side show' of growth, others believe in the considerable control that financial markets exercise on the real sector. The true nature and direction of the causality, however, is yet to be exposed. More recently, a parallel school of thought has developed that concentrate on the endogenous factors that generate dynamics within the stock market independent of the real sector. This growing body of the literature conjectures the global financial markets to be mostly deterministic, and in some cases, chaotic in nature. The markets are characterized by nonperiodic cycles and trends where volatility and fluctuations generate endogenously. The global markets, thus, are supposed to be inherently instable, or at best, stable on knife edge. Cycles and crashes are manifestations of this inherent instability. There is no determinate

A. Sarkar
Indian Statistical Institute, Kolkata 700203, India
e-mail: amitavasarkar@hotmail.com

G. Chakrabarti (✉)
Department of Economics, Presidency University, Kolkata 700073, India
e-mail: gagari_1974@yahoo.com

C. Sen
Auro University, Surat, Gujarat, India
e-mail: chitrakalpa@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_4, © Springer India 2013

equilibrium in the market and no external shocks would be required to gear financial crises at regular interval. The implications of these are tremendous. A chaotic financial market puts efficient market hypothesis on trial, renders traditional asset pricing models useless, and makes long-term forecasting less reliable. The most devastating implication of the fact that the financial markets implode from within is perhaps for the developing markets as it makes government intervention ineffective. Hence, an introspection of financial market dynamics following this rather offbeat line of thought could be a researcher's delight. The results obtained might lead to a reframing of old ideas and beliefs regarding financial market dynamics, boom and bust cycle, and predictability of financial crashes.

These issues are of immense significance for developing markets and hence for India. In this era of global financial integrity, it would be very difficult for any country to remain decoupled from the global market. If the financial market cycles are self-generating, financial crises turn out to be real 'black swans': events that are fatal but unpredictable. Now placed firmly within the chaotic financial markets, India would be susceptible to financial crises, the causes of which might be difficult to predict and the effects of which would be impossible to be controlled by effective government intervention. Even as policies would fail to protect the economy from crisis, they may not be able to pull it out of crisis either.

This chapter discusses some of the technical issues in this new field of financial market dynamics. Some of these tools are used frequently to explore the possible inherent instable nature of financial time series data and their immense implications at theoretical and policy levels. Our exploration starts from an analysis of stylized facts of high frequency time series data and is then extended to consider the inherent instability or otherwise of the Indian financial market in recent years.

High frequency financial time series are characterized by some unique regularity, known as the stylized facts. Some of the major stylized facts are:

1. *Conditional volatility*: Almost all high frequency time series are characterized by conditional volatility, that is, the current volatility is dependent on previous period's volatility as well as the news about volatility in the past.
2. *Volatility persistence or long memory*: The presence of long-term memory indicates that when a shock is propagated into the market, it does not die down immediately, but lingers for a very long time. Hence, past price changes can be used as significant information for the prediction of future price changes (Mantegna and Stanley 2000).
3. *Fat tails*: The unconditional distribution of returns follows a power law or positive excess kurtosis. The distributions are approximately bell shaped but assign more than normal probability to events in the center (more peaked) and at the extremes (exhibit heavy tails).
4. *Volatility clustering*: Different measures of volatility display a positive autocorrelation over several days, which quantify the fact that high and low volatility events tend to cluster in time; in other words, periods of intense fluctuations and mild fluctuations tend to cluster together.

5. *Nonlinearity and chaos*: Financial time series are, mostly, nonlinear in nature. Often, these series are characterized by chaotic dynamics: they look random, although they are inherently deterministic in nature.

The stylized facts are significant for more than one reason (Mukherjee et al. 2011). The stylized facts clearly show that the normal distribution is inadequate in describing the distribution of the financial time series. It is necessary to invoke a distribution which assigns more probability to the tails of the distribution, and thus immediately alludes to power-law type distributions (a strong signature of complexity). Second, the appearance of these facts in apparently unrelated and diverse systems suggests the existence of some common underlying mechanism responsible for the appearance of these facts (collective behavior and critical phenomena typical of complex systems). Further, the existence of these apparently disparate facts forces us to think about an integrated account of these facts, pointing once again toward a common mechanism to explain these facts in a unified manner. As an example, one finds that explaining volatility clustering and long memory will most likely explain the occurrence of the fat tails (Rickles 2008).

In the following section, we review some major theoretical and empirical works discussing these above stylized facts, focusing mainly on volatility, long memory, and chaos. We restrict our discussion within high-frequency time series data, especially in foreign exchange market and stock market.

Conditional Volatility in Financial Market

Conditional volatility has always been a favorite topic for economists. To capture the conditional volatility, the most widely used technique is the ARCH school of models. The ARCH of the Autoregressive Conditional Heteroscedasticity model was first introduced by Engle (1982) and farther developed into the GARCH or the Generalized Autoregressive Conditional Heteroscedasticity model by Bollerslev (1986). Engle (1990) incorporated into the model explicitly the nonlinear asymmetry and the model was called NAGARCH (Non linear Asymmetric GARCH) or Power GARCH (PARCH). The exponential general autoregressive conditional heteroskedastic (EGARCH) model by Nelson (1991) is another widely used GARCH model to capture asymmetry. The GJR-GARCH by Glosten, Jagannathan and Runkle (1993), and TGARCH or TARCH (Threshold GARCH model, a slight modification of the GJR GARCH) also is able to capture the asymmetry. The component GARCH (CGARCH) model by Engle and Lee (1999) is used to distinguish between the long run and short run time path of volatility. Apart from these, there are many other ARCH school models, IGARCH (Integrated GARCH) by Engle and Bollerslev (1986), GARCH-M (GARCH in mean) by Engle, Lilien and Robins (1987), APARCH (Asymmetric power ARCH) by Ding and Granger (1996), AVGARCH (Absolute value GARCH) of Taylor (1986) and Schwert (1989) are some of the popular types. Some of the widely used models in the literature are briefly discussed below.

Few Models to Test for the Presence of Conditional Volatility

ARCH Model

The q-th order ARCH model can be written as -

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{4.1}$$

or,

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 \tag{4.2}$$

L denotes the lag operator.

where σ_t is the conditional standard deviation, $\omega > 0$ and $\alpha_i \geq 0$ for $i = 1,\ldots,$ q for σ_t^2 to be positive and the model to be well defined.

The ARCH model effectively captures any presence of volatility clustering. A large shock tends to increase the conditional variance of the current period. As $\omega > 0$ and $\alpha_i \geq 0$, a large (small) shock tends to be followed by a large (small) σ_t^2. The direction of the shock however, i.e., negative or positive in sign does not matter in this regard because on the right hand side, only the square terms are considered.

This, however, is one of the weaknesses of the ARCH model. In reality, empirical studies have shown that the market does not react to negative and positive shocks in a similar fashion. This is not the only weakness of an ARCH model. The ARCH model tends to over predict the volatility as it reacts slowly to large isolated shocks. Also, although this model diagnoses the volatility in the system, it does not explain the source of it (Tsay 2001).

Another drawback of the ARCH model is that as more parameters are added in the variance equation, chances are high that their estimated values turn out to be negative, i.e., the nonnegativity conditions are violated.

In many empirical applications of ARCH (q) models, long lag length in the conditional variance equation is often required. An alternative and more flexible lag structure and longer memory is provided by the Generalized Autoregressive Conditional Heteroscedasticity or the GARCH model (Bollerslev 1986).

GARCH Model

According to Bollerslev (1986, 308), "the extension of the ARCH process to the GARCH process bears much resemblance to the extension of the standard time series AR process to the general ARMA process and… permits a more parsimonious description in many situations". The GARCH model includes past conditional variances in the variance equation.

The GARCH (*q*,*p*) model can be specified as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \quad \text{or,} \quad \sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2 \tag{4.3}$$

Where q is the order of autoregressive GARCH terms and p is the order of moving average ARCH terms.

And the restrictions are, $p \geq 0$, $q > 0$, $\omega > 0$, $\alpha_i \geq 0, i = 1, \ldots, q$, and $\beta_j \geq 0, j = 1, \ldots, p$.

The conditional variance at period t (σ_t^2) is not only dependent on the squared errors from the previous period, but also on the past conditional variance. When $\beta = 0$, the GARCH process is reduced to an ARCH process.

But, the GARCH model has some serious limitations. Nelson (1991) criticized the GARCH model on three grounds.

(i) First, volatility does not react similarly to good and bad news as assumed in the GARCH model. Volatility tends to increase more in response to a bad news than to good news of the same magnitude. This "leverage effect" was first identified by Black (1976). This is not captured by the GARCH model. As the conditional volatility is dependent on the square of the shock (innovation) term, the sign of the shock does not matter, only the magnitude does. This calls for an appropriate model that captures the asymmetric movement of the conditional variance to positive and negative shocks.

(ii) The second drawback in the GARCH process lies in the nonnegativity restriction. It means that any change in ε_t^2 always leads to an increase in the conditional volatility of the next period, thus ruling out any random oscillatory movement. These constraints unduly restrict the dynamics of the conditional variance process.

(iii) A third limitation of this model lies in the interpretation of volatility persistence of GARCH model. The GARCH model cannot properly explain the volatility persistence, especially in the cases where a shock in the time series persists for a long time.

Also, Alexander (2001) pointed out that if volatility is higher after a negative shock than it is after a positive one, the autocorrelation between last period and if current period's squared return will be large and negative. i.e., $\frac{\sum_{t=2}^{T} y_t^2 y_{t-1}}{\sqrt{\sum_{t=2}^{T} r_t^4 \sum_{t=2}^{T} r_{t-1}^2}}$ is negative and the result from the Box–Pierce test is significantly different from zero, the volatility clustering cannot be captured by a symmetric GARCH model. It shows an asymmetry in the volatility clustering. For that, an asymmetric GARCH model needs to be used.

EGARCH Model

Nelson (1991) first proposed an Exponential GARCH model that meets these limitations. The EGARCH (1,1) model can be specified as:

$$\log\left(\sigma_t^2\right) = \omega + \alpha(|z_{t-1}| - E(|z_{t-1}|)) + \gamma z_{t-1} + \beta\log(\sigma_{t-1}^2) \tag{4.4}$$

where $\varepsilon_{t-1} = \sigma_{t-1} z_{t-1}$

An EGARCH (q,p) model can thus be expressed as:

$$\log\left(\sigma_t^2\right) = \omega + \sum_{j=1}^{p} \beta_j \log\left(\sigma_{t-j}^2\right) + \sum_{i=1}^{q} \alpha_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}} - E\left(\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right)\right| + \sum_{k=1}^{r} \gamma_k \frac{\varepsilon_{t-k}}{\sigma_{t-k}} \tag{4.5}$$

The dependent variable is no longer the conditional variance; it is now the log of conditional variance. Hence, the leverage effect is exponential rather than quadratic in the EGARCH model.

This model has several improvements over the GARCH model. The EGARCH model overcomes the most important limitation of the GARCH model by incorporating the leverage effect. If $\alpha > 0$ and $\gamma = 0$, the innovation in $\log\left(\sigma_t^2\right)$ is positive (negative) when z_{t-1} is larger (smaller) than its expected value. And if $\alpha = 0$ and $\gamma < 0$, the innovation in $\log\left(\sigma_t^2\right)$ is positive (negative) when z_{t-1} is negative (positive). Another significant improvement of the EGARCH process is that it contains no inequality constraint, and by parameterizing the $\log\left(\sigma_t^2\right)$ can take negative value so there are fewer restrictions on the model. Lastly, the EGARCH process can capture volatility persistence quite effectively. $\log\left(\sigma_t^2\right)$ can easily be checked for volatility persistence by looking at the stationarity and ergodicity conditions. However, the EGARCH model is also not free from its drawbacks. This model is difficult to use for there is no analytic form for the volatility term structure.

TARCH Model

An alternative process for modeling the leverage effect was proposed independently by Glosten, Jagannathan and Runkle (1993) and Zakoian (1994). The threshold ARCH or TARCH (also known as GJR-GARCH) is a modification of GARCH process with a threshold term included. The model is of the following form:

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{k=1}^{r} \gamma_k \varepsilon_{t-k}^2 I_{t-k} \tag{4.6}$$

Where $I_{t-k} = \begin{cases} 1 \; if \; \varepsilon_{t-i} < 0 \\ 0 \; if \; \varepsilon_{t-i} \geq 0 \end{cases}$ where I is the indicator function

i.e. good news ($\varepsilon_{t-i} > 0$) and bad news ($\varepsilon_{t-i} < 0$) have asymmetric effect on the conditional variance. For a good news, i.e. $\varepsilon_{t-i} > 0$, the impact on σ_t^2 is of magnitude $\alpha_i \varepsilon_{t-i}^2$. And for a bad news, i.e. $\varepsilon_{t-i} < 0$, the impact is of magnitude $(\alpha_i + \gamma_i)\varepsilon_{t-i}^2$. So for a positive γ_i,a bad news increases volatility and a leverage effect is said to exist. For $\gamma_i \neq 0$, the news impact is asymmetric.

Component GARCH (CGARCH) Model

The component GARCH or the CGARCH model, proposed by Ding and Granger (1996) and modified by Engle and Lee (1999) is a better model for capturing mean reversion, because it models mean reversion to a time varying level m_t. Moreover, it incorporates two components, one is the long run or the permanent component and another is the noisier short run or the transitory component. This model is preferred for its ability to capture both short and long run volatility.

Mean reversion is said to exist if a time series data moves toward its long run average overtime. The GARCH (1, 1) model can be used to capture this mean reverting property. The conditional variance of GARCH (1, 1) model can be expressed as –

$$(\sigma_t^2 - \sigma^2) = \alpha\left(\varepsilon_{t-1}^2 - \sigma^2\right) + \beta(\sigma_{t-1}^2 - \sigma^2) \tag{4.7}$$

$\sigma^2 = \frac{\omega}{1-\alpha-\beta}$ is the unconditional variance.

This reconfigured GARCH model explains mean reversion to a time invariant constant ω.

The component GARCH model can be defined as follows:

$$\sigma_t^2 - m_t = \alpha\left(\varepsilon_{t-1}^2 - m_{t-1}\right) + \beta(\sigma_{t-1}^2 - m_{t-1}) \tag{4.8}$$

and

$$m_t = \omega + \rho(m_{t-1} - \omega) + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2) \tag{4.9}$$

(4.8) is the transitory equation. The transitory component $\sigma_t^2 - m_t$ converges to zero with powers of $\alpha + \beta$. (4.9) parameterizes the corresponding long run variance. It is also called the permanent equation and the permanent component m_t converges to ω with powers of ρ. (4.8) and (4.9) can be combined together as:

$$\sigma_t^2 = (1 - \alpha - \beta)(1 - \rho)\omega + (\alpha + \phi)\varepsilon_{t-1}^2 - (\alpha\rho + (\alpha + \beta)\phi)\varepsilon_{t-2}^2 + (\beta - \phi)\sigma_{t-1}^2 - (\beta\rho - (\alpha + \beta)\phi)\sigma_{t-1}^2 \tag{4.10}$$

with constraints on the coefficients to ensure a positive conditional variance. (4.11) is nothing but a GARCH (2,2) model.

Testing for the Presence of Conditional Volatility in Financial Market

Studies in foreign exchange market volatility, although quite large in number, fall well short of the vast body of work in the context of stock market. Volatility in foreign exchange market is a common phenomenon. In the modern era, when the global markets are financially well integrated, exchange rate volatility is present in almost all exchange rates. Some of the most significant works are discussed below.

Chou (1988) used the GARCH framework to investigate the volatility persistence and the changing equity premium in the US. His results showed a high degree of volatility persistence in the market. Engle and Ng (1993) in their seminar work, introduced the news impact curve to estimate the impact of news on volatility estimates. The study fitted several models to daily Japanese stock returns from 1980 to 1988. TARCH and EGARCH turned out to be the two best models to forecast volatility. Rabemananjara and Zakoian (1993) found support for volatility in French stock market using a TARCH model. In the context of India, Sarkar, Chakrabarti and Sen (2009) provided an analysis of Indian stock market volatility dynamics. Sen (2012) explored the conditional volatility in foreign exchange market. The study considered nine currencies across the globe against Indian Rupee and investigated the nature of volatility and regime switch over a period of 13 years from 1998 to 2010.

Long Memory in Financial Market: Theory and Empirics

Long Memory and its Implication for High Frequency Financial Time Series Data

The presence of significant volatility clustering in most of financial market returns suggests these returns to be independent across time. The nature of such interdependence, particularly whether this is 'short term' or 'long term' has been the object of many empirical studies. The idea that financial asset returns, particularly stock market returns could be characterized by the presence of long-term dependence was first suggested by Mandelbrot in 1971. The long-term dependence in a series, better known as 'long memory', describes the correlation structure of the series at long lags. If a series exhibits long memory (or the "biased random walk"), there is persistent temporal dependence even between distant observations. Such series will be characterized by significant, nonperiodic cyclical patterns. Presence of long memory in asset pricing dynamics would persuade us to conclude in favor of the denial of the Efficient Market Hypothesis at least in its weak form. This is particularly because presence of long memory would imply

nonlinear dependence in the first moment of the distribution. Resultantly, strong predictable pattern could be traced in the asset return series dynamics that might be exploited by the investors to earn substantial profit from the financial market. From the point of view of the market itself, in a financial market that has long memory, a shock at a particular point in time will remain (although in a decaying fashion) in the system affecting the future outcomes. Financial crises thus might have their origin rooted in distant past. These might also raise issues regarding linear modeling, forecasting, statistical testing of pricing models based on standard statistical methods, and theoretical and econometric modeling of asset pricing.

Tests for Long Memory

Long memory has long been persistent in the areas of natural sciences and thus, not surprisingly, the first test for long memory owes its origin to a natural scientist, Harold Edwin Hurst (1951). He first designed a test for long-term dependence to predict the pattern of flooding by the river Nile by introducing a test statistic called the *rescaled range statistic* or the *R/S statistic*. It was later modified by Mandelbrot (1972, 1975).

The R/S Statistic

The R/S statistic is defined as:

$$Q_T = \frac{1}{S_T}\left[\max_{1 \le k \le T} \sum_{j=1}^{k} (y_j - \bar{y} - \min_{1 \le k \le T} \sum_{j=1}^{k} (y_j - \bar{y})\right] \tag{4.11}$$

The term within the third bracket shows the range of partial sum of deviations, and it is rescaled by dividing with S_T, the standard deviation; and $\bar{y}$ is the mean. The R/S statistic is quite robust in itself. It has the ability to detect long-range dependence in nonGaussian time series with large skewness and kurtosis. Traditional methods like serial autocorrelation can be used to locate long range dependence only for models with near Gaussian effects. However, this method has severe shortcomings when used for nonGaussian time series. According to Mandelbrot and Wallis (1969), "when an ACF analysis program is used blindly for such t.s., the degree of dependence is grossly underrated". Further, R/S analysis has its advantage over other measures of long run dependency like the variance time analysis. For stochastic processes with infinite variances, the variance time analysis becomes inapplicable. This problem can be overcome by using the R/S analysis. The spectral analysis also becomes inappropriate for the typical economic time series due to its inability to capture the nonperiodicity. Again, R/S analysis can be used for such time series.

R/S statistic, however, has been criticized on many grounds. The classical R/S test often tends to indicate that a time series has long memory when it does not really have so. Annis and Lloyd (1976) criticized the test for its small sample bias. Andy Lo (1991, 1280) criticized the R/S statistic for being unable to distinguish between short- and long-term dependence and called it a "severe shortcoming in the applications of the R/S analysis". This observation was backed by Lo and McKinlay (1988, 1990), that found presence of significant short-range dependence in stock returns. Accordingly, Lo modified the R/S statistics by replacing the denominator, which now is the square root of a consistent estimator of the partial sum's variance.

The Modified R/S Statistic

The modified R/S Statistic is defined as:

$$\bar{Q}_T = \frac{1}{\bar{\sigma}^2_{(T)}}\left[max_{1\le k\le T}\sum_{j=1}^{k}(y_i-\bar{y})\right) - min_{1\le k\le T}\sum_{j=1}^{k}(y_i-\bar{y})] \tag{4.12}$$

where $\sigma^2_T(q) \equiv \frac{1}{T}\sum_{j=1}^{k}\left(y_j-\bar{y}\right)^2+\frac{2}{T}\sum_{j=1}^{q}\omega_j(q)\{\sum_{i=j+1}^{t}(y_i-\bar{y})\left(y_{t-j}-\bar{y}\right)\}$

$\bar{\sigma}^2_y$ and $\widehat{\gamma}_j$ are the usual sample variance and auto-covariance estimators of y. $\widehat{Q}_T$ is different from Q_T only in the denominator, which is the square root of a consistent estimator of the variance of the partial sum. The estimator $\widehat{\sigma}_T(q)$ involves not only sums of squared deviations of y_j, but also its weighted auto-covariances up to lag q. The weights $\omega_j(q)$ are those suggested by Newey and West (1994) and lead to a positive value of $\widehat{\sigma}_T(q)$, an estimator of 2π times the (unnormalized) spectral density function of y_t at frequency zero using a Bartlett window.

Lo's modified R/S test has been criticized for being too stringent. It has been shown numerically that even for a synthetic long memory time series, the Lo test cannot reject the null hypothesis of short-range dependence.

The rate of growth of the R/S statistic, as Hurst found it, was close to $n^{0.74}$, where n is the number of observations. This follows the rule proposed by Mandelbrot (1975) that a process with long memory converges toward a random variable at a rate n^H where H is the Hurst coefficient. For a process with no long memory, H = 0.5 and for a long memory process, the value of H lies between 0.5 and 1. This phenomenon was named *Hurst Phenomenon*. Moran in 1964 tried to explain this phenomenon by incorporating the assumption of an infinite variance. But as pointed out by Mandelbrot (1965) and Mandelbrot and Wallis (1969), assumption of infinite variance is not able to explain the Hurst phenomenon. For that, they suggested a model with a slowly decaying rate of autocorrelation. This brings us to the concept of Fractional Brownian Motion.

Fractional Brownian Motion

The fractional Brownian motion (fBm) is a Gaussian stochastic process $V_H(t)$ in continuous time with the property that for any t_1, t_2,$Var[V_H(t_2) - V_H(t_1)] \infty |t_2 - t_1|^{2H}$ with $0 < H < 1$

When $H = ½$, the process is the standard Brownian motion.

The fBm process is able to explain the Hurst phenomenon. The parameter H in the fBm is known as the *Hurst parameter* or *Hurst exponent*. Value of the Hurst exponent can be interpreted in different ways. If H lies in between 0 and 0.5, the series will be anti-persistent or mean reverting, characterizing a short memory process. An H value of 0.50 signifies a random and statistically independent or uncorrelated series: a random walk where the present is not influencing the future. For H lying between 0.5 and 1, i.e. for the fBm, the series would be a persistent or trend-reinforcing series, where the current trend will be followed in the next period. This process is mean averting and is often referred to as a biased random walk, in terms of nonlinear dynamics. Moreover, higher value of H would mean stronger persistence and lesser white noise in the series (Peters 1989).

However, "many natural stationary processes cannot be integrated even once, while preserving stationarity. It turns out, however, that sometimes one can integrate a process a fractional number of times, while preserving stationarity. This leads to a class of models known as fractionally integrated processes" (Samorodnitsky 2006, 204).

Testing for Fractional Integration: GPH Test

The GPH test, proposed by Geweke-Porter and Hudak (1983) is a semi-non-parametric approach to test for long memory and fractional integration. The spectral density of the fractionally integrated process y_t is given by:

$$f(\omega) = \left[4\sin^2\left(\frac{\omega}{2}\right)\right]^{-d} f_u(\omega) \tag{4.13}$$

ω is the Fourier frequency and $f_u(\omega)$ is the spectral density corresponding to u_t. The fractional difference parameter d can be estimated from the regression:

$$lnf\left(\omega_j\right) = \beta - dln\left[4\sin^2\left(\frac{\omega}{2}\right)\right] + e_j; \quad j = 1, 2, \ldots, \boldsymbol{n_f} \tag{4.14}$$

The least square estimate from Eq. 2.28 is normally distributed in large samples if$n_f(T) = T^{\alpha}$ with $0 < \alpha < 1$:

$$\widehat{d} \sim N\left(d, \frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \widehat{U})^2}\right) \tag{4.15}$$

where $U_j = ln\left[4\sin^2\left(\frac{\omega}{2}\right)\right]$ and $\bar{U}$ is the sample mean of U_j. Under the null hypothesis of no long memory, i.e. d = 0, the t-statistic $t_{d=0} = \widehat{d}\left(\frac{\pi^2}{6\sum_{j=1}^{n_f}\left(U_j-\bar{U}\right)^2}\right)^{-1/2}$ has a limiting standard normal distribution.

Testing for Fractional Integration: Whittle's Estimate of d

A long memory process y_t can be modeled parametrically into a fractionally integrated process (Granger and Joyeux 1980). Fractional integration in y_t can be described as:

$$(1 - L)^d(y_t - \mu) = u_t \tag{4.16}$$

Where L: lag operator, d: fractional integration or fractional difference parameter, μ: expectation of y_t, and u_t: stationary short memory disturbance with zero mean. For highly persistent economic and financial time series, integer difference of a time series is not sufficient and often d is allowed to be fractional. The fractional difference filter is defined as:

$$(1-L)^d = \sum_{k=0}^{\infty}\left(\binom{d}{k}(-1)^k L^K\right) \tag{4.17}$$

for any real $d > -1$ with binomial coefficients: $\binom{d}{k} = \frac{d!}{k!(d-k)!}$

The fractional difference filter can be equivalently treated as an infinite order autoregressive filter. With $|d| > 1/2$, y_t is non-stationary; when $0 < d < 1/2$, y_t is stationary and has long memory; when $-1/2 < d < 0$, y_t is stationary and has short memory, and is sometimes referred to as anti-persistent.

Long Memory in High Frequency Financial Time Series: Empirical Evidence

There have been a number of studies analyzing the cause, effect, and types of long memory in financial markets across the globe. This study discusses some of them.

Cheung (1993) applied the GPH test as well as ARFIMA to provide strong evidence for long memory in five nominal spot dollar exchange rates of Pound, Deutschemark, Swiss Franc, French Franc, and Japanese Yen. In the following year, Tschernig (1995) examined long memory in the US Dollar spot rates for Swiss Franc and Deutschemark using the ARFIMA model and came to a conclusion that weak long memory is present in US Dollar and it is specific to US Dollar only, as Swiss Franc/Deutschemark show no such property.

The early works of Greene and Fielitz (1977) used the rescaled range statistics (R/S statistics) to detect long-term memory in 200 daily stock returns. Most of the series were characterized by long-term dependencies. A significant work by Ding and Granger (1996) analyzed the SnP500 daily price from 1928 to 1991 and found evidence of strong autocorrelation in long lags. Barkoulas, Baum, and Travlos (2000) applied the spectral analysis method in the Greek stock market. They considered weekly stock returns of thirty stocks traded on the Athens stock exchange during 1988–1990. The result supported the presence of fractional integration and long memory in the Greek stock market. Barkoulas, Baum, and Christopher in their earlier study (2000) used the GPH test to prove the presence of long-term dependencies in the US stock return. Their study showed fractional integration in some firm's return but not in the market indices.

For Korean market, Lee et al. (2000) used a Fractionally Integrated GARCH model (FIGARCH) and showed the daily returns of KOSPI200 were characterized by long memory. Kang and Yoon (2007) used a ARFIMA-GARCH (Autoregressive fractionally integrated GARCH) model in Korean stock market for KOSPI200 and KOSDAQ returns and the result suggests a strong long-term dependency in the volatility process. ARFIMA-FIGARCH model was also used by Kasman and Torun (2007) to show the presence of long-term memory in the volatility of Turkish Stock Exchange (ISE).

Cavalcante and Assaf (2004) used several techniques to test for long memory in the Brazilian stock market. They used the R/S statistics, the rescaled variance statistic, a semi-parametric estimator as well as a FIGARCH model to examine long memory both in return and volatility. While the return series showed little or no long memory, the volatility did exhibit a significant long-term dependency.

While most of the studies are in favor of long memory, some studies conclude in contrary. Lux (1996, 701) used the Hurst coefficient and Rescaled range analysis to test for long memory in German stock return and according to him "virtually no evidence of such behavior is found". The most famous of these studies is probably Andy Lo's 1991 work, in which he developed a modified version of the famed R/S analysis and applied it to the US Stock market. Analysis of daily and monthly stock returns, showed no trace of long-term memory. Cheung and Lai (1995) took 18 countries stock indices and used the modified R/S test and fractional differencing to show that long memory does not exist in those markets. Jacobsen (1996) used the rescaled range statistics to five selected European indices along with one US and Japanese indices and concluded that long-term memory does not exist in those markets. Tolvi (2003) conducted his study in context of 16 OECD countries and concluded that long memory does not exist in those markets.

In context of India, Golaka Nath (2001) used the R/S analysis and variance ratio test to detect any long-term memory in Indian Rupee–US Dollar exchange rates. But his results were inconclusive. While the variance ratio test supported presence of only a short-term memory, the R/S analysis provided evidence for a long-term memory but with noise. Sen et al. (2009) provided evidence for long memory in some selected global stock markets including Indian market. Mukherjee et al. (2011) showed that Indian stock market is also characterized by long memory.

Sen (2012) considered nine currencies across the globe against Indian Rupee and using the techniques mentioned in the study tested for the presence of long memory or otherwise over a period of 13 years from 1998 to 2010.

Presence of long memory indicates possible nonlinearity in the underlying time series. One major characteristic of nonlinear time series is mutual dependence, which makes it inherently unstable, and very often chaotic in nature. This is not unnatural, as it has been an established fact that financial markets are complex in nature and determinism and chaos are two common characteristics of this kind of markets. The nonlinear dynamics, although indicated by the conditional volatility and long memory models, cannot be captured by these models, for that exist an entire field of the literature. In the next section, this study briefly discusses some major tools used in nonlinear market research.

Non-linearity, Determinism and Chaos in Financial Market: Theory and Empirics

Implication of Non-linearity for High Frequency Financial Time Series Data

The three most important revolutions in the field of science in the twentieth century have been the theory of relativity, the theory of quantum mechanics and the chaos theory. The first two theories rely significantly on calculus, which again stands on the premise of simple linear approximations of nonlinear behavior. But chaos theory is different from the first two in the sense that it deals with nonlinear behavior of a system. A chaotic system is essentially nonlinear; the time-dependent variables share a nonlinear relation. Assuming a system to be linear often resorts to some oversimplifications which are invalid in a real world. At the beginning, chaos started as a mere mathematical curiosity and any random or irregular behavior in the related system were treated as anomalies. But the next three decades saw an ever-increasing interest in this topic, which eventually revolutionized the way things were being perceived. Chaos, to be precise, surrounds us. Most of the natural phenomena are chaotic in nature. This concept has revolutionized the way scientists treat various systems as it gave them freedom to look at the more complex nonlinear dynamics.

For a system to be chaotic, it must be *nonlinear*, *deterministic,* and *sensitive to initial condition*. Future states of a chaotic system depend on some underlying rules. A chaotic system is sensitive to initial conditions, whereas iterations increase, the error in the system increases proportionally. A minuscule change in the initial condition can generate an error that increases exponentially with every iteration and eventually grows beyond 100 % after only a few iterations. A popular theory based on the premise of sensitive dependence on initial condition (SDIC) is the *Butterfly Effect*. It says that the flutter of the wings of a butterfly in Mexico can

cause or stop a tornado in New York. So at the end of the day it is the initial condition that matters the most.

The possibility that financial markets are chaotic has significant policy and theoretical implications. As pointed out by Peters (1991, 62), "…for too long, we seem to have been divided between "technicians" who believe that the market follows a regular cycle and 'quants' who believe there is no cycle at all. The truth lies somewhere in between". The markets are characterized by nonperiodic cycles and trends where volatility would be endogenous to the system. A small change in an index can have a major impact on the future and as the time goes by, economic forecasts become less reliable. Following the analysis of Peters (1991), we can say that even if it is possible to determine the equations of motion underlying the global stock market return, it would be difficult, if not impossible, to forecast beyond a short time frame, because current conditions can be hardly measured with significant accuracy. For the financial markets, instability will be intrinsic rather than aberration. Moreover, a chaotic financial market invalidates the assumption of efficient market. This defends some of those investment strategies that economic theories have earlier ruled out under efficient market hypothesis. Market timing, value investing and tactical asset allocation could be profitable investment strategies and investors might very well capitalize on market cycles (Peters 1991). A chaotic financial market thus suggests reframing of traditional asset pricing models to analyze stock price movements. The Gaussian assumption about probability distribution on which traditional methods of market analysis rest seems invalid for a chaotic series. Fractal distributions are characterized by infinite variance thus necessitating a reformulation of such models. The dynamic nature of equilibrium in a stock market further necessitates modifications in the class of models based on static mean reversion. The market analysts while explaining market movements, should act cautiously as the traditional econometric techniques often cannot capture the irregular cycles of a chaotic market. However, even with a dynamic quantitative technique, it would be unwise to predict future movements because long-term economic forecasting is no longer feasible.

Tests for Non-linearity, Determinism, and Chaos in Financial Market

Chaotic process, as is already mentioned, is a nonlinear, deterministic process that retains its dimensionality while placed in a higher embedding dimension and is sensitive to initial conditions. In a chaotic series, strange attractors exist characterized by fractal shape. The inquiry for possible presence of chaos could starts from the reconstruction of phase space. The methodology proposed by Kodba, Perc, and Marhl (2004) has been quite handy to test for the presence of chaos. The methodology has been followed by Chakrabarti (2010), Sen (2012) and by Chakrabarti and Sen (2012). The methodology is reproduced in the following sections.

The State-Space Reconstruction

A time series data is a series of scalars rather than a state-space object, thereby raising the necessity for state-space reconstruction of the data into state vectors (Kantz and Schreiber 1994). By Taken's (1981) embedding theorem, the reconstructed attractor of the original system will be given by:

$$\mathrm{p(i)} = \left(\mathrm{x_i},\ \mathrm{x_{i+\tau}},\ \mathrm{x_{i+2\tau}},\ \ldots,\ \mathrm{x_{i+(m-1)\tau}}\right) \tag{4.18}$$

Where m is the relevant embedding dimension. τ is the time delay parameter, or the time difference between adjacent components of p(i). In a state-space reconstruction state variables are replaced by a corresponding delayed variable. The reconstructed vector of the delayed variables will be identical to the original state variables in term of their intrinsic nature. With state-space reconstruction we can move over to a noise-free but larger dimension. Further, it is important to choose τ optimally. τ should neither be too small so as to make the set of information containing in x_i and $x_{\tau\ +\ i}$ essentially the same; nor it should be too large to lose all its memory of the initial states.

Mutual Information Criterion

One workable criterion to choose optimum τ could be the mutual information criteria (MIC) as introduced by Fraser and Swinney (1986). MI is nothing but the information available about the state $x_{i\ +\ \tau}$ given x_i. While calculating the MIC, the observations are first arranged in ascending order and then are divided in h equal intervals. The MIC is given as:

$$I(\tau) = -\sum_{h=1}^{j}\sum_{k=1}^{j} P_{h,k}(\tau) ln \frac{P_{h,k}(\tau)}{P_h P_k} \tag{4.19}$$

P_h: probability that the variable falls into the interval, P_k: probability that the variable falls into the interval k, $P_{h,k}()$: joint probability of one variable falls into the interval h and another falls into the interval $h\ +\ \tau$, τ being the time delay. We could find a reliable optimum τ provided we know the appropriate probability density function. The optimum time delay is that τ for which $I\ (\tau)$ is minimum for the first time. The criterion is such, because, if the system is chaotic system, as the correlation between x_h and x_k becomes negligible, $I\ (\tau)$ will tend to zero when τ tends to infinity. At the first minima, $x_{i+\tau}$ adds the most to the available information from x_i without losing the correlation between them completely.

False Nearest Neighborhood

This study reproduces the false nearest neighborhood (FNN) technique of Kennel et al. (1992) with or an optimum embedding dimension m, the reconstructed delay

space is topologically consistent with the original state space. Therefore, two points in the neighborhood, when subjected to a short forward iteration, continue to be in the neighborhood. If, a point i has a neighbor j that after a forward iteration does not remain in the neighborhood of i anymore, j is said to be a FNN of i. When the reconstructed delay space is formed with too small embedding dimension, the reconstructed space ceases to be topologically consistent with the original state space and points after forward iteration may appear to be in the neighborhood while actually they are far apart.

Let $\boldsymbol{p}(i)$ be a point on a m dimensional reconstructed space with a nearest neighbor $\boldsymbol{p}(j)$. If d be the Euclidean distance between two points, defined as:

$$d(m) = \|x_i(m) - x_j(m)\| \tag{4.20}$$

To get the nearest neighbors, the distance d will be minimized. Then, there would be iteration for a bigger dimension and it should be checked whether d is still minimized. The embedding dimension is increased by one so that

$$d(m+1) = \|x_i(m+1) - x_j(m+1)\| \tag{4.21}$$

If $\boldsymbol{p}(i)$ is a false nearest neighbor of $\boldsymbol{p}(j)$ then $d(m + 1)$ will not be minimized. This is characterized by the change of distance between the two points with the embedding dimension increased from d to $d + 1$ being larger than an acceptable tolerance level. This can be expressed as:

$$\frac{\left|x_{i+m\tau} - x_{j+m\tau}\right|}{d(m)} > R_t \tag{4.22}$$

Where $\boldsymbol{R_t}$ is distance ratio threshold. The value of the $\boldsymbol{R_t}$ will be chosen in such a manner for which the percentage of false nearest neighbors in the dataset falls to zero. The choice of R although subjective should be judicious (Rhodes and Morari 1997). If $\boldsymbol{R_t}$ is very small FNN will not drop to zero at the correct embedding dimension. On the other hand, if $\boldsymbol{R_t}$ is too large it would tend to accept a unusually lower embedding dimension.

Determinism Test

Kaplan and Glass (1992) proposed an effective technique to check whether the series is truly deterministic. In this technique, initially, the attractor is plotted in a $x(t)$ versus $x(t - \tau)$ space. The phase space is then subdivided into $q \times q$ grids, each of which has a part of the attractor passing through it. A directional vector of unit length, called a 'trajectory vector', is assigned to each grid that corresponds to the portion of attractor in it. If e_i be the unit vector passing through each box, then the resultant vector V_k from all the vector passes is a simple average given by:

$$V_k = \frac{1}{P_k}\sum_{i=1}^{P_k} e_i \tag{4.23}$$

Where P_k is the number of passes through the kth grid. Now for a deterministic system, the phase space offers a unique solution. For that, the unit vectors in a grid must follow the same direction, that is, the trajectories inside the grid must never cross, whereas for a stochastic system the trajectories inside the grid cross each other. As the crossing never occurs in deterministic system, V_k is of unit length and for stochastic system value of V_k is significantly lower than one.

Maximum Lyapunov Exponent

Lyapunov exponent (Λ) measures the degree of separation between infinitesimally close trajectories in phase space. For a chaotic system, the trajectories diverge continuously as the system is dependent on initial conditions. An m-dimensional system has m numbers of Lyapunov exponents, of which the maximum Lyapunov exponent (Λ_{max}) gives an indication of chaos. A system is chaotic and close trajectories diverge in phase space, if the maximum Lyapunov exponent is positive.

To calculate Λ_{max} (Wolf, Swinney and Vastano 1985) an initial point $\boldsymbol{p}(0)$ is chosen with a nearest neighbor. The distance between them be L_0. Now the points are iterated toward by time t_{evolv} (which is essentially equal to τ) and the distance after iteration is noted (L_{evolv}). For a chaotic system, the trajectories diverge in time. Hence $L_{evolv} > L_0$. The value of t_{evolv} has to be less than $m\tau$, a larger value will result in an underestimation of Λ_{max}. At the end of first iteration, a replacement is made to choose a new nearest neighbor for the evolved $\boldsymbol{p}(0)$. This process continues till $\boldsymbol{p}(0)$ reaches the end of the series. Maximum Lyapunov exponent can be presented as:

$$\Lambda_{max} = \frac{1}{Mt_{evolv}}\sum_{i=0}^{M} \ln \frac{L_{evol}^{(i)}}{L_0^{(i)}} \tag{4.24}$$

Non-linearity, Determinism, and Chaos in Financial Market: Few Empirical Evidences

Nonlinear time series analysis is a relatively new area of research and due to lack of computational power even a couple of decades ago, empirical works in this area do not date too far back.

Of the early studies dealing with nonlinearities in financial time series, Scheinkman and Lebarons' (1989) work needs special mention. The study considers over 5,200 daily returns and further constructed a less noisy weekly return series from it. It found evidence of strong nonlinearity in the series. Along this line, Hsieh's (1989) study took five exchange rates, British Pound, Canadian Dollar,

Japanese Yen, Deutsche Mark and Swiss Franc under consideration and used the BDS test to find evidence of nonlinearity in British Pound and Deutsche Mark. He went farther with his study of nonlinearity in stock returns (1991) where he considered weekly stock returns for the period 1963–1987 from 11 stock returns and concluded that although the results from the BDS test were consistent with low complexity chaos, the apparently random fluctuations in the series under consideration were not in reality caused by chaos. Bajo-Rubio et al. (1992) used the correlation dimension and the Lyapunov exponents and chaos in Spanish Peseta-US Dollar spot and forward rates. De Grauwe, Dewachter, and Embrechts (1993) used daily exchange rates for Deutsche Mark-US Dollar, Yen–US Dollar and British Pound–US Dollar for the period 1971–1990. The study subsequently divided the entire period into four subperiods and calculated the correlation dimensions. The result supported chaotic behavior in the Yen-Dollar and Pound-Dollar series.

De Grauwe et al. (1993); Lux (1996); Brock and Hommes (1998); De Grauwe and Grimaldi (2006) all provided strong evidence of chaotic dynamics in market behavior. Kočenda (1996) pointed toward nonlinear dependency between Czech Koruna and six other major currencies. Cuaresma (1998) conducted his study in context of Austrial Schilling/US Dollar and the result supported presence of strong nonlinearity in the market.

Bask (2002) tested for the presence of chaos in nominal exchange rate. This study considered daily Swedish Kroner rates against Deutsche Mark, ECU, US Dollar and Japanese Yen. The study used a block wise bootstrap method to calculate maximum Lyapunov exponent for the series. The exchange rate series were found to be deterministically chaotic in nature.

Chris Brooks (1995) tested for chaos in 10 currencies against the British pound for a period of over 20 years. He calculated both the correlation dimension as well as the Lyapunov exponent. While he found a somewhat nonlinear dependence in the series, he found no evidence of chaos. Guillaume (1995, 2000) tested for chaos in intraday exchange rates for four currency pairs using the GP algorithm. The study found no evidence of chaos in the series.

Rodriguez et al. (2004) used a new test based on the stability of the Lyapunov exponent and found strong evidence of deterministic chaos in French Franc/US Dollar, Canadian Dollar/US Dollar while the German Mark/US Dollar series showed no chaotic signature. Vandrovych (2007) used the methodology of correlation dimension and maximum Lyapunov exponent and found no evidence of chaos in four exchange rates against US Dollar. However, the series remains highly nonlinear.

In context of India, Sen, Chakrabarti and Sarkar (2011) showed that Indian foreign exchange market is characterized by chaotic dynamics—highlighting empirical evidence of chaos in such foreign exchange rates. Sen (2012) explored the chaotic behavior of nine currencies across the globe against Indian Rupee over a period of 13 years from 1998 to 2010. Chakrabarti and Sen (2012) while

analyzing the nature of global stock market crashes identified the inherent chaotic nature of global stock market as a possible factor behind its knife-edge stability.

Volatility, Long Memory, and Chaos: An Example from Indian Financial Market

The chapter now seeks to present a study where the presence of the three stylized facts described in this chapter, namely presence of conditional volatility, long memory, and nonlinearity (specifically, presence of chaos) could be explored in the context of Indian financial market using the methodologies described in the chapter. This might be helpful for those who seek to analyze the stylized facts for financial time series in a comprehensive fashion. This is particularly so, because the existing studies consider and analyzes these stylized facts in isolation. An analogous study in the context of global stock market has been performed by Chakrabarti (2010). Sen (2012) performed a similar analysis in the context of Indian foreign exchange market.

The study concentrates on Indian financial markets and considers high frequency time series data from both the stock market and foreign exchange market. The study selects six foreign exchange rates against namely, US Dollar, Brazilian Real, Chinese Yuan, Japanese Yen, Malaysian Ringgit and Mexican Peso against Indian Rupee. So far as the stock market is concerned, the study picks the bench market stock market index in India namely the BSE SENSEX. The markets and currencies are selected from across the globe in such a fashion that it consists of global as well as regional representatives, some of them being on the same competitive platform, and hence providing a better view of the comparative analysis.

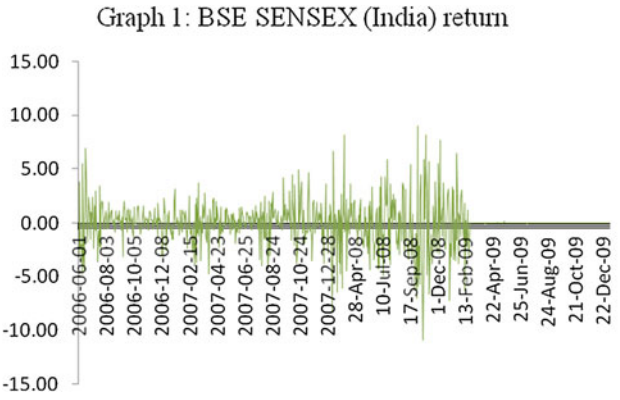
The sample period is carefully chosen for the study. With the liberalization in 1992, India saw a sea change in trade and stock market related policies. With lesser restrictions, foreign investment increased, capital markets were also reformed. In 1993, the unified exchange rate initiated the era of market determined exchange rate. In 1994, full current account convertibility was achieved and in 1997 capital account convertibility was formalized. Although India has resisted full capital account convertibility, by 1998 the exchange rate was truly market driven. This justifies the sample period. The daily data are collected from January 1998 to December 2011 for all currencies and stock prices from the websites of the central banks and stock exchanges of the corresponding countries. This study, however, does not take into consideration the possible presence of structural break during this long period of study.

Stylized Facts in Indian Financial Market

The study starts from a simple graphical analysis where financial market returns have been plotted against time. As is evident from Graphs 4.1–4.7, the Indian financial market (both the stock market as well as the foreign exchange markets) has been characterized by high volatility that has been time varying in nature. Further, the periods of low (high) volatility have been followed by a period of low (high) volatility providing evidence of volatility clustering. Of these seven graphs, only Graph 4.1 is presented here. The rest have been shown in appendix.

Apart from the time-varying nature of volatility and presence of volatility clustering, Indian financial market returns are characterized by some stylized facts of financial time series (Table 4.1).

The returns are skew, leptokurtic, and nonnormal. With these features of Indian financial market returns, volatility could best be analyzed by GARCH family models. This study emphasizes particularly on exploring the presence of volatility clustering, volatility mean reversion, and asymmetric response of volatility toward good and bad news.



Graph 4.1 BSE SENSEX (India) return

Table 4.1 Descriptive statistics of Indian financial market return

Financial market		Mean	Standard deviation	Skewness	Kurtosis	Jarque–Bera	Probability
Stock market	BSE SENSEX	−0.0033	1.92	−0.12	7.79	792.3	0
Foreign exchange market	RS/USD	0.01	0.001	−0.165	22.813	73659.38	0
	RS/Peso	0.01	0.002	−1.072	26.229	102083.4	0
	RS/Real	0.004	0.005	−0.728	18.565	70878.57	0
	Rs/Yuan	0.01	0.001	−0.082	22.438	8289.94	0
	RS/Yen	0.02	0.003	0.438	9.59	405560.2	0
	Rs/Ringgit	0.01	0.003	−0.59	49.483	1411.916	0

Conditional Volatility in Indian Financial Market

The Indian market is characterized by asymmetric response of volatility toward good and bad news. The EGARCH model turns out to be the best fit model for Indian financial market on the basis of AIC criterion. Consequently, volatility in Indian market responds asymmetrically to good and bad news in the financial market. This is evident from the statistical significance γ for all types of financial markets. Further, significantly negative γ values suggest presence of 'leverage effect' where volatility increases with unfavorable news. The statistically significant and positive β values suggest presence of volatility clustering in the market where present conditional variance is positively and significantly related to the immediate past variance (Table 4.2). The ARCH-LM test suggests absence of any further ARCH effect in residual.

Results from estimation of component ARCH model suggest presence of volatility persistence in the market. Long-term volatility approaches its steady state at a slow rate. This is suggested by significantly positive values for γ that are close to one for all the financial markets (Table 4.3).

Table 4.2 Results for estimating EGARCH model for BSE SENSEX return

Market		Mean equation Constant	Variance equation				ARCH-LM statistic (prob) F, Chi sq
			ω	α	γ	β	
Stock market	BSE SENSEX	0.12	0.04	−0.03	−0.25*	0.95*	0.79, 0.79
Foreign exchange market	RS/USD	0.078*	−0.01	−0.04	−0.12*	0.97*	0.24, 0.24
	RS/Peso	0.00114	−0.76*	0.09	−0.32*	0.92*	0.74, 0.74
	RS/Real	0.139	0.19*	−0.15	−0.31*	0.88*	0.88, 0.88
	Rs/Yuan	0.0006	−0.35*	−0.13	−0.15*	0.95*	0.71, 0.71
	RS/Yen	0.0003	−2.81*	−0.09	−0.54*	0.69*	0.91, 0.91
	Rs/Ringgit	0.0002	−0.30	0.02	−0.15*	0.97*	0.71, 0.71

* implies significance at 1 % level

Table 4.3 Results for applying Component ARCH model for BSE SENSEX

Market		Mean equation Constant	Variance equation				
			ω	ρ	φ	α	β
Stock market	India	0.15	3.28	0.98*	−0.04*	0.06*	−0.82*
Foreign exchange market	RS/USD	−0.20	6.63	0.97*	0.25	−0.19	1.09**
	RS/Peso	−0.001	0.0005	0.99*	0.11*	0.05	−0.97*
	RS/Real	−0.17	13.32	0.99*	0.10	0.14	−0.42
	Rs/Yuan	−0.001	0.0005*	0.96*	0.16*	−0.23	−0.28
	RS/Yen	−0.003	0.001*	0.90*	0.002	0.12*	−0.70*
	Rs/Ringgit	−0.15	18.50	1.00*	−0.02	0.15	0.83*

* implies significance at 1 % level

Long Memory in Indian Financial Market

As it can be seen from Table 4.4, Indian financial market returns are characterized by significant level of long-term memory. This is evident from the results of both R/S and modified R/S test. The test statistics are significant at 1 and 5 % levels, hence the null hypothesis of no long memory is rejected. Further, as suggested by the values of the GPH test statistic and Whittle's d, all the series are also fractionally integrated. Finally, values of all the Hurst coefficients from the R/S analysis are significantly above 0.5, indicating toward a trend reinforcing long memory process. This high value of Hurst coefficient is further indicative of possible presence of nonlinearity and sensitive dependence on initial condition in the market. This, in turn, provokes a researcher to test for possible presence of nonlinearity, specifically, chaos in Indian stock market. And, that is where we move next.

Non-linearity and Chaos in Indian Financial Market

For a time series to be deterministic and chaotic, it has to be nonlinear at first. Therefore, it is necessary to check for nonlinearity in the underlying series. For this, the BDS test, named after Brock, Dechert and Scheinkman, is used. The test developed by Brock, Dechert, and Scheinkman (1996) is a popular test for nonlinearity. It was initially used to test for the null hypothesis of independent and identical distribution (iid) for the purpose of detecting nonrandom chaotic dynamics. The rejection of the null of IID by the BDS statistic, although provides evidence of the series being nonlinear, it does not necessarily mean the time series exhibits chaotic behavior. Rejection of IID can be consistent with any of the following four types of nonIID behavior: linear dependence, nonstationarity, nonlinear stochastic processes, and nonlinear deterministic process (chaos). Since all financial time series data are nonstationary, only linear dependence should be removed by suitable filtering before proceeding further to discriminate between

Table 4.4 Presence of long memory in Indian stock market

Series	R/S test	Modified R/S test	Estimate of d			Hurst coefficient
			GPH d	GPH test stat.	Whittle's d	R/S Analysis
BSE SENSEX	2.24*	3.161*	0.478	2.1765*	0.228	0.739
RS/USD	11.654**	5.8158**	0.6585	7.5684**	0.215	0.657
RS/Yuan	11.9181**	5.9822**	0.6629	7.6189**	0.217	0.728
RS/Yen	8.4489**	5.2669**	0.54	6.2065**	0.172	0.634
RS/Peso	6.9591**	4.0923**	0.438	5.0323**	0.202	0.618
RS/Real	5.5997**	3.1748**	0.449	5.1648**	0.195	0.782
RS/Ringgit	12.3754**	6.7601**	0.586	6.7294**	0.279	0.664

* implies significance at 1 % level
* implies significance at 5 % level

Table 4.5 Detection of possible chaos in Indian financial market

	Rs–Dollar	Rs–Peso	Rs–Real	Rs–Ringgit	Rs–Yen	Rs–Yuan	BSE–SENSEX
τ (Emb. delay)	3	1	2	4	2	5	1
Shannon Entropy	2.597	3.493	2.808	2.562	3.643	2.594	3.406
m (Emb. Dimension)	5	5	5	5	5	5	5
Determinism	0.8074	0.8802	0.6845	0.8187	0.7411	0.7831	0.8862
Maximum Lyapunov	5.6709	7.4784	5.1092	5.9099	8.3267	2.89	5.0567

nonlinearities due to stochastic behavior and nonlinearities due to the existence of chaotic behavior. All financial data are characterized by significant autocorrelation, and hence need to be AR filtered. Otherwise the data will be less reliable while applied to a nonlinear model. Order of AR should be determined by minimum AIC criterion. This nonlinearity, however, can be generated from a stochastic (ARCH type) model or from chaotic behavior. To remove that, the AR filtered series should be filtered again by a suitable GARCH model. We have run BDS test for our data series and found the series to be nonlinear. The appropriate AR-GARCH filtered series then are passed through the tests proposed by Kodba, Perc, and Marhl (2004) and further by Perc (2005a, 2005b). The following table (Table 4.5) shows the results obtained when the tests are replicated for the Indian financial market data.

The Indian financial markets are thus deterministic and the positive Lyapunov exponents for all the series confirm the inherent chaotic nature of them. To be noted here, Rs-Yen series turns out to be most chaotic, while Rs-Yuan is the least chaotic series. This chaotic nature of the series provides a possible reason for the endogenous volatility and the ineffectiveness of interventions as by the nature of a chaotic series, any long run prediction is virtually impossible; only short run predictions can be made.

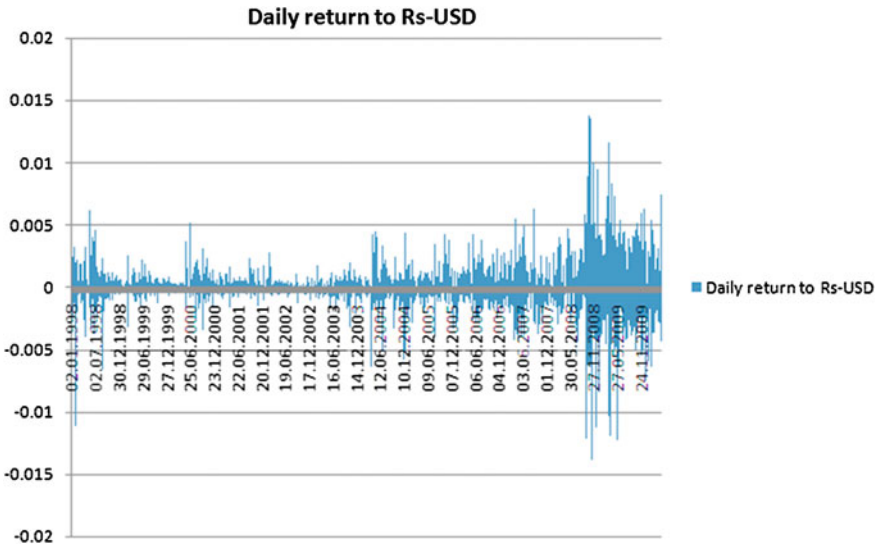
Conclusion

The chapter reviews some of the theoretical models to analyze few stylized facts in the context of high frequency financial time series data. The presence of such stylized facts in financial markets often necessitates reformulation of standard asset pricing models and redefining some of the rules on which theories of financial economics rest. Such an analysis therefore is important from the point of view of researchers as well as policy-makers. There are, however, some limitations of this review. First, it has considered only three most important stylized facts for a financial market data. Further, it has considered only a few models used in the literature. The researchers, while analyzing presence of conditional volatility in financial market may use alternative ARCH school of models like NAGARCH

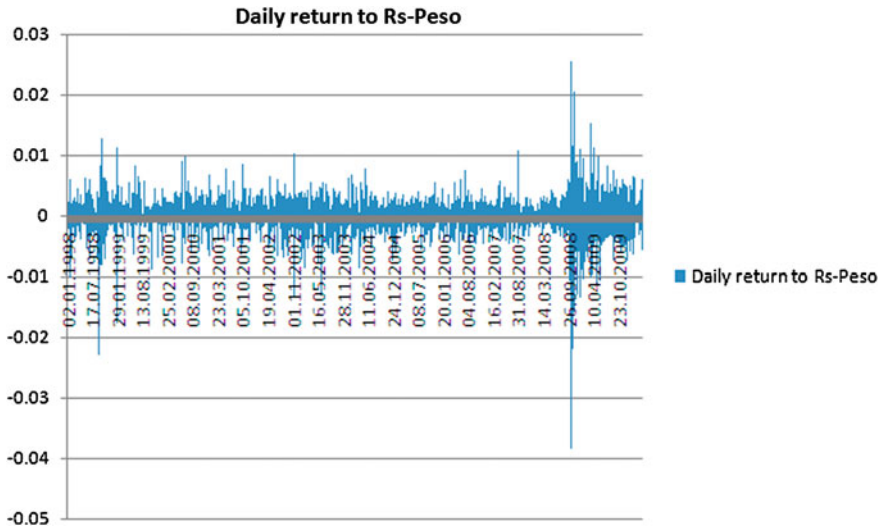
(Non linear Asymmetric GARCH), Power GARCH (PARCH), GARCH-M (GARCH in mean), APARCH (Asymmetric power ARCH), or AVGARCH. For analyzing presence of long memory researchers may apply the log-periodogram-based method, the Wavelet estimator or fractionally integrated GARCH model. The last model could particularly be used to test for the long memory property of volatility by combining the fractionally integrated process for the mean with the regular GARCH process for the conditional variance. The ARFIMA model could be used to explain the long memory property that exists in the conditional mean. Further, a chaotic system could be analyzed using the alternative method of correlation dimension. This study, however, attempts to provide a well-rounded review of analyzing a high frequency financial time series, so that all major stylized facts are highlighted and examined for. It might be beneficial for the researchers to analyze a given financial time series data using the steps and procedures discussed in the chapter.

Appendix

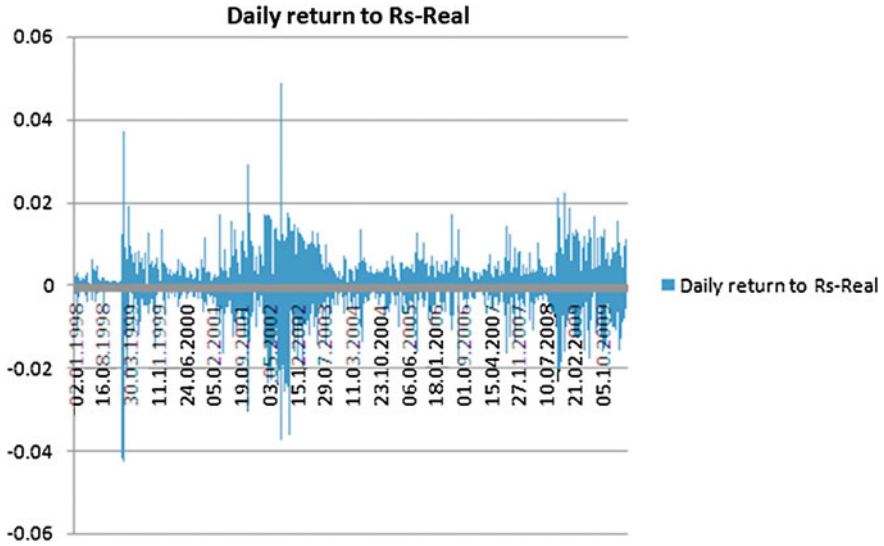
Graph 4.2, Graph 4.3, Graph 4.4, Graph 4.5, Graph 4.6, Graph 4.7



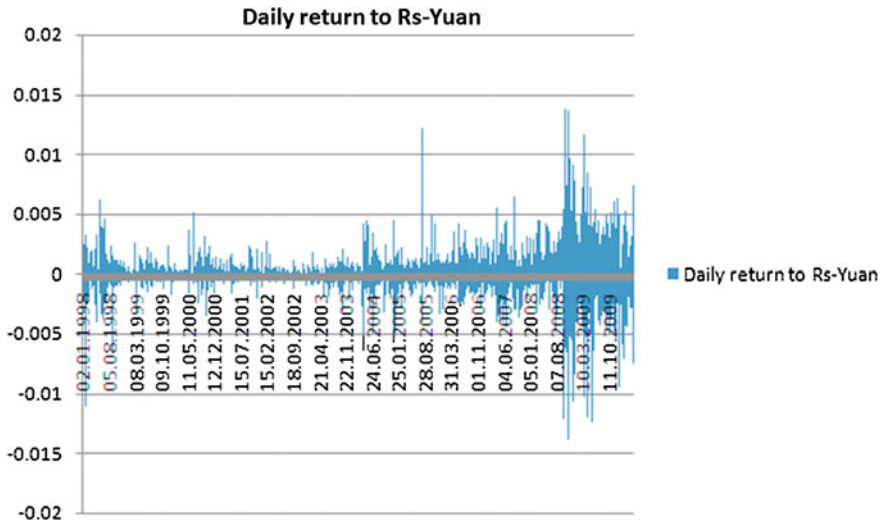
Graph 4.2 Daily return to Rs–US Dollar



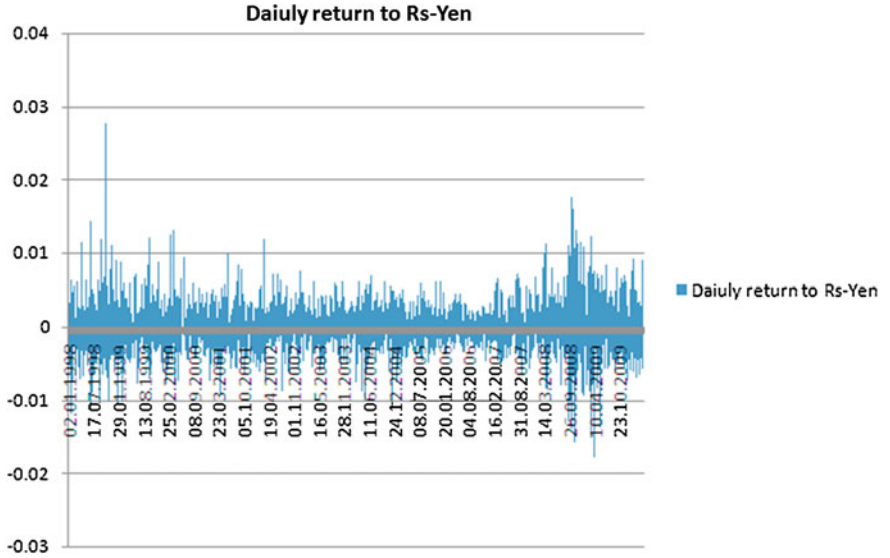
Graph 4.3 Daily return to Rs–Peso



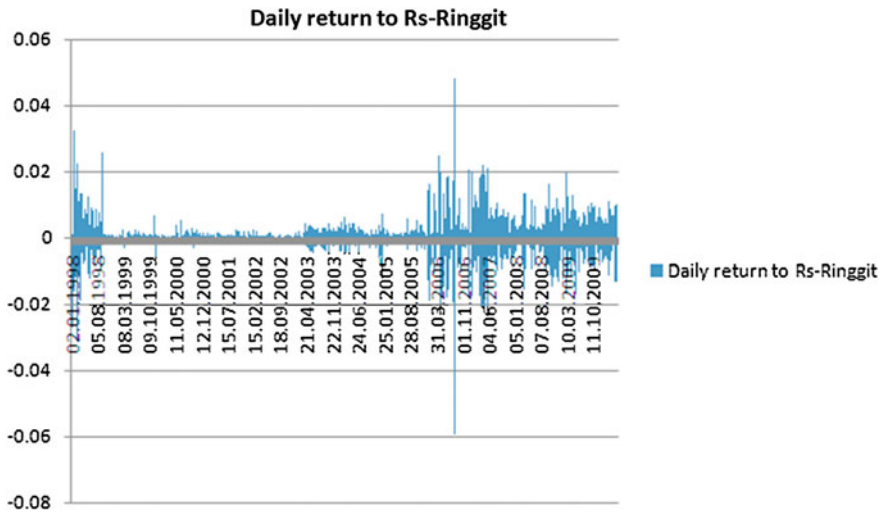
Graph 4.4 Daily return to Rs–Real



Graph 4.5 Daily return to Rs–Yuan



Graph 4.6 Daily return to Rs–Yen



Graph 4.7 Daily return to Rs–Ringgit

References

Alexander C (2001) Market Models: a guide to financial model analysis,Chichester, J. Wiley

Annis AA, Lloyd EH (1976) The expected value of the adjusted rescaled Hurst range of independent normal summands. Biometrika 63:111–116

Barkoulas J, Baum T, Christopher F, Travlos N (2000) Long memory in the Greek stock market. Appl Financ Econ 10(2):177–184

Bask M (2002) A positive Lyapunov exponent in Swedish exchange rates? Chaos, Solitons Fractals 14(8):1295–1304

Black, F (1976) Studies in stock price volatility changes. In: American Statistical Association, Proceedings of the Business and Economic Statistics Section, 177–181

Bollerslev T (1986) Generalized autoregressive conditional heteroscedasticity. J Econometrics 31:307–327

Brock W, Hommes C (1998) Heterogeneous beliefs and routes to chaos in a simple asset pricing model. J Econ Dyn Control 22:1235–1274

Brock WA, Dechert W, Scheinkman J (1996) A test for independence based on the correlation dimension. Econometric Rev 15(3):197–235

Brooks C (1995) A measure of persistence in daily pound exchange rates. Appl Econ Lett 2:428–431

Cavalcante J, Assaf A (2004) Long range dependence in the returns and volatility of the Brazilian stock market. Eur Rev Econ Finance 3:5–22

Chakrabarti G (2010) Dynamics of global stock market: crisis and beyond. (Vdm-Verlag Dr. Müller), Germany

Chakrabarti G, Sen C (2012) Anatomy of global stock market crashes: an empirical analysis. Springer, Newdelhi

Cheung Y (1993) Long memory in foreign-exchange rates. J Bus Econ Stat 11(1):93–101

Cheung Y, Lai K (1995) A search for long memory in international stock market returns. J Int Money Financ 14:597–615

Chou RY (1988) Volatility persistence and stock valuations: some empirical evidence using Garch. J Appl Econometrics 3(4):279–294
Cuaresma JC (1998) Deterministic chaos versus stochastic process: an empirical study on the Austrian Schilling-US Dollar Exchange Rate. *Economic Series* No. 60
De Grauwe P, Grimaldi M (2006) Exchange rate puzzles: a tale of switching attractors. Eur Econ Rev 50:1–33
De Grauwe P, Dewachter H, Embrechts M (1993) Exchange rate theory: chaotic models of foreign exchange markets. Blackwell, Oxford
Ding Z, Granger CWJ (1996) Modeling volatility persistence of speculative returns: a new approach. J Econometrics 73:185–215
Engle R (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica 50:987–1008
Engle RF, Lee GGJ (1999) A permanent and transitory component model of stock return volatility. In: Engle R, White H (eds.) Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger. Oxford University Press, New York
Engle RF, Victor K Ng (1993) Measuring and testing the impact of news on volatility. J Financ 48(5): 1749–1778
Engle RF, Bollerslev T (1986) Modeling the persistence of conditional variances. Econometric Rev 5:1–50
Engle RF (1990) Discussion: stock market volatility and the crash of '87. Rev Financ Stud 3:103–106
Engle RF, Lilien DM, Robbins RP (1987) Estimating time varying risk premia in the term structure: The ARCH-M model. Econometrica 55:391–407
Fraser AM, Swinney HL (1986) Independent coordinates for strange attractors from mutual information. Phys Rev A 33:1134–1140
Geweke J, Porter-Hudak S (1983) The estimation and application of long memory time series models. J Time Ser Anal 4:221–238
Glosten LR, Jagannathan P, Runkle DE (1993) On the relation between expected value and the volatility of excess returns on stocks. J Financ 48:1779–1801
Granger CWJ, Joyeux R (1980) An introduction to long-memory time series models and fractional differencing. J Time Ser Anal 1:15–30
Greene MT, Fielitz BD (1977) Long-term dependence in common stock returns. J Financ Econ 4(3):339–349
Guillaume DM (1995) A low-dimensional fractal attractor in the foreign exchange markets? In: Robert Trippi (ed) Chaos & nonlinear dynamics in the financial markets, IRWIN Professional Publishing, pp. 269–294
Guillaume DM (2000) Chaos in the foreign exchange market. In: Intra-daily exchange rate movements, Chapter 3. Springer, Boston, pp 59–79
Hsieh DA (1989) Testing for nonlinear dependence in daily foreign exchange rates. J Bus 62(3):339–368
Hurst H (1951) Long Term Storage Capacity of Reservoirs, Transactions of the American Society of Civil Engineers, 116: 770–799
Jacobsen B (1996) Long-term dependence in stock returns. J Empirical Financ 3:393–417
Kang SH, Yoon S (2007) Long memory properties in return and volatility: evidence from the Korean stock market. Physica A 385(2):591–600
Kantz H, Schreiber T (1997) Nonlinear time series analysis. Cambridge University Press, Cambridge
Kaplan DT, Glass L (1992) Direct test for determinism in a time series. Phys Rev Lett 68:427–430
Kasman A, Torun E (2007) Long memory in the Turkish stock market return and volatility. Central Bank Rev 20(2):13–27
Kennel MB, Brown R, Abarbanel HDI (1992) Determining embedding dimension for phase space reconstruction using a geometrical construction. Phys Rev A45:3403–3411
Kočenda E (1996) Volatility of a seemingly fixed exchange rate. East Eur Econ 34(6):37–67

Kodba S, Perc M, Marhl M (2004) Detecting chaos from a time series. Eur J Phys 26(2005):205–215
Lee J, Kim TS, Lee HK (2000) Long memory in volatility of Korean stock market returns. Proceedings of Informs & Korms
Lo AW (1991) Long-term memory in stock market prices. Econometrica 59:1279–1313
Lo AW, MacKinlay C (1988) Stock market prices do not follow random walks: evidence from a simple specification test. Rev Financ Stud 1:41–66
Lo AW, McKinlay AC (1990) Data snooping biases in tests of financial asset pricing models, Review of Financial Studies, Society for Financial Studies, 3(3): 431–67
Lux T (1996) Long-term stochastic dependence in financial prices: evidence from the german stock market. Appl Econ Lett 3(11):701–706
Mandelbrot BB (1965) Forecasts of Future Prices, Unbiased Markets, and "Martingale" Models, The Journal of Business, University of Chicago Press, 39: 242–255
Mandelbrot B, Wallis J (1969) Noah, joseph and operational hydrology. Water Resour Res 4:909–918
Mandelbrot BB (1971) When Can Price Be Arbitraged Efficiently? A Limit to the Validity of the Random Walk and Martingale Models, The Review of Economics and Statistics, MIT Press, 53(3): 225–236
Mandelbrot BB (1972) Statistical methodology for non-periodic cycles; From the covariance to R/S analysis. Ann Econ Soc Measur 1:259–260
Mandelbrot BB (1975) A fast fractional Gaussian noise generator. Water Resour Res 7:543–553
Mantegna R, Stanley E (2000) Introduction to econophysics. Cambridge University Press, Cambridge
Mukherjee I, Sen C, Sarkar A (2011) Long memory in stock returns: insights from the Indian market. Int J Appl Econ Financ 5(1):62–74
Nath G. (2001) Long memory and Indian stock market-an empirical evidence, UTIICM Conference Paper. http://www.utiicm.com/Cmc/PDFs/2001/golakanath_longmemory_indian_stock_market.pdf. Accessed 3rd Jan 2009
Nelson DB (1991) Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59:347–370
Newey WK, West KD (1994) Automatic Lag Selection in Covariance Matrix Estimation. Rev Econ Stud 61(4):631–653
Perc M (2005a) Nonlinear time series analysis of the human electrocardiogram. Eur J Phys 26:757–768
Perc M (2005b) The dynamics of human gait. Eur J Phys 26:525–534
Peters EE (1991) Chaos and order in the capital markets. John Wiley & Sons, New York
Peters EE (1989) Fractal structure in the capital markets. Financ Anal J 32–37 (July–August)
Rabemananjara R, Zakoian JM (1993) Threshold arch models and asymmetries in volatility. J Appl Econometrics 8(1):31–49
Rhode C, Morari M (1997) False-nearest-neighbors algorithm and noise-corrupted time series. Phys Rev E 55(5):6162–6170
Rickles D (2008) Econophysics and the complexity of financial markets. In Collier J, Hooker C (eds.) Handbook of the philosophy of science, Vol 10. Philosophy of complex systems. Elsevier, North Holland
Rodrìguez SA, Kharitonov V, Dion JM, Dugard L (2004) Robust stability of neutral systems: A Lyapunov-Krasovskii constructive approach. Int J Robust Nonlinear Control 14(16):1345–1358
Rubio BO, Rodriguez F, Rivero S (1992) Chaotic behavior in exchange rate series: first results for the Peseta-U.S. dollar case. Econ Lett 39(2):207–211
Samorodnitsky G (2006) Long range dependence. Found Trends Stochast Syst 1(3):163–257
Sarkar A, Chakrabarti G, Sen C (2009) Indian stock market volatility in recent years: transmission from global and regional contagion and traditional domestic sectors. J Asset Manage 10(4):63–71
Scheinkman J, LeBaron B (1989) Nonlinear dynamics and stock returns. J Bus 62(3):311–337

Schwert GW (1989) Business cycles, financial crises, and stock volatility. Carnegie-Rochester Conf Ser Public Policy 31:83–126
Sen C, Mukherjee I, Chakrabarti G, Sarkar A (2009) Long memory in some selected stock markets: an autocorrelation based approach. Indian J Capital Market 3(1):11–16
Sen C (2012) Indian foreign exchange market: a study into volatility and regime switch. LAP Lambert Academic Publishing, Germany
Sen C, Chakrabarti G, Sarkar A (2011) Evidence of chaos: a tale of two exchange rates. Empirical Econ Lett 10(8):777–784
Taylor SJ (1986) Modeling financial time series. John Wiley and Sons, Chichester
Tolvi J (2003) Long memory and outliers in stock market returns. Appl Financ Econ 13:495–502
Tsay RS (2001) Analysis of financial time series. Wiley-Interscience, New York
Tschernig R (1995) Long memory in foreign exchange rates revisited. J Int Financ Markets Inst Money 5(2/3):53–78
Vandrovych V (2007) Nonlinearities in exchange-rate dynamics: chaos? SSRN: http://ssrn.com/abstract=945797. Accessed 24 July 2008
Wolf A, Swift JB, Swinney HL, Vastano JA (1985) Determining Lyapunov exponents from a time series. Physica D 16:285–317
Zakoian JM (1994) Threshold heteroscedastic models. J Econ Dynam Control 18:931–955

Chapter 5
Globalization and Labour Markets

Ranajoy Bhattacharyya

Introduction

Income inequality and the level of globalisation are rising rapidly across most parts of the world. The obvious question to ask is: is this pure coincidence or are the two causally related? To answer this question one needs three things: A theoretical argument that links the two observations, a methodology for empirically implementing the theoretical argument and of course reliable data. This chapter is about the ways in which existing research has attempted to implement these steps in order to answer the above question.

Income inequality can be affected by a lot of factors, a small number of which are related to globalisation. Economic wisdom suggests that even this single cause has many dimensions: flow of goods, technology, jobs, factor migration—all of which can have a direct impact on the distribution of income. With so many competing factors to choose from, it is a challenge to zero in on a single factor and ascertain with an element of authority that *this* is the one that clearly counts. The literature on the effect of globalisation on income inequality has used a variety of methods to isolate the effect of globalisation and its dimensions on income inequality. Reading through this literature one feels rather encouraged to believe that this difficult task can fruitfully be undertaken. In this chapter, we travel through a small strand of this literature—the effect on relative income inequality coming through the goods market, especially those that pertain to international trade and technological change.

R. Bhattacharyya (✉)
Indian Institute of Foreign Trade, Kolkata Campus, Kolkata 700091, India
e-mail: branajoy@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_5,

The goods market influences some kinds of income inequality[1] by affecting the relative demand for skilled and unskilled labour. If unskilled labour intensive goods are imported and globalisation increases the influx of these goods then the unskilled labour intensive domestic import competing sector shrinks due to globalisation, reducing the demand of this kind of labour relative to skilled labour. As a result, the relative wage of unskilled labour at home decreases. Since unskilled labour always earns lower wages compared to skilled labour, income inequality sharpens in the post globalised situation. This phenomenon—the Stolper–Samuelson Theorem in the context of a two factor two commodity and two country world where the two factors are skilled and unskilled labour—has been alleged to be responsible for increasing wage inequality in developed countries especially the US from the 1990s when low price exports especially from underdeveloped countries (say, China) started to invade their markets.

It can be imagined that there will be many a proverbial slip between the cup and the lip if one were to be serious in ascertaining that *this* is what has happened at least for the US. First, what then has happened to income inequality in underdeveloped countries? By the logic of the Stolper–Samuelson Theorem export of unskilled labour intensive goods is expected to have increased the relative demand for unskilled labour intensive goods in these countries reducing income inequality there. Has that happened? Second, do the theoretical conclusions of the hypothetical two–dimensional world carry over to the real multidimensional world? Since empirical analysis cannot operate in a hypothetical vacuum (as the data that is its raw material derives from the real multidimensional world) it is important to know clearly what results we are seeking from the analysis. Finally as we have already hinted, what empirical strategy do we employ to delineate the effect of a particular factor from that of the rest? As we have already mentioned, the objective of this chapter is to summarise in a nutshell these three issues highlighted by the literature. Most of this can be achieved without any advanced level exposure to trade modelling and econometric analysis. It can thus be replicated quite easily provided one gets hold of the data. This in fact is what makes this literature especially useful for understanding the process by which a theory can be put into practice by applying data to it.

The rest of the chapter is arranged as follows: Sect. 2 reports what theory has to say. Since the theory mostly operates in two good two factor hypothetical world we restrict ourselves to this situation in this section. The following section uses the conclusions of this theory to zero in on the possible determinants of wage

[1] The discussion below follows the literature in defining income inequality by only looking at a broad feature associated with it: the *relative* income of high wage earners to low wage earners. In the literature 'High wage earners' and "low wage earners' have often been referred to 'skilled workers' and 'unskilled workers' or 'white collar workers' and 'blue collar workers' or even 'production workers' and 'managers'. We will however simply assume that there is a cut off income below which a worker is called 'low wage' and above which 'high wage'. The cut off income is assumed to be equivalently defined in terms of skill. Low skill workers earning low wages and highly skilled workers earning high wages.

inequality. The theory is exclusively based on goods supply and hence factor demand. However, we will look at both the demand and the supply sides of the factor market to arrive at these determinants. Once the theoretical points are clear and the determinants identified Sect. 4 turns to the empirical implementation of the results. This has several steps. Since the theory is constructed in a hypothetical world, it has first to be modified to yield results that are compatible to available data. After this is achieved an exact application of the method is demonstrated with data from the Korean economy. Section 5 goes beyond the exact empirical specifications discussed so far and very briefly reviews the results that have been reported in the literature. Since the main focus of this chapter is to demonstrate a method rather than make a comprehensive review of the existing literature we pick and choose only certain research works and attempt to arrive at broad conclusions from the huge list of papers that are available on the subject. The last section concludes the paper with a remainder of what we have intended to achieve in this chapter.

Demand for Labour in a Hypothetical Economy

To derive the relative labour demand function we will begin with the two good—two factor—two country open economy where the two economies are: home and rest of the world (aggregated as a single economy). Let the quantities of the two goods produced at home be denoted by X_1 and X_2 with P_1 and P_2 as their respective prices. Since the relative wage between low and high income people is our concern, we represent the two factors by 'high wage labour' (H) and 'low wage labour' (L) with w_H and w_L being their respective wages. Let $\bar{L}$ and $\bar{H}$ be the total availability of L and H in the home economy.

Production equilibrium for this economy is achieved at the point where the line showing the relative price of the two goods ($M = P_1\ X_1 + P_2\ X_2$) is exactly tangent to the production possibility frontier (PPF). This is represented by point P in Fig. 5.1. Since the value of this line (M) at the optimal point P corresponds to the total value of the two goods produced in equilibrium ($M^* = P_1 \leq X_1^* + P_2\ X_2^*$, where '*' represents optimal values) and since this economy produces only these two goods and nothing else, value of the line at the optimum point (M^*) is the value of GDP for the economy. Since stretching out this GDP line as far as possible subject to the resource constraint given by the PPF is the objective of any economy the production equilibrium in the two—good case mathematically looks like:

$$\text{Max GDP} : P_1X_1 + P_2X_2$$

$$\text{Subject to PPF} : X_1 = h(X_2, \bar{L}, \bar{H})$$

Or representing the PPF in this two good case is in its explicit (demand = supply of the two kinds of labour) terms:

$$\text{Max GDP} : P_1X_1 + P_2X_2$$

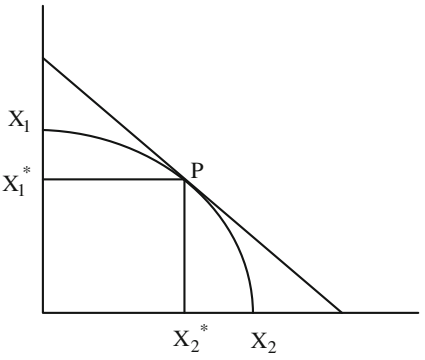


Fig. 5.1 The production equilibrium

$$\text{Subject to : } \begin{matrix} a_{L1}X_1 + a_{L2}X_2 \leq \bar{L} \\ a_{H1}X_1 + a_{H2}X_2 \leq \bar{H} \end{matrix} \tag{5.1}$$

Dual of the second exercise yields:
Min Cost : $w_L\bar{L} + w_H\bar{H}$

$$\text{Subject to : } \begin{matrix} a_{L1}w_L + a_{H1}w_H \geq \bar{P}_1 \\ a_{L2}w_L + a_{H2}w_H \geq \bar{P}_2 \end{matrix} \tag{5.2}$$

Production equilibrium point P in Fig. 5.1 is available either from the solution of the primal or the dual[2] or by considering, as Jones (1965) did, that the constraints (5.1) and (5.2) are satisfied in their binding forms:

$$\begin{aligned} a_{L1}X_1 + a_{L2}X_2 &= \bar{L} \\ a_{H1}X_1 + a_{H2}X_2 &= \bar{H} \\ a_{L1}w_L + a_{H1}w_H &= \bar{P}_1 \\ a_{L2}w_L + a_{H2}w_H &= \bar{P}_2 \end{aligned} \tag{5.3}$$

Note that if we arrive at (5.3) from this perspective, P_1 and P_2 are necessarily constant (as they are the constraints of the dual) and the economy is by definition a small open economy. Under the circumstances, the effect of trade on the economy is reflected by a change in the international price of the imported good.

The labour demand curve can be derived from (5.3). We begin by assuming that the a_{ij}'s in (5.3) are constant and that the production function is that of the fixed coefficient type. The shape of the relative labour demand curve depends on whether there is complete specialisation. If two good are produced then the first two equations and the last two equations in (5.3) are independent of each other. X_1 and X_2 are determined by the first two equations while the factor prices w_H and w_L are independently determined by the last two equations. Thus w_H/w_L is independent of $\bar{L}$ and $\bar{H}$ and hence $\bar{H}/\bar{L}$ and the relative labour demand curve is horizontal as in Fig. 5.2. If only one good (say X_1) is produced instead then (5.3) looks like:

[2] See Dorfman et al. (1958) Chaps. 2 and 3 and Dixit and Norman (1980) Chap. 2.

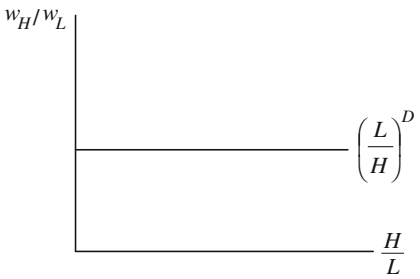


Fig. 5.2 Labour demand with incomplete specialisation

$$\begin{aligned} a_{L1}X_1 &= \bar{L} \\ a_{H1}X_1 &= \bar{H} \\ a_{L1}w_L + a_{H1}w_H &= \bar{P}_1 \end{aligned} \tag{5.4}$$

The first two equations in (5.4) are over determined and the last equation is under determined. Solutions for this case can be of two kinds: 1) a knife edge situation $\frac{\bar{L}}{a_{L1}} = \frac{\bar{H}}{a_{H1}}$ with full employment of both the factors. 2) Given the factor endowments $\bar{L}$ and $\bar{H}$, the output levels can be solved as $X_1 = \min\left[\frac{\bar{L}}{a_{L1}}, \frac{\bar{H}}{a_{H1}}\right]$. Suppose, $X_1 = \frac{\bar{L}}{a_{L1}}$ which implies skilled labour H remains unemployed. With flexible factor prices, this would drive down w_H to zero. Given the commodity price $\bar{P}_1$ and with $w_H = 0$, the unskilled wage rate w_L can be solved for from the pricing equation. Thus one of the factors (skilled labour) turns out to be a free factor.[3] Clearly, therefore, a meaningful relative labour demand curve cannot be derived. For want of a better term we call these cases "indeterminate" below.

To make things, a little bit more realistic let us now drop the assumption of the fixed coefficients production function. Instead, we assume that the isoquants are smooth and concave and factor usage responds to changes in factor prices. (5.3) and (5.4) will then have an additional equation[4] :

$$a_{ij} = a_{ij}(w_H, w_L) \tag{5.6}$$

With (5.6) the indeterminacy in (5.4) vanishes immediately as the three equations in (5.4) now become simultaneous equations and can together solve for w_L, w_H and X_1. However, (5.3) still remains dichotomous with the last two equations determining w_L and w_H independently of the first two equations. Thus w_H/w_L is still independent of $\bar{L}$ and $\bar{H}$ and hence $\bar{H}/\bar{L}$ in (5.3) and the relative labour demand curve is horizontal as in Fig. 5.2. However, when one good is produced and the equation system is represented by (5.4) along with (5.6) w_H/w_L and $\bar{H}/\bar{L}$ become dependent on each other and the labour demand curve will be of its normal down ward slope: to find employment, skilled workers' relative wages must fall as $\bar{H}/\bar{L}$ rises.

[3] I am deeply indebted to an anonymous referee for pointing this out to me.

[4] Note that with (5.6) (5.1) and (5.2) are no more linear programming exercises.

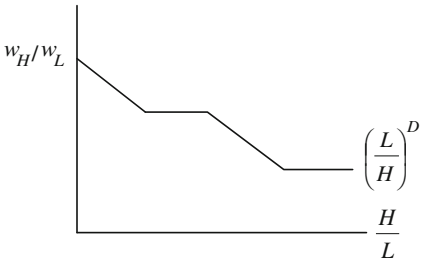


Fig. 5.3 The labour demand curve

Now consider a situation where there is a continuum of goods but only the above two factors. And assume that goods are arranged according to their skill intensities so that good 1 is the least skill intensive, good 2 is second and so on:

$$a_{1H}/a_{1L} < a_{2H}/a_{2L} < a_{3H}/a_{3L} \ldots \ldots \tag{5.5}$$

In such a situation as $\bar{H}/\bar{L}$ increases the labour demand curve looks as in Fig. 5.3:

1. If $\bar{H}/\bar{L}$ is very low only good 1 is produced
2. If $\bar{H}/\bar{L}$ rises a little both goods 1 and 2 are produced
3. If $\bar{H}/\bar{L}$ rises further only good 2 is produced
4. If $\bar{H}/\bar{L}$ rises further both goods 2 and 3 (but not good 1) is produced and so on.

Since labour is demanded by producers, we derived the relative demand curve of labour exclusively from the production equilibrium or the supply side for goods. There was no reference to commodity demand. Neither was there any reference to the word 'autarky'. However, this does not mean that commodity demand has no role in determining the relative demand for labour curve. Nor is autarky irrelevant to the above analysis. What is the labour demand curve under autarky? The answer to this question flags in these missing links succinctly.

Let us assume that both goods are produced under autarky. Now prices cannot be exogenously fixed. So (5.3) is underdetermined with four equations and six variables P_1, P_2, w_H, w_L X_1 and X_2. We need two more equations. These are the demand = supply equations for the two commodities X_1 and X_2. Assuming homothetic preferences (the so called Heckscher—Ohlin—Vanek model), a constant proportion (say α) of income ($M = w_H H + w_L L = P_1 X_1 + P_2 X_2$) goes to each commodity the two additional equations are:

$$\begin{aligned} \alpha(w_H H + w_L L) &= P_1 X_1 \\ (1-\alpha)(w_H H + w_L L) &= P_2 X_2 \end{aligned} \tag{5.7}$$

Note that (5.3) and (5.7) are always a set of simultaneous equations with or without the fixed coefficient production function assumption (with or without 5.6) and the labour demand curve, if it can be derived, is necessarily downward sloping (see Table 5.1).

Also missing is the case of a large open economy. In this case there are two sets of equations like (5.3) one for the home and the other for the foreign country (the 'rest of the world'). Further (5.7) changes to:

Table 5.1 Labour demand curves under various assumptions

Nature of home economy	Nature of technology	Number of goods produced (maximum of 2)	Shape of relative labour demand curve
Closed	Fixed coefficient	One	Indeterminate
		Two	Negatively sloping
	Variable coefficient	One	Negatively sloping
		Two	Negatively sloping
Small open economy	Fixed coefficient	One	Indeterminate
		Two	Horizontal
	Variable coefficient	One	Negatively sloping
		Two	Horizontal
Large open economy	Fixed coefficient	One	Indeterminate
		Two	Negatively sloping
	Variable coefficient	One	Negatively sloping
		Two	Negatively sloping

$$\begin{aligned}\alpha^H(w_H^H H^H + w_L^H L^H) + \alpha^F(w_H^F H^F + w_L^F L^F) &= P_1^H X_1^H + P_1^F X_1^F \\ (1-\alpha^H)(w_H^H H^H + w_L^H L^H) + (1-\alpha^F)(w_H^F H^F + w_L^F L^F) &= P_2^H X_2^H + P_2^F X_2^F\end{aligned} \quad (5.7')$$

The superscript H is for the home country and F for the foreign country. The resultant ten equations solve for the following ten variables:

$$w_H^H, w_L^H, w_H^F, w_L^F, P_1^H (= P_1^F), X_1^H, X_1^F, X_2^H, P_2^H (= P_2^F), X_2^F$$

Note that trade induces commodity prices to be equalised so that $P_1^H = P_1^F, P_2^H = P_2^F$.

Given the simultaneous nature of the system of equations both for the fixed and variable coefficient cases the labour demand curve for each country is once more negatively sloping under all circumstances for the large open economy.

Since there are three states of the economy (autarky, small open and large open), two assumptions regarding technology (fixed coefficients and variable coefficients) as well as the single good as well as the diversified (two) product case, there are as many as 12 situations under which the relative labour demand curve can be separately derived. These situations are represented in Table 5.1.

Possible Determinants of Income Inequality: A Demand–Supply Approach

Supply Side Determinates

In any country, relative wage between low and high wage individuals and hence income inequality is determined by the interaction between the relative demand and relative supply of low and high wage people. That is precisely how w_H/w_L is

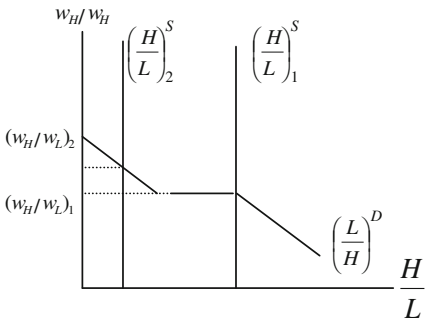


Fig. 5.4 A change in relative wage caused by a change in labour supply

determined in our model. However, to arrive there we need to figure out the relative supply curve in this model. This is easily done. In our model, H and L are available in fixed quantities $\bar{H}$ and $\bar{L}$ so that the relative labour supply curve is vertical (Fig. 5.4). If supply is $(\bar{H}/\bar{L})_1^S$ the equilibrium relative wage will then be $(w_H/w_L)_1$.

Figure 5.4 suggests that relative wage can change due to a host of factors either from the demand or the supply side. Changes in supply are easier to trace as the supply curve can only shift in a parallel fashion to the right or to the left. For example, suppose the country in question is India and there is a sudden surge in inflow of low income unskilled people from Bangladesh. The labour supply curve will then move to the left enhancing income inequality.

The labour supply curve can change due to many other reasons as well. Capital flows, level of female labour force participation, a rise in the proportion of old age people in the work force, changing returns to education leading to a change in the supply of educated workers, a rise in the number of engineering and medical colleges are some other factors that will shift the labour supply curve and hence affect income inequality (see Topel 1997 for a survey). Note that whether the economy will specialise completely or produce one good depends on the supply curve (the demand curve is clearly incapable of doing this). For instance, if it is at $(\bar{H}/\bar{L})_1^S$ then only one good is produced. On the other hand if it shifts to $(\bar{H}/\bar{L})_2^S$ then both are produced. Also note that if the labour demand curve is horizontal, the relative wage is insensitive to changes in factor supply. Referring back to Table 5.1 this implies that supply side factors never have any role on factor prices in a diversified small open economy. But in all other non-indeterminate cases, the role of factor supply cannot be discounted.

The position of the supply curve also determines the relative skill intensity of the goods. As the supply curve gradually shifts to the right the country graduates to higher and higher skill intensive goods. Thus as the supply curve continuously shift to the right the level of income inequality continuously falls. Since it is expected that a developed country will have a higher level of H/L than an underdeveloped country, a rightward shift of the supply curve is a sign of economic development. It is then expected that the wage inequality will be fall from

the supply side as the country becomes developed. However this is not what has happened in this world. Many developed countries have levels of income inequality that is higher than many underdeveloped countries. Clearly, therefore, the supply side of the picture is not adequate to explain reality. The demand side has to be factored in. To this we turn now.

Demand Side Determinants

Whether the labour demand function is horizontal or negatively sloping the reasons for its shift are same: anything that induces producers to offer changed relative wages to its workers *given the factor supply ratio H/L*. Three such changes are at the heart of the globalisation debate around the world. These are:

1. Change in technology
2. Outsourcing of jobs/intermediate inputs/production networking
3. International trade

In the first case either of two things can happen: unskilled workers are replaced by skilled workers in the production of the same goods or vice versa (as this implies that one factor is being replaced by another factor, this is called factor-biased technical change) or skill intensive goods experience higher levels of technical progress than unskilled labour intensive goods or the other way round (sector-biased technical progress). If the technology change comes through foreign direct investment, it is globalisation that lies at the heart of such changes. However note that technology may change due to domestic innovation as well. If this is so it is unrelated to globalisation.

The second situation may arise when production units in developed countries decide to locate a part of its assembly line to foreign countries. Since the developed country has a relative advantage in skill intensive production, those parts of the assembly line that are unskilled labour intensive are usually outsourced by them. Thus unskilled jobs 'fly out' and relative inequality rises. This is equivalent to replacing unskilled workers with skilled workers in the factory floor as the entire floor is now taken over by the skill intensive assembly lines. Even though no explicit technological change is required for parts of the shop floor to be vacated and parts of production processes to be shipped out, the result is effectively similar to a factor-biased technical progress.[5]

In the third case, reduced relative prices of unskilled labour intensive goods leads to shrinkage of the unskilled labour intensive sector. If that sector is the import competing sector then it is possible that this fall is due to trade-enhanced

[5] Note that in a small open economy only the sector bias of technology change matters for relative wages as opposed to the factor bias but for outsourcing only factor bias so outsourcing has no effect in a small open economy,.

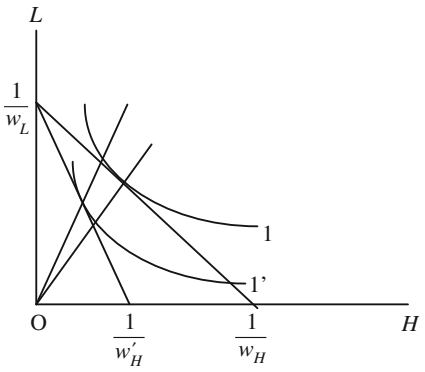


Fig. 5.5 Factor-biased technical progress

inflow of cheaper substitutes of goods produced by the import competing sector. Like outsourcing, this case also has the same impact on the labour demand curve as a change in technology—this time only of the sector-biased type—if the economy is small like in (5.3).

Technical progress thus turns out to be the most fundamental in terms of ideas that affects labour demand. To this let us turn now.

Since there is no scope for sector or good intensive technical progress in a one-good economy, it is a nice place to study the implications of factor-biased technical progress. This is done in Fig. 5.5 in which 1 is the unit value isoquant for the good. Since they are of unit value the common cost line is the unit cost line. Therefore, they have intercepts $1/w_L$ and $1/w_H$. The slope of this line determines w_H/w_L. In Fig. 5.6 the difference in this relative wage is caused by the change in the slope of isoquant 1 as it moves down due to technical progress. The change in slope is captured both through the changed relative factor price ratio that is tangent to the two isoquants and the fact that the expansion paths for 1 and 1' are different (the later using up less H as its relative price has increased). Since w_H has increased with a constant w_L, wage inequality has increased. This is a clear case of an upward shift of the negatively sloping labour demand curve (Fig. 5.6). Note that this is exactly as it would happen with outsourcing. 1' is steeper than 1, therefore

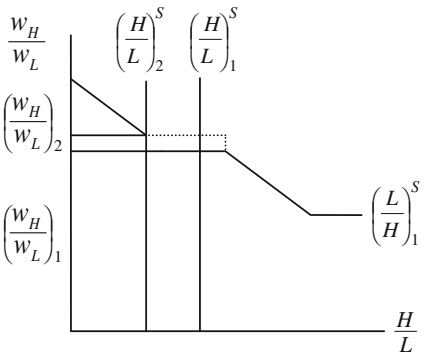


Fig. 5.6 A change in relative wage caused by a change in labour demand

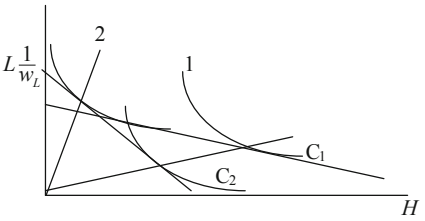


Fig. 5.7 Sector-biased technical progress

high wage labour is more effective in 1' than in 1. If jobs of low wage labour disappear from the developed country's shop floor low wage labour will obviously become less effective and will be paid a relatively lower wage. Cheaper low wage labour will automatically imply higher level of employment of that labour and it appears that though the relative wage of low wage labour falls their employment will rise. This is a limitation of the one sector model we are considering here. Let us therefore turn to the two sector economy.

To understand sector-biased technical progress we turn to the two good economy of Fig. 5.7. Here good 1 or sector 1 is skilled labour intensive and sector 2 is unskilled labour intensive. Technical progress occurs in the skilled labour intensive sector. However, there is no factor bias in this technical progress—isoquant 1 shifts down parallelly. Since no technical progress takes place in good 2 the new cost line tangent to both isoquants is C_2 rather than C_1 and income inequality once more rises.[6] Since once again this change has occurred without any change in H/L, the labour demand curve shifts up as in Fig. 5.6. This happens irrespective of whether the country is large or small and the demand curve is horizontal or negatively sloping.

Why is this sector-biased technical progress equivalent to trade? To understand this note that if the country in question is a developed country, with unskilled labour intensive imports, an increase in import due to globalisation will reduce the domestic price of the unskilled labour intensive good shifting the unit value isoquant 2 up with a similar effect on relative wages. In fact, as is well known, the exact relationship between prices and wages is given by the following Stolper–Samulson magnification effect:

$$\hat{w}_L > \hat{P}_2 > \hat{P}_1 > \hat{w}_H \tag{5.8}$$

where a '^' over a variable implies relative change.

Though the above statement cannot be generalised to the empirically relevant many (say, n) goods and many (say, m) factors case (see Jones and Schinekman 1977), a weaker statement can be made when $m = n$: while every good is a 'natural enemy' to some factor, a given good need not be a 'natural friend' to any factor.

[6] Figure can be drawn by the reader.

Empirical Application

Preparing the Theory So That it Can Receive Data

If one knows about the ideas just laid out one would still need to understand how to make it amenable to data. The theory that we discussed in the previous section approaches the issue of wage inequality exclusively from the factor demand side (the supply side of the *goods* market). The set of equations in (5.3) are the only set of usable equations. Of the four equations in (5.3) the last two are the key ones that determine wages if the economy is small. Assuming a small open economy we will thus have to manipulate these two to arrive at something decisive about the role of different factors on wages. In these equations, there are three variables: the wages, the prices and the a_{ij}'s. Thus only two determinants of wage inequality can be considered: prices (and anything that impacts upon it, like, for instance a sudden rise in the inflow of goods through trade) and technology (with a similar adage). In fact, as we will see in the next section technological change has turned out to be a major irritant in figuring out a story that attributes wage changes to trade.

Finally, the equations must be consistent with reality: (a) all variables must be measurable and data regarding them must be available (b) the number of goods and factors must be large enough to capture reality and (c) since all production processes involve the use of intermediate inputs, provision must be made for them—a tall order for theory as the structure will become much too complicated. But as we will see, meaningful empirical analysis can still be done.

Generalising the last two equations in (5.3) with (a)–(c) in mind we first write it as follows:

$$p_i = \sum_j a_{ji} w_j + \sum_k \gamma_{ki} p_k, i = k = 1, 2, \ldots, n. j = 1, 2, \ldots, m \tag{5.9}$$

Where goods are indexed by i (when they are used as outputs) and k (when they are used as inputs). Note that $i = k$. Factors are indexed by j. p_i and p_k are the prices of good i and good k. w_j is the price of factor j. a_{ji} is the amount of factor j used to produce one unit of good i. Similarly γ_{ki} is the amount of good k used to produce one unit of good i. Totally differentiating (5.9):

$$dp_i = \sum_j a_{ji} dw_j + \sum_j da_{ji} w_j + \sum_k \gamma_{ki} dp_k + \sum_k d\gamma_{ki} p_k$$

Thus,

$$\widehat{p}_i = \frac{dp_i}{p_i} = \sum_j \frac{a_{ji} w_j}{p_i} \widehat{w}_j + \sum_j \frac{a_{ji} w_j}{p_i} \widehat{a}_{ji} + \sum_k \frac{\gamma_{ki} p_k}{p_i} \widehat{p}_k + \sum_k \frac{\gamma_{ki} p_k}{p_i} \widehat{\gamma}_{ki}$$

Let $\frac{a_{ji} w_j}{p_i} = \theta_{ji}$ and, $\frac{\gamma_{ki} p_k}{p_i} = \Gamma_{ki}$.

Thus:

$$\hat{p}_i = \sum_j \theta_{ji}\hat{w}_j + \sum_j \theta_{ji}\hat{a}_{ji} + \sum_k \Gamma_{ki}\hat{p}_k + \sum_k \Gamma_{ki}\hat{\gamma}_{ki} \tag{5.10}$$

Assume that the production function is of the following form:

$$Q_i = A_i f(L_{1i}, L_{2i}, \ldots, L_{mi}, Q_{1i}, Q_{2i}, \ldots, Q_{ni})$$

where A_i is a constant. Q_i is the output of good i. L_{ji} is the amount of factor j used in the production of good i and Q_{ki} is the amount of good k used in the production of good i. Taking log derivative:

$$\hat{Q}_i = \hat{A}_i + \sum_j \frac{\delta f_i(.)}{\delta L_{ji}} \frac{dL_{ji}}{f_i(.)} + \sum_k \frac{\delta f_i(.)}{\delta Q_{ki}} \frac{dQ_{ki}}{f_i(.)}$$

or,

$$\hat{Q}_i = \widehat{A}_i + \sum_j \frac{\delta f_i(.)}{\delta L_{ji}} \frac{L_{ji}\hat{L}_{ji}}{f_i(.)} + \sum_k \frac{\delta f_i(.)}{\delta Q_{ki}} \frac{Q_{ki}\hat{Q}_{ki}}{f_i(.)} \tag{5.11}$$

Now the factor and intermediate good market equilibrium conditions are:

$$\begin{aligned} \frac{w_j}{p_i} &= \mathrm{MP}_{ji} = A_i \frac{\delta f_i(.)}{\delta L_{ji}} \\ \text{and,} \frac{p_k}{p_i} &= \mathrm{MP}_{ki} = A_i \frac{\delta f_i(.)}{\delta Q_{ki}} \end{aligned} \tag{5.12}$$

where MP is the marginal product. Putting (5.12) in (5.11):

$$\hat{Q}_i = \hat{A}_i + \sum_j \frac{w_j}{p_i A_i} \frac{L_{ji}\hat{L}_{ji}}{f_i(.)} + \sum_k \frac{p_k}{p_i A_i} \frac{Q_{ki}\hat{Q}_{ki}}{f_i(.)}$$

But, $A_i = \frac{Q_i}{f_i(.)}$.

Thus:

$$\hat{Q}_i = \hat{A}_i + \sum_j \frac{w_j L_{ji}}{p_i Q_i} \hat{L}_{ji} + \sum_k \frac{p_k Q_{ki}}{p_i Q_i} \hat{Q}_{ki} \tag{5.13}$$

Also note that $L_{ji} = a_{ji}Q_i$ and,

$$Q_{ki} = \gamma_{ki} Q_i \tag{5.14}$$

Thus (5.13) becomes:

$$\hat{Q}_i = \hat{A}_i + \sum_j \frac{a_{ji} w_j}{p_i} \hat{L}_{ji} + \sum_k \frac{\gamma_{ki} p_k}{p_i} \widehat{Q}_{ki} = \hat{A}_i + \sum_j \theta_{ji}\hat{L}_{ji} + \sum_k \Gamma_{ki}\hat{Q}_{ki}$$

(5.14) also implies $\hat{L}_{ji} = \hat{a}_{ji} + \hat{Q}_i$ and $\hat{Q}_{ki} = \hat{\gamma}_{ki} + \hat{Q}_i$. Putting these in (5.14):

$$\hat{Q}_i = \hat{A}_i + \sum_j \theta_{ji}\hat{a}_{ji} + \hat{Q}_i \sum_j \theta_{ji} + \sum_k \Gamma_{ki}\hat{\gamma}_{ki} + \hat{Q}_i \sum_k \Gamma_{ki}$$

Now from (5.8): $1 = \sum_j \frac{a_{ji}w_j}{p_i} + \sum_k \frac{\gamma_{ki}p_k}{p_i} = \sum_j \theta_{ji} + \sum_k \Gamma_{ki}$ so that: $\sum_j \theta_{ji} = 1 - \sum_k \Gamma_{ki}$. Putting this in (5.8) we have:

$$\hat{A}_i = -\sum_j \theta_{ji}\hat{a}_{ji} - \sum_k \Gamma_{ki}\hat{\gamma}_{ki} \tag{5.15}$$

Putting (5.15) in (5.10) we have:

$$\hat{p}_i = -\hat{A}_i + \sum_j \theta_{ji}\hat{w}_j + \sum_k \Gamma_{ki}\hat{p}_k$$

Thus,

$$\hat{p}_i - \sum_k \Gamma_{ki}\hat{p}_k = -\hat{A}_i + \sum_j \theta_{ji}\hat{w}_j$$

Defining $\hat{p}_i - \sum_k \Gamma_{ki}\hat{p}_k = \hat{P}_i$ as the 'value added prices' and noting that A_i is the total factor productivity (TFP) of good i we get: $\hat{P}_i = -T\hat{F}P_i + \sum_j \theta_{ji}\hat{w}_j$

We can separate this relative price change ($\hat{P}_i$) in two parts (a) that caused by technology ($\hat{P}_i^{te}$) and (b) that caused by other factors including international trade ($\hat{P}_i^{tr}$). Accordingly, the total change in wages ($\hat{W}_j$) can be separated into that due to technology ($\hat{w}_j^{te}$) and that due to trade ($\hat{w}_j^{tr}$):

$$\hat{P}_i = \hat{P}_i^{te} + \hat{P}_i^{tr} \tag{5.16}$$

where

$$\widehat{P}_i^{te} = \sum_j \theta_{ji}\widehat{w}_j^{te} - \mathrm{T}\widehat{\mathrm{F}}\mathrm{P}_i \tag{5.17}$$

And

$$\widehat{P}_i^{tr} = \sum_j \theta_{ji}\widehat{w}_j^{tr} \tag{5.18}$$

Let us further assume that a certain proportion (λ) of productivity improvements are directly passed on to prices:

$$\widehat{P}_i^{te} = -\lambda\mathrm{T}\widehat{\mathrm{F}}\mathrm{P}_i \tag{5.19}$$

so that (5.17) yields:

$$(1-\lambda)\widehat{\mathrm{TFP}}_i = \sum_j \theta_{ji}\widehat{w}_j^{te} \tag{5.20}$$

Now, from (5.16) $\hat{P}_i^{tr} = \hat{P}_i - \hat{P}_i^{te}$ and using (5.19):

$$\widehat{P}_i^{tr} = \widehat{P}_i + \lambda\widehat{\mathrm{TFP}}_i \tag{5.21}$$

Putting (5.20) in (5.18) we have:

$$\widehat{P}_i + \lambda\widehat{\mathrm{TFP}}_i = \sum_j \theta_{ji}\widehat{w}_j^{tr} \tag{5.22}$$

Leamer (1996) estimates $\hat{w}_j^{te}$ from Eq. (5.21) and $\hat{w}_j^{tr}$ from (5.22) and compares them with the observed $\hat{w}_j$ to determine the effect of technology and trade on wages.

It is therefore interesting to note that we are seeking a completely different genre of results than the popular magnification effect associated with the Stolper–Samuelson Theorem. We are not even seeking friends and enemies a la Jones and Schinekman (1977) in the empirically more relevant many goods many factors case. Question asked is: is there a pass through from price change to wage change? The presumption is that if we indeed find such a link then whatever pattern is observed for wage changes (like a rise or fall in wage inequality) can be attributed to price changes and hence international trade.

Application: Estimating Equations (5.21) and (5.22) for the Korean Data[7]

Korean Wage Inequality

Figure 5.8 shows the time series trend of Korea's relative (w_H/w_L) and absolute ($w_H - w_L$) wage during the last 25 years of the twentieth century—the initial phase of globalisation. Interestingly, relative wages of high income people have fallen over time reducing relative income inequality in Korea, but absolute income inequality has risen. This observation can be explained as follows.

Let W_A and W_R be the absolute difference in wages between skilled and unskilled workers ($w_H - w_L$) and the skill premium (w_H/w_L), respectively. Let us further assume that all derivatives are with respect to time ($dx = dx/dt$ for all x) and denote

[7] Research support from the Asian Scholarship Foundation, Bangkok, is gratefully acknowledged. I am indebted to Jong Eun Lee for discussions. I am also deep; y indebted to Se Kwon Joo for research assistance.

all relative changes over time by '^' $\left(\hat{x} = \frac{dx/dt}{x}\right)$

Now since, $W_A = w_L(W_R - 1)$ therefore, the absolute wage gap will rise if:

$$\hat{W}_A = \hat{w}_L + \frac{W_R}{W_R - 1}\hat{W}_R$$

Assuming $\hat{w}_L$ to be positive (unskilled wages increase over time) if relative wages are rising, absolute wages must rise. However, if relative wages are falling absolute wages may still rise provided the fall is small enough so as not to violate the above equation. In Korea, therefore even though relative wages are falling they are not falling fast enough for absolute wages to fall as well. In other words The Korean case (as depicted in Fig. 5.8) is:

$$\hat{w}_L > \hat{w}_H > \hat{W}_A > 0 > \hat{W}_R$$

and $\hat{W}_R$ is not negative strongly enough for it to materialise into equality. The result can also be stated in terms of the absolute level of skilled and unskilled wages. For this note that $\hat{W}_A > 0$ if $w_H > (2 - \beta)w_L$ where $\beta = \hat{w}_H/\hat{w}_L$. Since in Korea $\hat{w}_H < \hat{w}_L$ therefore, $0 < \beta < 1$ and hence $1 < (2 - \beta) < 2$. Therefore, in Korea the high absolute gap between skilled and unskilled wages has resulted in the simultaneous occurrence of a rising W_A and a falling W_R. In the $2 \times 2 \times 2$ economy, Stolper–Samuelson framework this is then what has happened in Korea (assuming commodity 1 to be H intensive and 2 to be L intensive):

$$\hat{w}_L > \hat{P}_2 > \hat{P}_1 > \hat{w}_H > \hat{W}_A > 0 > \hat{W}_R$$

However, this implies that as far as absolute wages are concerned the Stolper–Samuelson result cannot be generalised to the m factor and n commodities case even when $m = n$ as mentioned above. How do we define 'natural' enemies and friends? This can be understood by taking a look at the last inequality stated above. According to the definition employed by these authors in this inequality good 2 is a 'natural friend' of labour U and good 1 is its 'natural enemy'. Why? Because if $\hat{P}_2 > \hat{P}_1$ then $\hat{W}_R$ is negative. However it is clear that this is not sufficient for making $\hat{W}_A$ negative. The Stolper–Samuelson result in the form that it has been treated here, therefore, cannot generalise to defining 'friends' and 'enemies' in terms of W_A.

Applying Data to (5.21) and (5.22)

Before turning to the equations, Table 5.2 reports the simple correlation between changes in different factor prices[8] and (a) simple price inflation (b) $\hat{P}_{it}$ (c) $\widehat{\text{TFP}}_{it}$.

[8] The Wage Structure Survey reports wages of males, females and total. Throughout this section we have used the figures reported under 'total'.

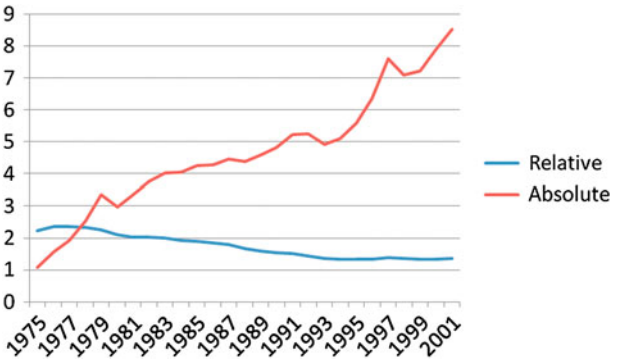


Fig. 5.8 Relative and absolute wage differential in Korea*Note* 1. Relative = Wages of University (UNIV) + Wages of college graduates (COL)/Wages of middle school (MS) + Wages of high school pass outs (HS). 2. (UNIV + COL) − (MS + HS). 3. Absoulte difference in wages is divided by 1,00,000 so that it can be reported in the same diagram as wage ration. 4. All wages are monthly earnings in Won *Source* Calculated from data in National Statistical Office, Korea

First, note that throughout this section $\hat{x}_{it}$ is determined as follows: $\ln x_t - \ln x_{t-1}$. For this reason, we write $\Delta \log x_t$ instead of $\hat{x}_t$ in Tables 5.2–5.4. Turning now to Table 5.2 ,it is clear that all the factor prices have a higher correlation with TFP than with price inflation whether adjusted to intermediate good prices or without such adjustments. However, 'senior college and over' is the only factor that has negligible correlation with *all the variables*. This of course does not mean that mandated changes of wages for this variable would always be insignificant. Mandated changes in Eqs. (5.21) and (5.22) determine the extent to which price of a particular factor has to change corresponding to the actual changes in $\hat{P}_{it}$ or $\widehat{\text{TFP}}_{it}$ and θs to satisfy the zero profit conditions dictated by the assumption of perfect competition. Thus the correlation between returns to a particular factor with $\hat{P}_{it}$, $\widehat{\text{TFP}}_{it}$ and their mandated changes are un-correlated. However, it should be noted that the simple correlation results point to a greater promise for $\widehat{\text{TFP}}_{it}$ than for $\hat{P}_{it}$.

Table 5.2 Simple correlations between factor prices, price inflation and change in TFP

Changes in wages per hour ($\Delta \log w_{it}$)	$\Delta \log P_{it}$	$\Delta \log \text{TFP}_{it}$
Under middle school	0.11 (0.26)	0.13
High school	0.13 (0.15)	0.20
Junior college	0.14 (0.18)	0.31
Senior college and over	−0.02 (−002)	−.005
Unskilled	0.15 (0.22)	0.19
Skilled	0.03 (0.08)	0.27

Note The figures in the brackets are correlations with prices unadjusted to intermediate good prices

Table 5.3 Simple correlations between θ_{jit}, price inflation and change in TFP

Factor	$\Delta \log P_{it}$	$\Delta \log TFP_{it}$
Under middle school	0.09	0.15
High school	0.05	−0.03
Junior college	−0.04	−0.06
Senior college and over	0.01	−0.31
Unskilled	0.09	0.19
Skilled	−.002	−0.23

Table 5.4 The mandated wage regressions (Eqs. 5.2 and 5.3), 1983–1996

Factor	$\lambda = 0$		$\lambda = .5$		$\lambda = .9$	
	$\Delta \log P_{it}$	$\Delta \log TFP_{it}$	$\Delta \log P_{it}$	$\Delta \log TFP_{it}$	$\Delta \log P_{it}$	$\Delta \log TFP_{it}$
Under middle school	−0.13	0.005	−0.12	0.002	−0.12	0.0001
	(−0.82)	(3.27)	(−0.73)	(3.27)	(−0.75)	(3.27)
High school	0.81	0.006	0.78	0.19	0.79	0.0001
	(1.18)	(2.99)	(1.11)	(2.99)	(1.12)	(2.99)
Junior college	−8.05	0.05	−7.60	0.02	−7.87	0.004
	(−1.51)	(2.75)	(−1.35)	(2.75)	(−1.40)	(2.75)
Senior college and over	0.63	0.01	0.54	0.006	0.62	0.001
	(0.48)	(2.74)	(0.39)	(2.74)	(0.44)	(2.74)
Unskilled	0.003	0.003	0.006	0.001	0.009	0.0001
	(0.05)	(3.44)	(0.07)	(3.44)	(.12)	(3.44)
Skilled	−0.25	−0.07	−0.27	−0.065	−0.27	−0.071
	(−0.78)	(2.27)	(−0.82)	(2.27)	(−0.82)	(2.27)
Number of observations	48	48	48	48	48	48

Note The figures in the brackets are *t*-values

Table 5.4 estimates Eqs. (5.2) and (5.3) for three different values of λ and Table 5.3 gives an indication of the R^2s of these regressions. There are several important points to note before reading the table. First and most importantly, the table does not report the result of all the basic industries reported. We have taken only three basic industries for this table: mining, manufacturing and 'electricity, gas and water supply'. This is because price data for the other industries are not available. The price data used for this table is the producer's price index reported by the National Statistical Office, Korea. Second, input–output tables are calculated every five years in Korea (by the Bank of Korea). Hence data for a particular year (for example 1995) was used in calculating value added prices for the previous four years also (1991–1995). Third, we have not taken capital as an additional variable though this is customary in the literature. This is because the capital stock data was found to be unreliable. Also there is no clear proxy for the rental data. However, the TFP calculation uses capital as a factor. Fourth, the labour data is used from the Wage Structure Survey and it has all the limitations of sampling that the survey has (especially the fact that it leaves out the informal sector). Thus the labour data comes with bias. Fifth, since gross output data by industry could not be found, therefore GDP accounted for by the industry is taken as a proxy for it.

Let us first turn to the first four rows of the table. The question that is posed here is: what is the mandated wage if each kind of labour is taken as a separate factor of production. However, there are serious problems of running such a regression. The correlation between θ_{HSi} and θ_{COLi} is .84 and that between θ_{COLi} and θ_{UNIVi} is .90. It is therefore obvious that if we keep all of these in one regression the sampling variances and covariances will be large and the coefficients will be sensitive to small data changes. The strategy that we adopt therefore is to determine whether the mandated changes are significant or not. Thus in the first four rows of Table 13 each of the independent variables are taken one at a time. This makes the actual values of their coefficients meaningless as far as indicators of mandated changes are concerned. However, we can determine whether these coefficients will be significantly different from zero by looking at their t-values. A look at the t-values reported in Table 5.4 clearly shows that mandated changes for TFP are significant for the factors but the mandated changes for trade are not so. We can thus conclude from the table that under the assumptions stated above international trade does not appear to mandate any statistically significant change on wages of Korean workers. The variable that appears to be important was technology and not trade.

The problem of multicollinearity disappears when we take skilled and unskilled wages that is θ_{Ski} and θ_{USKi}. The correlation between these variables is −.02. Thus we have taken them together in the last two rows of the table. Thus *if skilled labour and unskilled labour were to be the only two factors of production in our perfectly competitive system* then trade (technology) accounted for between .3 (.01) and .9 (.3) % of unskilled and between –25 (−.7) and –27 (−6.5) % for skilled wages. However, since the t-values for trade are all insignificant we can consider these coefficients to be zero and it can be stated that international trade mandated an insignificant change in wages in Korea. For technology, the changes, though significant, are very small for unskilled labour but relatively high for skilled labour. Since the simple correlation results, the mandated regression results in the first four rows of Table 5.4 and the mandated regression results in the next two rows of Table 5.4 independently point in the same direction we conclude that trade did not mandate any significant change on Korean wages. However, technology mandated a higher fall on unskilled wages than for skilled wages though their magnitudes are small.

What can We Say About the Effect of Globalisation on Wage Inequality?

Haskel and Slaughter (1999) and Leamer (1996) perform an analysis of the link between trade, technology and wage inequality by using exactly the above method for UK and US, respectively. First, both observe that wage inequality in their respective countries has increased though for different periods of time. Second, both find evidence that trade did play a role in this change in both countries.

For UK, the role of technology is more robust while the reverse is true for the US. Generally in European countries where stronger unions resist wage declines, the effect of globalisation is reflected more in unemployment of low skilled workers rather than their wages while the reverse is true for the US. One of the dominant arguments for the developed countries has been that the volume of trade between developed and underdeveloped countries have been too low for it to have an effect on developed country wages (Krugman 2000). However, recent evidence seems to suggest that in spite of the smallness in volume the effect has been rather significant (see Corsini 2011)

Contradicting the predictions of the Stolper–Samuelson result in (5.8) some of the sharpest rises in income inequality during the period of globalisation has been in the developing countries. Zhu and Trefler (2011) point this out and explain this in terms of an increase in the export shares of skill intensive goods of the underdeveloped countries. Acharyya (2006) also writes about the occurrence of the same phenomenon for India and other countries and try to theoretically rationalise it. Further to what these authors have pointed out, looking at this unexpected symmetric effect of globalisation on both sets of countries it appears that 'skill' and 'unskilled' labour intensive goods cannot be unequivocally defined across countries. With vast difference in their base skill levels, goods that are termed 'skilled' labour intensive for underdeveloped countries usually turn out to be unskilled labour intensive for developed countries. Thus while wage inequality may well rise in developing countries due to the rise in exports of 'skilled' labour intensive goods, these same goods may turn out to be unskilled labour intensive as they enter the markets of developed countries increasing wage inequality there as well. This aspect of the skilled–unskilled debate has not been addressed adequately in the empirical literature on wage inequality.

For India, the ratio of skilled to unskilled wage increased moderately form the mid-1970s to the early 1980s and then after a long period when it was roughly constant once again increased rather sharply from the late 1990s to the early 2000s (see Srivastava and Mathur 2011). The rather long period of near constancy in this ratio from the mid-1980s to the late 1990s—a period of rapid trade liberalisation—has made causality analysis from trade liberalisation to wage inequality difficult for India. There are thus conflicting conclusions with researchers swinging from one extreme position—like trade liberalisation decreased rather than increased income inequality in India (Mishra and Kumar 2005)—to the other extreme, that trade liberalisation increased income inequality (Acharyya 2006). There are authors who have also brought in the technology angle as well as other aspects of globalisation like foreign direct investment (FDI). For instance, Banga (2005) find skill-biased technological progress and FDI as more responsible than trade in affecting wages in India. Indirect aspects of globalisation—like the replacement of permanent workers by contractual workers as labour laws become lax have also been brought in by researchers. For example, Ramaswamy (2008) find that the positive contribution of change in output (scale effect), capital-output ratio and contract-worker intensity to wage inequality in Indian manufacturing. To the best of my knowledge, the above approach of calculating the pass-through of change in

prices and technology to wages has not been applied to India as well as most other developing countries. Since technology has changed quite rapidly in these countries during the period of globalisation, it would be interesting to know what their effect has been relative to trade.

Putting Theories Into Practice: Concluding Comments

Theory and data make for strange bedfellows. Theory cannot be done away with altogether while analysing data for fear of losing a mental bearing on it, but taking theory's help while analysing data allows too little leeway in explaining the data satisfactorily. Since economists take the first part of the above statement more seriously than the later, what they ultimately end up doing is estimating theories. In this chapter, we provided an example of how this can be done: understanding what the theory suggests in its own hypothetical cocoon, preparing it to face the world of data as best as is possible and then subjecting it to the data to see how the theory fares. The limitation of the approach is that the estimable equation derived from the theory has too much of theory embedded in it to be flexible enough to interpret all kinds of data. For instance (5.21) and (.5.22) have all the usual assumptions associated with the Heckscher–Ohlin theory: constant returns to scale, perfect competition in the factor market, costless inter sector mobility of factors of production etc. Without these assumptions (5.9) is not valid. Therefore, (5.21) and (5.22) are not valid and the empirical exercise cannot be conducted. But there is a lot that has been achieved as well. The main contribution of the approach is that it gives us a methodology to subject the data to a logical process provided by the theory and squeeze some sense from it so that it looks less chaotic than before. Further since the initial conditions (assumptions) under which the empirical exercise is conducted are well known, the exact context of the conclusions is clear to the researcher. From both counts the approach is capable of generating many more research insights: either by replicating the method for different data sets (as we did for Korea) or by taking up the more challenging option of changing the assumptions and re-deriving the estimable equations in the context of the new assumptions (say, increasing returns to scale) and then subjecting it to the old data sets to find out how sensitive the conclusions are, to the changes in initial conditions set for the exercise. It is thus a good launching pad for further research – a great opportunity for young researchers to test their skills.

References

Acharyya R (2006) Trade liberalisation, poverty, and income inequality in India, INRM Policy Brief No. 10 ADB

Banga R (2005) Liberalization and wage inequality in India. Working Paper No. 156, Indian Council For Research on International Economic Relations, New Delhi

Corsini L (2011) Is there really no link between international trade and wage differentials? A cross-country analysis. Discussion Paper no. 118, Collana di E-papers del Dipartimento di Scienze Economiche—Università di Pisa

Dixit A, Norman Victor (1980) Theory of international trade: a dual general equilibrium approach. Cambridge University Press, Cambridge

Dorfman R, Samuelson Paul, Solow A, Robert M (1958) Linear programing and economic analysis. The RAND Corporation, New York

Haskel J, Slaughter MJ (1999) Trade technology and UK wage inequality. NBER Working Paper No. 6978

Jones RW (1965) The structure of general equilibrium Models. J Polit Econ 73(6):557–572

Jones RW, Schinekman JA (1977) The relevance of the two-sector production model in trade theory. J Polit Econ 85(5):909–935

Krugman PR (2000) Technology, trade and factor prices. J Int Econ 50(1 Feb):51–71

Leamer EE (1996) In search of Stolper Samuelson effects on US wages. NBER Working Paper No. 5427

Mishra P, Kumar U (2005) Trade liberalization and wage inequality: evidence from India. International Monitary Fund, Working Paper no. WP/05/20

Ramaswamy KV (2008) Wage inequality in Indian manufacturing: is it trade, technology or labour regulations? WP-2008-021. Indira Gandhi Institute of Development Research, Mumbai

Srivastava A, Mathur SK (2011) Rising wage inequality in India: a translog cost function analysis, J Bus Policy Res 6(1):1–15

Topel RH (1997) Factor proportions and relative wages: the supply-side determinants of wage inequality. J Econ Perspect 11(2) (Spring, 1997):55–74

Zhu SC, Trefler D (2011) Trade and inequality in developing countries: a general equilibrium analysis. Michigan State University, Mimeo

Chapter 6
Production Performance of Indian Agriculture in the Era of Economic Reforms

Sankar Kumar Bhaumik and Sk. Abdul Rashid

Introduction

Agriculture was the predominant sector of the Indian economy at the time of the independence. About 55 % of GDP came from agriculture and about 70 % of the workforce was engaged in this sector at that time. So the planners and policy-makers emphasized on improving the performance of the agriculture sector in the initial Five-Year Plans as they realized that overall development of vast majority of the people in the country and achievements of developmental goals such as reduction of unemployment, poverty, malnutrition, and so on could not be fulfilled without it.

There is now a vast literature that describes the policies adopted since the independence for agricultural development in India. Broadly speaking, there are three distinct phases through which agricultural development in the country proceeded since the independence (Venkateswarlu 2008; Chand 2009). In the first phase, during the pre-green revolution period (till mid-1960s), the implementation of land reforms was undertaken along with the policies and programmes to increase food grains production; in the second phase, during the green revolution period (1966–1991), the bio-chemical and mechanical technologies were adopted to increase agricultural production and productivity; and in the third phase, the post-reform period, the Indian agriculture was influenced by the Structural Adjustment Programme (SAP) and India's entry into the WTO.

The land reforms measures comprising abolition of intermediaries, tenancy reforms, and imposition of ceiling laws were implemented throughout the 1950s and the early half of 1960s to change the feudal institutional agrarian structures.

S. K. Bhaumik (✉)
Department of Economics, University of Calcutta, 56A, BT Road, Kolkata 700050, India
e-mail: bhaumiksk@yahoo.co.in

Sk. A. Rashid
Raja Birendra Chandra College (affiliated to University of Kalyani), Kandi, Murshidabad, West Bengal, India
e-mail: rashidsa123@rediffmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_6, © Springer India 2013

Simultaneously, increasing food grains production was accorded top priority as the country was facing acute food shortage at that time. To begin with, the Grow More Food Campaign was launched. Several other programmes such as the Community Development Programme, National Extension Programme, Intensive Agricultural District Programme, Intensive Agricultural Areas Programme etc. were also implemented. In spite of all these efforts, the country had to depend on import of food grains during this period as the levels of production were inadequate to meet domestic demand (Venkateswarlu 2008, p. 30).

To boost agricultural production and to make India self-sufficient in food grains production, the green revolution strategy was adopted since mid-1960s. The hallmark of such a strategy was introduction of high yielding variety (HYV) seeds, fertilizers, and irrigation technology package (known as bio-chemical inputs), specifically in the resource-rich areas of the country (Punjab, Haryana and western Uttar Pradesh, mostly). Subsequently, several other supports were extended toward the agricultural sector such as extension services, credit facilities, price support, cheap (subsidized) supply of inputs, and so on. Further, the public investment for agricultural development (especially for building irrigation infrastructures) was raised which also encouraged expansion of private investment, thanks to the complementary relation between these two types of investment. Consequent upon all these efforts, the country became self-sufficient in food grains production by the late 1970s, overcoming the painful memories of previous food crises. Even during the decade of 1980s, the progress of the agriculture sector as a whole, as measured through growth rate of value of agricultural output, was commendable. As shown by Bhalla and Singh (2012, p. 27), while the value of agricultural output in India grew at the rate of 2.24 % per annum during 1962–1965 to 1980–1983, the corresponding figure became 3.37 % during 1980–1983 to 1990–1993.

Nevertheless, the whole approach toward agricultural development underwent a fundamental change ever since India's adoption of economic reforms policies in 1991. The post-1991 period witnessed implementation of the SAP policies, followed by India's entry into WTO in 1995. As is well-known, the package of macroeconomic and trade policy reforms introduced in 1991 consisted of macroeconomic policy changes, change in exchange rate, trade liberalization, devaluation of currency, gradual dismantling of the industrial licensing system and controls, reduction of tariffs, reform of public enterprises, increasing privatization, and so on (Joshi and Little 1996). It is true that no direct reference was made to agriculture in this new policy package. However, the protagonists of economic reforms policies argued that the new policy framework would definitely benefit tradable agriculture by ending discrimination against it and by turning the terms of trade in its favor.[1] This, in turn, is supposed to promote exports leading to rapid agricultural growth.

The consequence of such a policy shift on the Indian economy in general and agricultural economy in particular has naturally drawn widest possible research

[1] See, for example, Singh (1994), Ahluwalia (1996), Gulati (1998), Gulati and Kelley (1999).

attention in recent years. This is because of several reasons: First, a good number of scholars observed that, contrary to expectations, the Indian agricultural performance actually deteriorated in the era of economic reforms so much, so that the country has eventually landed into a phase of 'agrarian crisis'. Second, it has been noted that compared to the pre-reforms period (mid-1960s through end of 1980s), the rate of employment expansion in rural areas as well as poverty reduction declined in the era of economic reforms (1990s and beyond). This has happened primarily due to decline in growth of agricultural production. Third, and more importantly, as India adopted the policy of 'inclusive growth' in the 11th and 12th Five-Year Plans, it would be necessary to improve agricultural performance significantly in the absence of which the growth processes would continue to remain exclusionary, as in the past, and the government will not be able to attain significant reduction in poverty.[2] In fact, the projections done by the policy-makers show that the country will not be able to achieve its targeted 8–9 % growth rate of overall GDP (which is necessary to ensure success of inclusive growth strategy) unless the GDP accruing from agriculture and allied activities grow at the rate of 4 % per annum. In any case, the performance of the agriculture sector has to be improved substantially if at all India has to fulfill its dream of attaining 'inclusive growth'.

It is in the above context that we take a fresh look at the production performance of agriculture in India during the period of economic reforms. Specifically, we seek to examine the production performance of agriculture in India and her 17 major states during the period of economic reforms (1995–2010), and compare the same with the performance during pre-reforms period (1980–1995). Although several other scholars have also studied production performance of Indian agriculture in the era of economic reforms, our study is different from them at least in three respects. First, unlike most other scholars, we look into production performance of agriculture in India at a disaggregated level—individual states and individual crops, which is necessary to devise more meaningful policies to improve agricultural performance in the country, and reducing inter-state disparity in this regard. Second, most scholars studied the Indian agricultural performance by dividing both the pre- and post-reforms period into different sub-periods (mostly of 5 years duration).[3] In our opinion, such exercises provide an understanding of the short-run behavior of Indian agriculture.[4] However, to obtain an overall view of the impact of economic reforms on agriculture, we need to adopt a medium to long-term perspective, which is possible when we compare the whole

[2] Recent research has shown that the elasticity of poverty with respect to agricultural GDP is greater compared to elasticity of poverty with respect to non-agricultural GDP. In other words, expansion of agricultural GDP reduces poverty more compared to expansion of non-agricultural GDP. See Cristiaensen et al. (2006, 2010a), Cristiaensen et al. (2010b), de Janvry and Sadoulet (2009), Ravallion (2010).

[3] See Chand et al. (2007) and Chand and Parappurathu (2012).

[4] This is not to say that this issue is unimportant. It is indeed important if our objective is to understand the fluctuations in agricultural performance in the era of reforms.

of post-reforms period with a pre-reform period of comparable duration.[5] This is precisely what we have attempted to do in this chapter. Third, by considering data for latest available year, our study is able to offer an up-to-date understanding of the impact of economic reforms and globalization policies on the agricultural sector in India. Apart from discussing production performance, this chapter also highlights briefly some of the constraints faced by agriculture in India which impinge upon its performance in the era of economic reforms.

This chapter is divided into five sections. Following by the introduction, we provide a description of data sources and methodology followed in this study in Sect. 2. In Sect. 3, we seek to understand production performance of Indian agriculture during pre- and post-economic reforms periods by comparing the growth rates of agricultural production for these periods. Such a comparison has been carried out for all-India and 17 major states, both at levels of individual crops and groups of crops. Some of the constraints faced by Indian agriculture have been discussed in Sect. 4. The final section summarizes our main findings and suggests some policies to improve production performance of Indian agriculture in future.

Data Sources and Methodology

The data that we have used in this study are those of area and production of principal crops for the period of 1980–1981 to 2009–2010 for all-India as well as 17 major states. These data have been collected mostly from the Centre for Monitoring the Indian Economy (CMIE) reports (various issues). Some data have also been gathered from the Ministry of Agriculture and Cooperation, Government of India, New Delhi and the website www.indiastat.com. For land and inputs use data, apart from these sources, we have depended on the works of some past researchers.

In order to ensure inter-temporal comparability among the states, we made some initial adjustment to data. As the states of Jharkhand, Uttaranchal, and Chhattisgarh have been carved out of Bihar, Uttar Pradesh and Madhya Pradesh respectively, we have merged Bihar with Jharkhand, Uttar Pradesh with Uttaranchal, and Madhya Pradesh with Chhattisgarh. The absolute figures on area and production of various crops for individual states helped to arrive at the figures for the combined states. Thus, the figures of Bihar, Uttar Pradesh, and Madhya Pradesh refer to those of the undivided states.

In this study, we are required to compute, inter alia, the growth rates of food grains, non-food grains and all-crop production for the two periods (pre-reform and post-reform) to provide an overall understanding of changes in growth performance of agriculture in India and her states between these periods. As the series representing indices of these crop groups are not available, particularly at the state level,

[5] This is also the approach adopted by Bhalla and Singh (2001, 2012) in their studies on Indian agriculture.

we have constructed these series for the entire period of our study (1980–1981 to 2009–2010). To construct these series, we have first converted physical quantities of production of various crops into value terms, and added those. For this purpose, we have used all-India prices for those crops for the triennium ending 2003–2004.[6] The values of production obtained here have been also utilized for estimating the growth rates of production of individual crops and groups of crops.

The growth rates estimated in our study are all exponential growth rates. The growth rates of production for individual crops/groups of crops have been computed for two sub-periods—1980–1981 to 1994–1995 and 1995–1996 to 2009–2010. The former period represents the pre-reforms period while the later represents post-reforms period. It is expected that some changes would be reflected in the production performance of agriculture in India following adoption of economic reforms policies and globalization of agriculture in the aftermath of India's signing of Agreement on Agriculture in 1994 as a member of the WTO.

To ensure comparability of growth rates for two different sub-periods, we have estimated the 'kinked exponential model' as proposed by Boyce (1987). The main advantage of this model is that it eliminates 'discontinuity' between different sub-periods which automatically arises when separate exponential functions are fitted for different sub-periods. Another advantage of Boyce-methodology is that it enables us not only to obtain comparable growth rates for two different sub-periods by fitting a single function but also to examine the presence of a 'trend-break' in the growth rate of production.

The kinked exponential model estimated by us takes the following form:

$$ln\ Q = a + b_1 D_1 t + b_2 D_2 t$$

where $D_1 = 1$ in 1980–1981 to 1994–1995 and 0 in 1995–1996 to 2009–2010[7];

$$D_2 = 1 - D_1;\ \text{and}$$

$$t = -14.5, \ldots, -0.5,\ 0.5, \ldots, 14.5$$

After estimating the above model, the growth rates for pre- and post-reforms periods are obtained as $(\hat{b}_1 * 100)$ and $(\hat{b}_2 * 100)$, respectively.[8]

---

[6] We used the price data available from Singh (2007). To construct all-India price series for individual crops, Singh considered the data on value of production of various crops as available from *National Accounts Statistics* (published by the Central Statistical Organization) which have been divided by the physical quantities of respective crop-production levels as available from *Area and Production of Principal Crops in India* (published by the Ministry of Agriculture, Government of India). To overcome the yearly fluctuation, an average of 3-year prices (for 2001–2002, 2002–2003, and 2003–2004) has been considered.

[7] On the basis of visual inspection of data series for all-crop production, we assumed a 'break' between the crop years 1994–1995 and 1995–1996.

[8] It, however, needs mention that the 'kinked exponential model' suffers from the limitation that it considers the breakpoint in a given time series as exogenously determined. To overcome such a limitation, Quandt (1960) and Andrews (1993) developed a methodology (popularly known as the

To examine the statistical significance of difference in growth rates for the two sub-periods, we further conducted the Wald Chi-square test wherein the null hypothesis has been $H_N : b_1 = b_2$.

Production Performance of Indian Agriculture

All-India Scenario

To begin with, we consider the growth rates of food grains, non-food grains and all-crop production in all-India during pre- and post-reforms periods. It appears from Table 6.1 that, for all these categories, annual growth rates of production decreased significantly during post-reforms period compared to pre-reforms period. While annual growth rate of food grains production in the country was 2.56 % during pre-reforms period (which was higher than annual growth rate of population during 1980s), it declined to a very low level of 1.02 % during post-reforms period (which is much below annual growth rate of population during 1990s and beyond[9]). The growth rate of non-food grains production also bears the same story, declining from 4.14 % during pre-reforms period to 2.15 % during post-reforms period. Thus, the growth rate of all-crop production during post-reforms period (1.50 %) is reduced to less than 50 % of its level during pre-reforms period (3.15 %). It also needs mention that decline in growth rates of food grains, non-food grains and all-crop production were all statistically significant (as revealed by statistical significance of Wald Chi-square statistics), which also indicates presence of negative trend-break in production after adoption economic reforms policies. On the basis of these results, it may be concluded that Indian agriculture

(Footnote 8 continued)
Quandt–Andrews methodology) to endogenously determine the breakpoint in the time series. This methodology has been further extended by Bai and Perron (1998, 2003) to examine presence of multiple breakpoints in the time series. In this study, we have not applied these methodologies to determine the break-year for the following reasons: First, the methodologies for endogenous determination of breakpoints are ideally suited for sufficiently long (historical) data series. Second, these methodologies when applied to different crop-production series do not produce identical result with regard to the presence of the breakyear. This causes difficulty in dividing our study period into two sub-periods (pre- and post-reform). However, our assumption regarding presence of a 'break' in 1995 is broadly in conformity with Chand and Parappurathu (2012) who applied Bai-Perron methodology to a 60-year long data series on value of agricultural GDP and found that 1995–1996 was one of the six breakyears for such a series. Third, in a study such as this, the year of break should not be construed as one that merely fulfills a statistical criterion; rather, it should also correspond to some definitive shift in economic policy regimes. Viewing from this later perspective, our choice of 1995 as the break-year seems justifiable as efforts toward globalization of Indian agriculture formally started from early 1994 the effect of which on agricultural production performance is likely to be reflected after a time-lag.

[9] The annual compound growth rates of population in India during 1991–2001 and 2001–2011 were 1.95 and 1.65 %, respectively.

Table 6.1 Kinked exponential growth rates of food grains, non-food grains and all-crop production in All-India

Crop	Annual growth rate		Diff. in growth rate	Wald Chi-square statistic
	Pre-reform (1980–1981 to 1994–1995)	Post-reform (1995–1996 to 2009–1910)		
Food grains	2.56	1.02	−1.55	10.67*
Non-food grains	4.14	2.15	−1.99	12.00*
All crop	3.15	1.50	−1.65	11.97*

Note * implies significance at 1 % level
Source computed by the authors using data from sources mentioned in Sect. 6.2 of the chapter

suffered a setback with regard to production performance after adoption of economic reforms and globalization policies.

State-Level Performance

Let us now look into the state-level performances with regard to food grains, non-food grains and all-crop production for the two periods—pre-reform and post-reform. The information presented in Table 6.2 show that, compared to pre-reforms period, the decline in growth rate of food grains production during post-reforms period has been widespread in India. Thus, for 12 out of 17 major states considered by us, the annual growth rates of food grains production decelerated during post-reforms period compared to pre-reforms period. What is more, in 11 of these states, the decline in growth rate of food grains production was statistically significant indicating the presence of negative trend-break in food grains production.

As regards growth of food grains production, the five states are found to have improved their position in the post-reforms period. These states are Gujarat, Andhra Pradesh, Orissa, Karnataka, and Jammu and Kashmir. Among these states, the most noticeable improvement in growth of food grains production has taken place in Gujarat (increasing from −0.03 % per annum during pre-reforms period to 3.90 % per annum during post-reforms period). In fact, Gujarat topped the list of states in terms of growth of food grains production during post-reforms period. The other states to perform more or less satisfactorily in terms of growth of food grains production during post-reforms period are Andhra Pradesh (2.53 %), Karnataka (2.28 %), Haryana (1.96 %), and Rajasthan (1.62 %). On the other hand, one of the green-revolution states, namely Punjab (0.36 %), recorded a dismal performance with regard to growth of food grains production during post-reforms period. Further, West Bengal, which topped the list of states in terms of growth of food grains production during pre-reforms period (which was 4.49 %) dropped down to sixth position and suffered a significant decline in growth of food grains

Table 6.2 Kinked exponential growth rates of food grains, non-food grains and all crops production in 17 major states during pre- and post-reforms periods

State	Annual growth rate		Diff. in growth rate	Wald Chi-square statistic
	Pre-reform	Post-reform		
Foodgrains				
Andhra Pradesh	1.33 (13)	2.53 (2)	1.19	1.32
Assam	2.51 (7)	0.69 (7)	−1.83	6.32**
Bihar	3.20 (6)	0.40 (10)	−2.80	4.58**
Gujarat	−0.03 (15)	3.90 (1)	3.93	2.49
Haryana	4.01 (2)	1.96 (4)	−2.05	10.16*
Himachal Pradesh	2.11 (11)	0.30 (13)	−1.81	3.57***
Jammu and Kashmir	0.46 (14)	0.68 (8)	0.22	0.08
Karnataka	1.80 (12)	2.28 (3)	0.48	0.16
Kerala	−2.40 (17)	−3.91 (17)	−1.50	7.14*
Madhya Pradesh	3.87 (4)	−0.74 (15)	−4.61	13.70*
Maharashtra	2.21 (10)	−0.12 (14)	−2.32	3.09***
Orissa	−0.04 (16)	0.52 (9)	0.57	0.11
Punjab	3.88 (3)	0.36 (11)	−3.53	9.93*
Rajasthan	2.40 (8)	1.62 (5)	−0.78	0.15
Tamil Nadu	2.35 (9)	−2.17 (16)	−4.52	8.76*
Uttar Pradesh	3.25 (5)	0.33 (12)	−2.91	34.02*
West Bengal	4.49 (1)	1.32 (6)	−3.17	16.93*
Non-foodgrains				
Andhra Pradesh	4.53 (8)	1.26 (9)	−3.27	10.17*
Assam	1.84 (13)	0.67 (13)	−1.17	17.57*
Bihar	0.17 (15)	2.90 (5)	2.73	2.25
Gujarat	2.29 (11)	4.42 (1)	2.12	0.70
Haryana	5.62 (4)	1.78 (8)	−3.84	10.47*
Himachal Pradesh	7.50 (3)	0.13 (15)	−7.37	7.96*
Jammu and Kashmir	−2.78 (17)	−0.46 (16)	2.33	0.22
Karnataka	5.36 (5)	0.80 (12)	−4.56	28.16*
Kerala	3.85 (10)	1.06 (10)	−2.79	32.13*
Madhya Pradesh	10.73 (1)	2.60 (6)	−8.12	37.27*
Maharashtra	3.94 (9)	4.27 (3)	0.33	0.09
Orissa	−1.73 (16)	−1.35 (17)	0.38	0.06
Punjab	1.99 (12)	0.24 (14)	−1.75	2.96***
Rajasthan	9.22 (2)	3.16 (4)	−6.07	8.87*
Tamil Nadu	4.83 (6)	0.86 (11)	−3.97	11.61*
Uttar Pradesh	1.78 (14)	4.36 (2)	2.58	9.91*
West Bengal	4.63 (7)	2.37 (7)	−2.26	19.68*
All crops				
Andhra Pradesh	2.72 (11)	1.96 (4)	−0.76	0.62
Assam	2.13 (14)	0.68 (13)	−1.45	17.36*
Bihar	2.58 (12)	1.04 (9)	−1.54	1.62
Gujarat	1.57 (15)	4.27 (1)	2.70	1.20
Haryana	4.36 (4)	1.91 (5)	−2.45	17.47*
Himachal Pradesh	2.43 (13)	0.29 (15)	−2.13	5.08**

(continued)

Table 6.2 (continued)

State	Annual growth rate		Diff. in growth rate	Wald Chi-square statistic
	Pre-reform	Post-reform		
Jammu and Kashmir	0.20 (16)	0.76 (10)	0.56	0.48
Karnataka	3.66 (7)	1.51 (7)	−2.14	5.81**
Kerala	3.00 (8)	0.71 (11)	−2.28	25.89*
Madhya Pradesh	5.37 (1)	0.70 (12)	−4.67	16.64*
Maharashtra	2.94 (9)	2.31 (3)	−0.63	0.31
Orissa	−0.56 (17)	0.01 (16)	0.57	0.14
Punjab	3.71 (6)	0.37 (14)	−3.34	11.60*
Rajasthan	4.55 (2)	2.57 (2)	−1.99	1.20
Tamil Nadu	3.75 (5)	−0.19 (17)	−3.95	10.42*
Uttar Pradesh	2.83 (10)	1.51 (8)	−1.31	12.36*
West Bengal	4.52 (3)	1.71 (6)	−2.81	23.67*

Notes *, ** and *** imply significance at 1, 5 and 10 % levels respectively. Figures in parentheses are ranks among 17 major states
Source As in Table 6.1

production during post-reforms period (1.32 %). In four of the remaining states (Kerala, Tamil Nadu, Madhya Pradesh, and Maharashtra), the growth rate of food grains production turned negative during post-reforms period while it was less than 1 % in others. Thus, it can be said that a good majority of the states in India, including some green-revolution states, suffered significant decline in growth of food grains production in the era of economic reforms. The growth rate of food grains production during the post-reforms period was lower than the growth rate of population not only in India but also in majority of the states. This raises concern about maintaining food security in the country in future unless significant breakthroughs are attained with regard to performance of the food grains sector.

A more or less similar picture is obtained as regards growth of non-food grains production during pre- and post-reforms periods. It is observed from Table 6.2 that, compared to pre-reforms period, 11 out of 17 major states in India suffered decline in growth rate of non-food grains production during the period of economic reforms. The states which improved their positions with regard to growth of non-food grains production during post-reforms period are Gujarat, Uttar Pradesh, Maharashtra, Bihar, Jammu and Kashmir, and Orissa. Among these states, once again, Gujarat not only recorded significant improvement in growth rate of non-food grains production during post-reforms period (which is 4.42 % per annum) but also occupied the topmost position in terms of growth rate of non-food grains production during this period. The other states to display satisfactory performance with regard to growth of non-food grains production during post-reforms period are Uttar Pradesh (4.36 %), Maharashtra (4.27 %), Rajasthan (3.16 %), Bihar (2.90 %), Madhya Pradesh (2.60 %), and West Bengal (2.37 %). On the other hand, some of the states that displayed poor performance with regard to growth of non-food grains production during this period are Orissa (−1.35 %), Jammu and

Kashmir (−0.46 %), Himachal Pradesh (0.13 %), Punjab (0.24 %), Assam (0.67 %), and Karnataka (0.80 %).

As regards all-crop (food grains + non-food grains) production, it is found that an overwhelming majority of the states (14 out of 17) suffered decline in growth rate during the post-reforms period compared to pre-reforms period. In 10 out of these 14 states, the decline in growth rate of all-crop production has been statistically significant suggesting the presence of negative trend-break in production. The three states to improve growth rate of all-crop production during this period are Gujarat, Jammu and Kashmir, and Orissa. Among all states, the growth rate of all-crop production during post-reforms period was the highest in Gujarat (4.27 %), which is followed by Rajasthan (2.57 %), Maharashtra (2.31 %), Andhra Pradesh (1.96 %), Haryana (1.91 %), West Bengal (1.71 %), Karnataka (1.51 %), Uttar Pradesh (1.51 %), and Bihar (1.04 %). In remaining seven states (Jammu and Kashmir, Kerala, Madhya Pradesh, Assam, Punjab, Himachal Pradesh, and Orissa) the growth rate of all-crop production was less than 1 % while it was negative in Tamil Nadu. It clearly appears that while a vast majority of the states in India (including the green-revolution state like Punjab), suffered decline in growth of crop production during the period of economic reforms, a handful of states displayed some improvement in this regard. Of course, the most notable turn around with regard to crop production (both food grains and non-food grains) has happened in the state of Gujarat, which is currently being cited as a model for emulation by other states.

Crop-Level Performance in All-India

In order to have a more disaggregated view of agricultural performance, we have computed growth rates of production of major crops/crop groups for our study periods. Table 6.3 presents such information for all-India. It clearly emerges that the growth rates of production declined significantly for majority of the crops (both food grains and non- food grains) during the period of economic reforms compared to pre-reforms period. The important crops within the group of food grains to suffer decline in growth rate of production during post-reforms period are rice, wheat and pulses. For all these crops (except pulses), the decline in growth rate of production during the post-reforms period has been statistically significant. In other words, a negative trend-break in respect of production growth of these crops is observed during the period of economic reforms. Among the non-food grains and commercial crops, decline in growth rate of production during post-reforms period is noticeable in respect of groundnut, rapeseed and mustard, soybean, oilseeds (total), plantation crops (comprising tea, coffee and rubber), condiments and spices, banana, potatoes, and sugarcane. For all these crops (except potatoes) decline in growth rate of production has been statistically significant indicating the presence of negative trend-break in production. Three crops which experienced improvement in growth rate of production during post-reforms period are coarse cereals,

Table 6.3 Kinked exponential growth rates of production of some selected crops in all-India

Crop	Annual growth rate		Diff. in growth rate	Wald Chi-square statistic
	Pre-reform	Post-reform		
Rice	3.20	0.93	−2.27	15.06*
Wheat	3.75	1.35	−2.41	31.87*
Coarse cereals	0.05	1.14	1.09	1.66
Total pulses	0.86	0.43	−0.43	0.35
Groundnut	2.23	−1.84	−4.07	5.67**
Rapeseed and mustard	6.55	0.98	−5.58	14.29*
Soybean	18.14	4.66	−13.49	84.87*
Total oilseeds	5.28	0.83	−4.45	17.04*
Cotton	2.71	4.94	2.22	3.09***
Total fibers	1.70	3.90	2.20	5.89**
Plantation crops	3.78	2.55	−1.23	15.10*
Condiments and spices	4.29	3.03	−1.26	4.24**
Banana	6.88	4.03	−2.85	3.82***
Potatoes	4.59	3.36	−1.23	1.99
Sugarcane	3.71	0.77	−2.94	12.93*

Note *, ** and *** imply significance at 1, 5 and 10 % levels respectively
Source As in Table 6.1

cotton and fibers (total); the last two of these crops displaying statistically significant improvement in production growth rate during post-reforms period compared to pre-reform period. However, on the whole, it appears that majority of the crops in India suffered decline in growth rate of production during the period of economic reforms compared to the same during pre-reforms period.

Crop-Level Performance in States

We have also computed growth rates of production during pre- and post-reforms periods for different crops in 17 major states considered for this study. Such information has been presented in Table 6.4.

Consistent with the pattern observed for all-India, the growth rates of production of two most important food crops, namely rice and wheat, declined in a large number of states during the period of economic reforms compared to pre-reforms period. As regards rice, such a decline is visible in 14 out of 17 major states. Further, in half of these states, the growth rate of production for rice has been statistically significantly lower during post-reforms period compared to pre-reforms period which indicates the presence of negative trend-break in production. These states are Madhya Pradesh, Tamil Nadu, Uttar Pradesh, West Bengal, Assam, Punjab, and Kerala (in descending order of rate of reduction in growth rate of rice production). It is also to be noted that all these states, except Kerala, belong

Table 6.4 Annual growth rates of production of Individual crops in 17 major states of India during pre- and post-reforms periods

State	Rice				Wheat				Coarse Cereals			
	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic
	Pre-reform	Post-reform			Pre-reform	Post-reform			Pre-reform	Post-reform		
Andhra Pradesh	1.95	1.43	−0.52	0.18	−1.45	1.06	2.51	1.11	−2.98	5.66	8.65	48.55*
Assam	2.70	0.76	−1.94	6.35**	−1.51	−3.01	−1.50	1.73	0.82	0.60	−0.22	0.14
Bihar	3.09	0.26	−2.84	2.19	4.43	−0.50	−4.93	35.77*	2.38	2.64	0.25	0.03
Gujarat	1.26	3.69	2.43	1.30	−1.26	7.88	9.14	8.33*	−0.22	0.80	1.01	0.16
Haryana	4.04	3.81	−0.24	0.08	5.17	1.99	−3.18	39.28*	−0.03	4.13	4.15	2.98***
Himachal Pradesh	0.72	0.60	−0.12	0.01	2.72	0.09	−2.63	1.79	2.32	−0.12	−2.44	4.64**
Jammu and Kashmir	−0.92	0.02	0.94	0.93	4.32	2.19	−2.13	1.38	1.04	0.41	−0.63	0.28
Karnataka	2.84	1.64	−1.20	0.80	−0.82	3.35	4.16	2.47	1.32	2.06	0.74	0.29
Kerala	−2.49	−3.75	−1.25	4.67**					4.70	−10.85	−15.55	26.10*
Madhya Pradesh	4.41	−1.09	−5.50	7.86**	7.24	−6.89	−14.13	6.04**	−0.51	−1.40	−0.89	0.42
Maharashtra	0.93	0.18	−0.75	0.43	2.24	−2.58	−4.82	0.51	2.08	−2.97	−5.05	7.33*
Orissa	2.49	0.91	−1.59	0.86	−10.62	−13.34	−2.73	0.92	−11.69	0.12	11.81	23.45*
Punjab	4.81	2.51	−2.30	12.45*	5.56	−6.86	−12.42	3.36***	−3.60	1.66	5.26	31.40*
Rajasthan	0.73	2.14	1.41	0.22	4.86	1.80	−3.06	5.89**	0.67	4.31	3.64	1.50
Tamil Nadu	2.99	−2.46	−5.45	11.13*					−3.08	1.02	4.11	4.47**
Uttar Pradesh	5.09	0.32	−4.77	30.09*	3.75	1.07	−2.68	37.70*	2.01	−2.12	−4.13	27.95*
West Bengal	5.00	1.28	−3.72	19.29*	1.98	1.63	−0.35	0.06	−0.46	4.54	5.01	2.96***

(continued)

Table 6.4 (continued)

State	Pulses (total)				Oilseeds (total)				Fibers (total)			
	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic
	Pre-reform	Post-reform			Pre-reform	Post-reform			Pre-reform	Post-reform		
Andhra Pradesh	2.54	5.46	2.93	6.28**	5.16	−2.47	−7.63	13.88*	2.57	5.21	2.64	2.02
Assam	1.07	0.51	−0.56	0.52	1.58	−1.38	−2.97	9.42*	−1.23	−2.39	−1.16	0.92
Bihar	−1.06	3.03	4.09	6.87*	1.73	0.91	−0.82	0.26	0.95	0.59	−0.36	0.05
Gujarat	0.86	0.96	0.10	0.00	2.90	3.20	0.30	0.01	−0.01	11.76	11.77	9.79*
Haryana	−0.82	−10.18	−9.36	5.76**	12.08	−0.50	−12.58	34.16*	4.77	1.46	−3.31	0.09
Himachal Pradesh	−2.01	5.88	7.89	6.12**	4.69	−3.11	−7.80	9.49*				
Jammu and Kashmir	−5.49	−1.70	3.79	9.21*	−2.71	−0.38	2.32	0.21				
Karnataka	0.85	3.97	3.12	5.30**	5.46	−1.55	−7.01	38.32*	1.31	−2.11	−3.41	1.82
Kerala	−0.55	−8.91	−8.36	7.87*	5.78	−1.08	−6.86	44.74*	7.72	−18.39	−26.11	115.05*
Madhya Pradesh	3.08	1.31	−1.77	2.19	12.37	1.89	−10.48	59.47*	−0.09	8.47	8.57	15.35*
Maharashtra	4.30	1.56	−2.73	2.55	5.18	3.59	−1.59	0.85	2.89	6.54	3.65	2.61
Orissa	−8.51	−1.89	6.62	4.06**	−6.92	−5.15	1.77	0.29	−4.92	−4.17	0.75	0.13
Punjab	−6.10	−10.03	−3.93	9.32*	2.66	−8.25	−10.91	23.87*	1.35	0.36	−0.98	0.09
Rajasthan	0.83	−1.64	−2.47	0.45	11.55	2.58	−8.97	15.68*	5.45	−2.44	−7.88	5.32**
Tamil Nadu	3.64	−4.48	−8.11	19.17*	5.72	−2.27	−7.99	48.44*	1.34	−3.47	−4.80	2.12
Uttar Pradesh	−2.22	−4.31	−2.10	6.61**	3.50	0.51	−2.99	1.45	−13.03	−12.79	0.24	0.02
West Bengal	−2.76	−0.35	2.41	3.77***	6.39	2.05	−4.34	7.87*	3.06	2.45	−0.61	0.26

(continued)

Table 6.4 (continued)

State	Plantation Crops (total)				Potatoes				Sugarcane			
	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic	Annual gr. rate		Diff. in growth rate	Wald Chi-square statistic
	Pre-reform	Post-reform			Pre-reform	Post-reform			Pre-reform	Post-reform		
Andhra Pradesh									2.89	1.13	−1.76	1.47
Assam	2.29	1.14	−1.15	13.10*	5.94	0.16	−5.78	21.22*	−3.58	−3.02	0.57	0.28
Bihar					0.90	4.28	3.38	1.58	2.59	−1.48	−4.06	6.33**
Gujarat					5.20	8.11	2.91	2.39	2.83	2.32	−0.52	0.49
Haryana					0.11	7.88	7.77	24.70*	3.66	−0.76	−4.42	9.28*
Himachal Pradesh					9.42	−0.37	−9.79	9.44*	2.90	1.05	−1.84	0.20
Jammu and Kashmir									−10.02	−26.99	−16.97	10.86*
Karnataka	3.97	2.64	−1.33	0.88	9.81	−0.99	−10.80	21.48*	6.50	−1.63	−8.14	17.90*
Kerala	6.86	3.58	−3.28	27.46*					0.50	−3.88	−4.38	3.68***
Madhya Pradesh					4.86	3.83	−1.03	0.30	−0.27	4.48	4.74	9.77*
Maharashtra					−0.03	6.10	6.13	5.22*	2.53	2.95	0.42	0.03
Orissa					0.91	0.15	−0.76	0.06	−7.58	−4.61	2.97	1.05
Punjab					2.72	7.35	4.63	3.39***	3.43	−2.73	−6.16	10.96*
Rajasthan									−1.20	−8.06	−6.86	7.56*
Tamil Nadu	3.43	2.34	−1.10	7.65*					4.40	0.16	−4.23	6.82*
Uttar Pradesh					3.64	3.92	0.28	0.07	0.82	5.22	4.40	8.62*
West Bengal	1.24	2.77	1.53	17.63*	7.74	1.71	−6.03	28.32*	1.31	1.81	0.49	0.03

Notes *, ** and *** imply significance at 1, 5 and 10 % levels respectively. Blank spaces imply that the area under the crop is insignificant in the state

Source As in Table 6.1

to the group of major rice producing states of the country (CMIE 2010, p. 54). During the post-reforms period, annual growth rate of rice production was highest in Haryana (3.81 %). The other states to display impressive performance in rice production during this period are Gujarat (3.69 %), Punjab (2.51 %), and Rajasthan (2.14 %). On the other hand, while the growth rate of rice production in Tamil Nadu turned negative (−2.46 %) during post-reforms period, in West Bengal it is reduced to a very low level of 1.28 % (which was 5 % during pre-reforms period). Even in Uttar Pradesh, the highest contributor in the rice bowl of the country, the growth rate of rice production during post-reforms period is reduced to a meager 0.32 %, which was highest among all states at 5.09 % during pre-reforms period.

The picture obtained for wheat is no different from that is observed for rice. Thus, in 12 out of 17 major states the growth rate of wheat production during post-reforms period was lower than the same during pre-reforms period. Of these states, six states displayed clear signs of negative trend-break in production. These states, in descending order of rate of decline in growth rate of wheat production, are Madhya Pradesh, Punjab, Bihar, Haryana, Rajasthan, and Uttar Pradesh. All these states belong to the group of major wheat producing states of the country (Ibid., p. 96). Another important observation in respect of growth rate of wheat production is that it turned negative during post-reforms period in six states that include Punjab (others are Assam, Bihar, Madhya Pradesh, Maharashtra, and Orissa). On the other hand, Gujarat recorded the highest growth rate in wheat production during post-reforms period (7.88 %) which was negative during pre-reforms period (−1.26 %). As a consequence, it became the only state that enjoyed positive trend-break in wheat production in the era of economic reforms. Apart from Gujarat, other states to display relatively good performance with regard to growth of wheat production during post-reforms period are Karnataka (3.35 %), Haryana (1.99 %), Jammu and Kashmir (2.19 %), and Rajasthan (1.80 %). In Uttar Pradesh, which is the largest producer of wheat in the country, the growth rate of wheat production dropped down to 1.07 % during post-reforms period from a rather high level of 3.75 % during pre-reforms period.

In respect of production of coarse cereals, we have a mixed picture. While six states enjoyed positive trend-break in production of coarse cereals during post-reforms period compared to pre-reforms period (these are Andhra Pradesh, Haryana, Punjab, Orissa, Tamil Nadu, and West Bengal), four states suffered from negative trend-break in production (which are Kerala, Uttar Pradesh, Maharashtra, and Himachal Pradesh). It also needs mention that, among all states, most impressive performance (more than 4 % per annum growth rate) with regard to production of coarse cereals during post-reforms period has been displayed by Andhra Pradesh, Haryana, Rajasthan, and West Bengal.

The growth rate of total pulses production improved significantly in seven states during the post-reforms period compared to pre-reforms period. These states, in descending order of incremental growth rate, are Himachal Pradesh, Orissa, Bihar, Jammu and Kashmir, Karnataka, Andhra Pradesh, and West Bengal. On the other hand, the growth rate of pulses production decreased significantly in five states during post-reforms period, which are Tamil Nadu, Punjab, Haryana,

Kerala, and Uttar Pradesh. Among all states, quite impressive growth rate of pulses production has been recorded in Himachal Pradesh (5.88 %), Andhra Pradesh (5.46 %), Karnataka (3.97 %), and Bihar (3.03 %). However, the growth rate of pulses production in major pulses producing states like Maharashtra (1.56 %) and Madhya Pradesh (1.31 %) was much low during post-reforms period compared to pre-reforms period while in Uttar Pradesh it turned negative (−4.31 %).

Apart from rice and wheat, another glaring example of agricultural production suffering during the period of economic reforms is observed in respect of oilseeds (total). As shown in Table 6.4, in 14 out of 17 major states, the growth rate of oilseeds production during post-reforms period was lower than the same during pre-reforms period. Again, in 11 out of these 14 states, the decline in growth rate of production during post-reforms period was statistically significant indicating clear sign of negative trend-break in oilseeds production in vast majority of the states. Gujarat is the only exception among major oilseeds producing states (others are Rajasthan, Uttar Pradesh, Karnataka, Andhra Pradesh, Tamil Nadu, Madhya Pradesh etc.) to have recorded some improvement in growth rate of oilseeds production during post-reforms period (3.20 %) compared to pre-reform period (2.90 %).

More than one-half of the states considered by us suffered decline in fibers (comprising cotton, jute, mesta and sannhemp) production during post-reforms period compared to pre-reforms period. This fall in growth rate of fibers production has been statistically significant in Kerala and Rajasthan. On the other hand, Gujarat and Madhya Pradesh attained statistically significant improvement (positive trend-break) in growth of fibers production during post-reforms period compared to pre-reforms period. This has happened due to significant improvement in cotton production in these two states. The other states to attain some improvement in growth rate of fibers production (especially cotton) during post-reforms period are Andhra Pradesh and Maharashtra.

The plantation crops (tea, coffee and rubber) suffered decline in growth rate of production during post-reform period in all states except West Bengal. Significant decline in this respect (especially coffee and rubber) is noticeable in Kerala while in Tamil Nadu and Assam it is with respect of tea. On the other hand, West Bengal recorded significant improvement in production of plantation crops (primarily tea) during post-reforms period compared to pre-reforms period.

Production growth of a commercial/cash crop like potatoes suffered decline during post-reforms period in six states of which four states witnessed negative trend-break in production. These states, in descending order of rate of decline in growth rate, are Karnataka, Himachal Pradesh, West Bengal, and Assam. On the other hand, six states recorded higher growth of potatoes production during post-reforms period of which three displayed positive trend-break (these are Haryana, Maharashtra, and Punjab). It also needs mention that while West Bengal, the second largest producer of potatoes in the country, suffered from a very high reduction in growth rate of potatoes production during post-reforms period compared to pre-reforms period (declining from 7.74 to 1.71 % per annum), Uttar Pradesh (the largest producer of potatoes in the country) failed to make any significant improvement in the same during post-reforms period. This largely

explains overall (all-India level) fall in growth rate of potatoes production in the country, as observed earlier.

As regards sugarcane, another important cash crop, compared to pre-reforms period, the growth rate of production declined during post-reforms period in 11 out of 17 major states. Among these states, the decline in growth rate has been statistically significant in eight states which are Jammu and Kashmir, Karnataka, Rajasthan, Punjab, Haryana, Kerala, Tamil Nadu, and Bihar. The two states to improve growth rate of sugarcane production significantly during post-reforms period are Uttar Pradesh and Madhya Pradesh. It also appears that although, among major sugarcane producing states of the country, both Uttar Pradesh and Maharashtra recorded some improvement in growth rate of sugarcane production during post-reforms period, this has been significant only in the case of the former.

Some Constraints Faced by Indian Agriculture

The above discussion clearly brought out a story of dismal production performance of agriculture in India in the era of economic reforms. This has been revealed through declines in growth rates of production of a wide array of crops (both food and non-food), both in all-India and majority of the states. It would, therefore, be in the fitness of things to look into some major constraints faced by Indian agriculture which are likely to explain its poor performance during the post-reforms period. Some of those constraints are the following[10]:

Declining Land Availability

There is clear sign of amount of land available for cultivation declining, thanks to increasing industrialization and urbanization in the country. In fact, a process of agricultural land getting shrunk, albeit slowly, has been in operation right from the start of 1980s and continued unabated in subsequent decades in all-India and majority of the states. This point becomes clear when we look into the volume of net sown area and its growth rate, presented in Table 6.5. It is found that the net sown area (in thousand hectares) in all-India declined from 1,33,627 during 1980–1983 to 1,29,859 during 1990–1993, and further to 1,28,510 during 2005–2008. It is also found that net sown area in all-India declined at an annual rate of 0.29 % during pre-reforms period (1980–1983 to 1990–1993) and 0.08 %

[10] Our discussion here is too broad and synoptic, and needs to be expanded substantially. We also admit that the list of constraints provided by us is not exhaustive. Actually, the purpose of this section has been to initiate discussion on this vital issue. In any case, more detailed research should be undertaken to complete our understanding about the obstacles faced by the agriculture sector in different states/regions of India in the ongoing phase of globalization and liberalization.

during post-reforms period (1990–1993 to 2005–2008). When we look into the growth rate of net sown area in the states, it is found that it was negative in 13 out of 17 major states during post-reforms period (the exceptions being Gujarat, Rajasthan, Punjab, and Haryana). All these clearly indicate that no future scope for expansion of areas under cultivation exists so as to increase agricultural production in the country through this route.

Table 6.5 further shows that, despite non-availability of more lands for crop cultivation, cropping intensity[11] in India have not expanded to any great extent over the years. Cropping intensity, which was 124 in all-India during 1980–1983, expanded to 130 only during 1990–1993 and 138 during 2005–2008. Moreover, cropping intensity remained at abysmally low levels in majority of the states even in 2005–2008. In more than one-half of the states, cropping intensity hardly exceeded 150 today. These states are Andhra Pradesh, Assam, Bihar, Gujarat, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Rajasthan, and Tamil Nadu. When we consider the growth rate of cropping intensity, it appears that, at the all-India level, it remained almost stagnant all though our study periods—pre- and post-reforms. Further, at the state level, decline in growth rate of cropping intensity is visible during post-reforms period compared to pre-reforms period in 10 out of 17 major states. These states, in descending order of decline in growth rate of cropping intensity are Assam, Punjab, West Bengal, Tamil Nadu, Orissa, Jammu and Kashmir, Kerala, Karnataka, Andhra Pradesh, and Uttar Pradesh.

Low Growth of Inputs Used

Perhaps the most important constraint faced by Indian agriculture in the era of economic reforms is low progress in the matter of utilization of various inputs, which is likely to explain large part of observed decay in growth of agricultural production in this era. To understand the levels of inputs utilization, we have presented data on some indicators in Table 6.6, which are: (a) number of tractors per 10,000 ha, (b) Pumpsets per 1,000 ha, (c) fertilizers consumption (Kg./ha), and (d) percentage of gross cropped area irrigated. It is found that, in absolute terms, some increase in utilization of these inputs has taken place both in all-India and the individual states. At the all-India level, the number of tractors (No./0000 ha) increased from 37 during 1980–1983 to 86 during 1990–1993 and 167 during 2005–2008. Similarly, pumpsets (No./000 ha) increased from 49 to 79 and 111. Fertilizers consumption (Kg./ha) for the three periods are found to be 44, 91 and 153, respectively. As regards percentage of gross cropped area irrigated, the figures are 29, 36, and 44 for the three periods, respectively. The increases in utilization of all these inputs, in absolute terms, are also visible in the states (an

[11] Cropping intensity = (Gross cropped area/Net sown area) * 100.

Table 6.5 Net sown area and cropping intensity in India and major states

State	Net sown area (000 ha)			Cropping intensity			Annual growth rates of				
							Net sown area		Cropping intensity		
	1980–1983	1990–1993	2005–1908	1980–1983	1990–1993	2005–2008	Pre-reform	Post-reform	Pre-reform	Post-reform	Change
Andhra Pradesh	10608	10213	9525	115	120	126	−0.38	−0.54	0.43	0.38	−0.05
Assam	2598	2567	2461	128	142	139	−0.12	−0.32	1.04	−0.16	−1.21
Bihar	7665	7216	7198	133	133	135	−0.60	−0.02	0.00	0.11	0.11
Gujarat	8076	7510	8120	113	114	122	−0.72	0.60	0.09	0.52	0.43
Haryana	3311	3248	3309	153	164	181	−0.19	0.14	0.70	0.76	0.06
Himachal Pradesh	547	529	480	166	170	176	−0.34	−0.75	0.24	0.27	0.03
Jammu & Kashmir	658	666	631	137	146	152	0.11	−0.41	0.64	0.31	−0.33
Karnataka	9532	10089	9760	108	115	124	0.57	−0.25	0.63	0.58	−0.05
Kerala	1944	1907	1742	132	135	137	−0.19	−0.69	0.22	0.11	−0.11
Madhya Pradesh	17930	18640	18166	116	121	132	0.39	−0.20	0.42	0.67	0.25
Maharashtra	17685	16838	16466	109	117	129	−0.49	−0.17	0.71	0.75	0.04
Orissa	5555	5208	3944	141	152	159	−0.64	−2.12	0.75	0.35	−0.41
Punjab	3824	3726	3813	158	180	188	−0.26	0.18	1.31	0.34	−0.98
Rajasthan	14012	15335	16140	117	118	129	0.91	0.39	0.09	0.69	0.60
Tamil Nadu	5298	5210	4623	119	121	115	−0.17	−0.92	0.17	−0.39	−0.56
Uttar Pradesh	18114	16440	16072	143	148	154	−0.97	−0.17	0.34	0.31	−0.04
West Bengal	5353	4933	4565	132	160	182	−0.81	−0.60	1.94	1.00	−0.95
All India	133627	129859	128510	124	130	138	−0.29	−0.08	0.47	0.46	−0.01

Notes (i) Pre-reform period here represents 1980–1983 to 1990–1993 and Post-reform represents 1990–1993 to 2005–2008; and (ii) Growth rates are point-to-point growth rates

Source Computed by the authors using data available from Bhalla and Singh (2012)

Table 6.6 Levels of utilization of various inputs in agriculture—all India and 17 major states

State	Tractors (No./0000 ha)			Pumpsets (No./000 ha)			Fertilizers consumption Kg./ha)			% of GCA irrigated		
	1980–1983	1990–1993	2005–2008	1980–1983	1990–1993	2005–2008	1980–1983	1990–1993	2005–2008	1980–1983	1990–1993	2005–2008
Andhra Pradesh	19	52	85	62	101	148	56	137	243	36	40	46
Assam	1	3	5	1	2	2	5	16	75	17	15	4
Bihar	18	19	130	47	89	117	24	77	166	34	43	49
Gujarat	29	70	150	59	67	92	41	75	147	23	29	41
Haryana	170	444	549	71	143	155	71	175	324	62	76	85
Himachal Pradesh	16	45	130	3	4	20	33	62	90	17	18	19
Jammu and Kashmir	11	18	70	1	5	28	36	65	116	40	41	41
Karnataka	20	37	60	30	58	79	37	82	146	13	23	29
Kerala	6	9	10	45	88	196	49	111	98	13	12	16
Madhya Pradesh	13	24	130	22	47	107	14	50	80	12	21	30
Maharashtra	12	50	60	33	66	62	27	69	125	13	15	19
Orissa	2	4	28	3	6	19	14	33	74	22	26	35
Punjab	254	508	704	158	170	170	209	318	404	87	95	98
Rajasthan	35	90	184	28	54	88	10	30	55	21	27	36
Tamil Nadu	26	52	102	211	212	210	80	136	214	49	48	56
Uttar Pradesh	82	201	397	64	132	191	75	129	219	47	62	74
West Bengal	3	12	34	37	54	119	49	136	250	25	54	56
All India	37	86	167	49	79	111	44	91	153	29	36	44

Source as in Table 6.5

exception being Assam where the percentage of gross cropped area irrigated declined over the years).

However, a different story unfolds when we compare growth rates of these inputs for the two periods—pre- and post-reforms.[12] As shown in Table 6.7, for all inputs, at the all-India level, the growth rates of various inputs used in agriculture declined during the period of economic reforms compared to pre-reforms period. Thus, annual growth rate of gross irrigated area, which was 2.38 % during pre-reforms period, reduced to 1.94 % during post-reforms period; growth rate of fertilizers consumption, reduced from 7.74 % to 4.47 %; pump sets from 5.09 % to 3.04 %; and tractors from 9 % to 5.64 %.

The state level scenarios with regard to growth of various inputs are broadly similar to the one observed for all-India. In 13 out of 17 states considered, the growth rate of gross irrigated area declined during post-reforms period compared to pre-reforms period (the exceptions are Gujarat, Kerala, Maharashtra, and Tamil Nadu). As regards growth of fertilizers consumption, decline in growth rate during post-reforms period over the previous period is visible in all but one state (the exception here is Gujarat). Similarly, growth rate of pumpsets declined during post-reforms period in 13 states (Gujarat, Himachal Pradesh, Orissa and West Bengal are the exceptions). For tractors, 13 states are found to have suffered decline in growth rate during post-reforms period (exceptions are Bihar, Jammu and Kashmir, Madhya Pradesh, and Orissa).

It clearly emerges that a vast majority of the states in India (including green-revolution states like Punjab and Haryana) suffered declines in growth rates of various agricultural inputs during the period of economic reforms which possibly explain such a widespread decline in growth rates of agricultural production in the era of economic reforms, as observed earlier. However, it also needs mention that, in all these respects, a notable exception during post-reforms period has been Gujarat, and it is no wonder that its breakthrough on the agricultural front has drawn wide research attention in recent years.[13]

---

[12] For calculation of growth rate of an input (X), the following formula has been used:

$$X_j = X_i[1 + r]^n$$

$$\Rightarrow r = \left[X_j/X_i\right]^{1/n} - 1$$

where 'r' is the rate of growth and 'n' is the time length between two periods (i and j). Using this formula, 'r' when multiplied by 100 provides growth rate in percentage terms.

[13] Several scholars have attempted to explain Gujarat's success on the agricultural front in the era of economic reforms. They observed public investment in infrastructure like irrigation, power, roads, watersheds, check dams, technology like BT cotton, and diversification in agriculture played crucial roles in raising agricultural growth rate in Gujarat since mid-1990s (see Gulati 2009; Shah et al. 2009).

Table 6.7 Growth rates of various inputs used in agriculture—all India and 17 major states

State	Annual growth rates of											
	Tractors			Pumpsets			Fertilizers consumption			Gross irrigated area		
	Pre-reform	Post-reform	Change	Pre-reform	Post-reform	Change	Pre-reform	Post-reform	Change	Pre-reform	Post-reform	Change
Andhra Pradesh	10.64	3.68	−6.96	5.05	2.82	−2.23	9.41	4.34	−5.07	1.11	0.92	−0.19
Assam	12.64	3.50	−9.14	8.17	−0.31	−8.48	13.37	12.07	−1.30	−0.33	−10.11	−9.77
Bihar	−0.06	16.05	16.12	5.95	2.22	−3.73	11.69	6.19	−5.50	1.76	1.11	−0.65
Gujarat	8.52	7.24	−1.28	0.63	3.63	2.99	5.55	6.50	0.95	1.69	3.86	2.17
Haryana	10.63	2.57	−8.06	7.79	1.53	−6.26	9.99	5.80	−4.19	2.57	1.78	−0.79
Himachal Pradesh	10.78	7.98	−2.81	2.82	12.63	9.81	6.40	2.41	−3.99	0.47	−0.07	−0.54
Jammu and Kashmir	5.84	10.90	5.06	18.34	14.05	−4.29	6.88	4.45	−2.43	1.00	−0.10	−1.10
Karnataka	7.62	4.13	−3.50	8.10	2.74	−5.36	9.59	4.88	−4.71	7.14	2.13	−5.01
Kerala	4.17	0.23	−3.95	6.97	5.74	−1.24	8.56	−1.53	−10.09	−0.76	1.64	2.41
Madhya Pradesh	7.19	14.42	7.23	8.76	7.04	−1.73	14.50	4.17	−10.33	6.62	3.27	−3.35
Maharashtra	15.59	2.00	−13.59	7.41	0.10	−7.31	10.08	5.29	−4.79	1.66	2.43	0.76
Orissa	7.29	14.08	6.79	7.29	7.33	0.04	9.07	4.52	−4.55	1.79	0.50	−1.30
Punjab	8.30	3.07	−5.23	1.79	0.51	−1.28	5.38	2.38	−3.00	1.94	0.75	−1.19
Rajasthan	11.00	6.80	−4.19	7.85	4.95	−2.89	12.72	5.91	−6.81	3.56	3.35	−0.22
Tamil Nadu	7.18	3.95	−3.23	0.05	−1.37	−1.42	5.45	2.20	−3.25	−0.21	−0.13	0.08
Uttar Pradesh	8.70	5.51	−3.18	6.84	3.02	−3.82	4.91	4.29	−0.62	2.17	1.50	−0.66
West Bengal	16.15	8.77	−7.38	5.01	6.69	1.68	11.98	5.21	−6.77	9.21	0.68	−8.53
All India	9.00	5.64	−3.37	5.09	3.04	−2.05	7.74	4.47	−3.27	2.38	1.94	−0.44

Source as in Table 6.5

Falling Investment

Investment is considered to be the main driver of growth in all sectors of economy, including agriculture. Past studies have revealed that gross capital formation (GCF) in agriculture and allied sector (which is a measure of investment) moved upwards right from mid-1960s, which picked up further in mid-1970s. At that time, the share of public sector in GCF in agriculture and allied sector was quite high. Table 6.8 shows that it was as high as 48.39 % in 1980–1981. However, the share of public sector in GCF dwindled over the years, especially after the adoption of economic

Table 6.8 Gross capital formation (GCF) in agriculture and allied sector in India

Year	GCF in agrl. and allied sector (at current prices) (Rs. crore)			Share in GCF of agrl. and allied sector (%)		Share of agrl. and allied sector in total (%)		
	Public sector	Private sector	Total	Public sector	Private sector	Public sector	Private sector	Total
1980–1981	1892	2018	3910	48.39	51.61	15.6	12.9	14.1
1981–1982	2041	2263	4304	47.42	52.58	12.0	8.3	9.7
1982–1983	2263	2636	4899	46.19	53.81	11.2	10.9	11.1
1983–1984	2466	2902	5368	45.94	54.06	11.6	13.0	12.3
1984–1985	2678	3413	6091	43.97	56.03	10.5	10.5	10.5
1985–1986	2818	3768	6586	42.79	57.21	9.4	8.8	9.1
1986–1987	2895	4033	6928	41.79	58.21	8.3	9.4	8.9
1987–1988	3304	4642	7946	41.58	58.42	9.8	12.6	11.2
1988–1989	3442	5577	9019	38.16	61.84	8.6	9.2	8.9
1989–1990	3354	5833	9187	36.51	63.49	7.2	8.8	8.2
1990–1991	3628	8596	12224	29.68	70.32	6.8	11.7	9.7
1991–1992	3653	9000	12653	28.87	71.13	6.3	9.7	8.4
1992–1993	4175	10804	14979	27.87	72.13	6.5	10.4	8.9
1993–1994	4918	10331	15249	32.25	67.75	6.9	9.4	8.4
1994–1995	6002	12381	18383	32.65	67.35	6.8	7.7	7.4
1995–1996	6762	14605	21367	31.65	68.35	7.4	5.9	6.3
1996–1997	7296	17119	24415	29.88	70.12	7.6	7.9	7.8
1997–1998	6921	19087	26008	26.61	73.39	6.9	8.2	7.8
1998–1999	7583	19131	26714	28.39	71.61	6.6	8.7	8.0
1999–2000	8670	41481	50151	17.29	82.71	6.0	11.9	10.2
2000–2001	8176	38558	46734	17.49	82.51	5.7	11.1	9.5
2001–2002	10353	51285	61638	16.80	83.20	6.6	13.5	11.5
2002–2003	9563	52319	61882	15.45	84.55	6.4	11.5	10.2
2003–2004	12218	49249	61467	19.88	80.12	7.0	9.1	8.6
2004–2005	16187	59909	76096	21.27	78.73	6.7	7.8	7.5
2005–2006	20739	69213	89952	23.06	76.94	7.1	7.4	7.3
2006–2007	25606	73993	99599	25.71	74.29	7.2	6.5	6.7
2007–2008	27638	94836	122474	22.57	77.43	6.3	6.8	6.6
2008–2009	29378	132835	162213	18.11	81.89	5.6	9.7	8.5
2009–2010	34521	188842	223363	15.46	84.54	5.7	9.1	8.2

Source Ministry of Agriculture, Government of India (downloaded from www.indiastat.com)

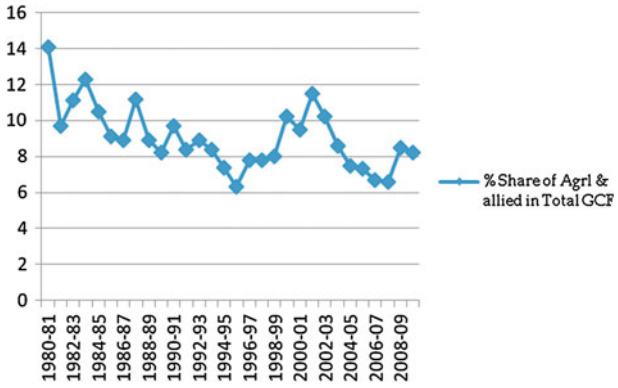


Fig. 6.1 Percentage share of agrl. and allied in total GCF

reforms policies in 1991. This produced a dampening effect on the private sector capital formation as well. As a consequence, GCF in agriculture and allied sector as a whole (public + private) suffered a decline in relative terms. This becomes clear from the fact that while the share of agriculture and allied sector in total GCF was as high as 14.1 % in 1980–1981, it reduced to 9.7 % in 1990–1991 and 6.6 % in 2007–2008 after which it recovered a bit to reach 8.2 % in 2009–2010. Another point that becomes clear from the series representing the share of agriculture and allied sector in total GCF is that ever since India's adoption of economic reforms policies, it followed a declining trend excepting a few years around 2,000 (see Fig. 6.1 below). Several scholars have held this falling tendency of agriculture investment (particularly by the public sector) responsible for falling agricultural growth in India during post-reforms period.

Unfavorable Terms of Trade

It has been observed that the terms of trade for agriculture, which is measured as a ratio of the implicit price index of GDP accruing from agriculture to that of GDP accruing from non-agriculture, witnessed a gradual increase during 1980s and early years of 1990s. However, the terms of trade for agriculture deteriorated sharply during late 1990s, which continued till about 2004–2005 (Chand and Parappurathu 2012). This not only depressed the farmers' incomes but also increased farm debt considerably. Notwithstanding the terms of trade displaying some improvement in recent years, it is considered as one of the factors responsible for deceleration in agricultural growth since mid-1990s. This is more so when the costs of crop-cultivation have been increasing with reduction of subsidies given to various inputs by the government as also rising labor costs (Raghavan 2008).

Credit Constraint

There has been gradual reduction in the share of credit for agriculture to total credit advanced by the commercial banks, which happened with implementation of financial liberalization policies. The Expert Group on Financial Reforms (Government of India 2008) noted that only 27 % of farmers in the country have access to institutional credit. Again, among all farmers, access to institutional credit by the small and marginal farmers remained low although they constituted more than three-fourths of all farmers in the country. The decline in credit-deposit ratio in rural areas during post-reforms period is also well-known.[14] Above all, a significant inter-regional inequality in the matter of distribution of agricultural credit persists.

Research, Extension, and Technology Fatigue

We observed earlier that the production growth for many crops declined in the post-reforms period. This happened due to stagnation in yield growth as the possibility of area growth is virtually absent. Technology plays an important role in improving the yields. The National Commission on Farmers indicated that there is a large 'knowledge gap' between yields in research stations and actual yields in farmers' fields. The Planning Commission (2007, p. 3) provided an account of yield gaps with regard to various crops, which are as follows[15]:

Wheat: 6 % (Punjab) to 84 % (Madhya Pradesh).
Rice: Over 100 % in Assam, Bihar, Chhattisgarh, and Uttar Pradesh.
Maize: 7 % (Gujarat) to 300 % (Assam).
Jowar: 13 % (Madhya Pradesh) to 200 % (Karnataka).
Mustard: 5 % (Rajasthan) to 150 % (Chhattisgarh).
Soybean: 7 % (Rajasthan) to 185 % (Karnataka).
Sugarcane: 16 % (Andhra Pradesh) to 167 % (Madhya Pradesh).

It is clear from above figures that there exists a 'technology fatigue' in Indian agriculture.[16] The failure of agricultural research system and extension services is understandable in such a situation.

[14] The rural credit-deposit ratios of commercial banks for all-India were 57.98, 59.98, 39.04 and 56.26 in 1980–1981, 1990–1991, 2000–2001 and 2005–2006, respectively (Chatterjee 2006).

[15] Chand (2009) compared India's yield levels for some crops with the same for China. He found that average yield of paddy in India is less than half of the yield level in China; for wheat, it is about two-thirds of that in China.

[16] The term 'technology fatigue' has been coined by M. S. Swaminathan, the Chairman of the *National Commission on Farmers* constituted by the Government of India in early 2000s.

Summary of Findings and Policy Suggestions

Our study has clearly brought out a story of production performance of Indian agriculture deteriorating significantly in the era of economic reforms and globalization. We found that the growth rates of food grains, non-food grains and all-crop production were statistically significantly lower during post-reforms period compared to pre-reforms period. In fact, for all these categories, negative trend-break in production occurred in the period of economic reforms. During this period, the growth rate of food grains production has been reduced to a low level of 1.02 % per annum, which is far below the rate of growth of population (1.65 % per annum). The emerging situation indeed raises serious concern about food security of the country.

The decline in growth rate of agricultural production in India during post-reforms period has been widespread, so that a large number of states in India suffered from declines in growth rates of food grains, non-food grains and all-crop production during this period compared to previous period. As regards food grains, such a decline is visible in 12 out of 17 major states. All but one of these states revealed negative trend-break in food grains production during post-reforms period. Interestingly, the list of states suffering from negative trend-break in food grains production during reforms era includes both green-revolution states like Punjab and Haryana and West Bengal which ranked first among all states in terms of growth rate of food grains production during pre-reforms period. Although compared to pre-reforms period, the growth rates of food grains production were higher in five states during post-reforms period, only one of them, namely Gujarat, displayed noticeable improvement in this regard.

Almost a similar story gets repeated in respect of growth of non-food grains production. Once again, as many as 11 out of 17 major states suffered from decline in growth rate of non-food grains production during post-reforms period compared to pre-reforms period. In all these states, the growth rates of non-food grains production were statistically significantly lower during post-reforms period compared to the earlier period, implying negative trend-break in production. The list of these states includes Punjab and Haryana, the two green-revolution states. Although six states improved their growth rates of non-food grains production during post-reforms period, only one of them enjoyed statistically significant improvement (positive trend-break).

During post-reforms period, the growth rate of all-crop production declined in as many as 14 states. In 10 such states, the decline was statistically significant. The three states to improve growth rate of all-crop production during post-reforms period compared to pre-reforms period are Gujarat, Jammu and Kashmir, and Orissa. Among all states, Gujarat ranked first in growth of all-crop production during post-reforms period, which was lowest in Tamil Nadu.

When we looked into the performance of individual crops, it appeared that, as expected, majority of the crops suffered decline in growth rates of production during post-reforms period compared to pre-reforms period. At the all-India level, the list of crops suffering from such decline includes both food crops and non-food crops.

Some of the important crops to witness statistically significant decline in growth rate of production (negative trend-break) during post-reforms period are rice, wheat, groundnut, rape and mustard, soybean, oilseeds (total), plantation crops, condiments and spices, banana, and sugarcane. On the contrary, fibers (total) experienced significantly positive trend-break in the growth rate of production. Among other important crops to improve growth rate of production (though not statistically significant) during post-reforms period are coarse cereals and cotton.

The crop-level discussion for the states revealed that most of the major rice and wheat growing states suffered decline in production growth rates during post-reforms period compared to pre-reforms period. For example, states like Punjab, Uttar Pradesh, Tamil Nadu, Kerala, West Bengal, Madhya Pradesh and Assam in respect of rice, and Punjab, Haryana, Uttar Pradesh, Bihar, Madhya Pradesh and Rajasthan in respect of wheat observed negative trend-break in production during post-reforms period. The picture is no different in respect of other important crops such as oilseeds and pulses. Among non-food grains, decline in growth rates is clearly visible in fibers and plantation crops for the states known for producing these crops. However, an important exception here is Gujarat that recorded positive trend-break in fibers (mostly cotton) production during the post-reforms period. Other commercial crops like potatoes and sugarcane also suffered from negative trend-break in production in the post-reforms period. All in all, it may be concluded that the decline of Indian agriculture was so pervasive that majority of the states and a wide array of crops suffered from negative trend-break in production in the era of economic reforms.

We have looked into some important constraints faced by Indian agriculture which might explain observed fall in growth rate of production of various crops/crop groups during post-reforms period. In this context, we find that, over the years, there has been a tendency of land available for cultivation declining, though slowly. An implication of this observation is that future growth of agricultural production has to be attained through expansion of cropping intensity as also growth of yield/productivity. However, cropping intensity, at the all-India level, has been growing at a very slow pace, both during pre- and post-reforms periods. What is more, in a sizeable number of states (10 out of 17 major) the growth rate of cropping intensity declined during post-reforms period compared to pre-reforms period.

In our view, the most important constraint faced by Indian agriculture during post-reforms period is the downward trend of growth rates of various inputs. In respect of inputs like tractors, pump sets, fertilizers, and irrigation, the growth rates during post-reforms period were lower than their growth rates during pre-reforms period. This has been the picture both for all-India and majority of the states. The problem got magnified with relative decline in agriculture investment in the era of economic reforms. Our data revealed that the share of agriculture and allied activities in total investment (gross capital formation) declined over the years from its level in early 1980s.

Further, excepting very recent years, the terms of trade for agriculture was unfavorable, which reduced farmers' incomes, and hence their ability to use purchased inputs like HYV seeds, fertilizers, irrigation, electricity, and so on.

Added to this, the prices of these inputs have been rising steadily with gradual withdrawal of subsidies given on them by the government. There is also the credit constraint faced by the farmers with the government's pursuance of financial liberalization policies. We must also mention here the inadequacy of country's agriculture research system to develop new seeds and other technologies suitable for different agro-climatic conditions, not to speak of poor extension services creating a large 'knowledge gap' between yields in research laboratories and farmers' fields. Under the circumstances, it is no wonder that growth of agricultural production suffered in India and vast majority of her states in the era of economic reforms.

Policy Implications

In the light of above discussion, we may suggest a few policies to improve production performance of Indian agriculture in future: First, and foremost, steps should be taken to expand areas under irrigation which would raise cropping intensity and encourage diversification of crop cultivation. However, for expansion of irrigation, more public investment would be necessary, especially to build new medium and large irrigation structures. Pubic investment would also be needed to improve other infrastructures (roads, electricity supply etc.) in rural areas. In any case, public investment in agriculture has to be enhanced substantially if at all this sector has to return to high growth paths and contribute toward fulfillment of India's goal of attaining inclusive growth.

Second, as there exists a large 'yield gap', efforts should be made to bridge the same. In this respect, a fresh look at the priorities of Indian agricultural research system seems necessary. More funds for agricultural research should be allocated. At present, India spends only about 0.5 % of GDP on agricultural research as against more than 1 % in other developing countries (Mahendra Dev 2009). Thus, apart from the central government, the state governments are also required to support agricultural research and extension programs in a bigger way.

Third, timely supply of agricultural inputs, including credit, at reasonable prices has to be ensured. The government should review its policy of agricultural subsidy in the light of the fact that marginal and small farmers are badly affected due to rise in prices of various inputs (especially fertilizers) ever since the government's introduction of structural adjustment and globalisation policies. At a time when net returns from cultivation has been falling gradually, the production performance of Indian agriculture could not be improved unless appropriate support mechanisms (price support, insurance against risks and uncertainties, steady supply of vital inputs, marketing infrastructures etc.) are in place to incentivize the farmers.

As agriculture is a state subject, more effective planning for agricultural development is required by the states keeping in view the regional specificities. Of course, the central government recently launched some special schemes to accelerate the pace of agricultural development in the country and ensure food

security. Some such schemes are National Food Security Mission (NFSM), Rashtriya Krishi Vikas Yojana (RKVY), National Mission for Sustainable Agriculture (NMSA), National Horticultural Mission (NHM) etc. As the responsibility for implementation of these schemes rest on the state governments, they will have to play a more proactive role for their successful implementation.

Last, but not the least, attention should be focused on the issue of natural resource management. It is well-known that a large part of the un-irrigated (rainfed) areas in India are characterized by low and fluctuating productivity, weak institutional network and environmental degradation (Ibid.). This requires formulation of separate development strategies for irrigated and un-irrigated areas. Further, the decline in productivity in many areas is attributed, among other things, to deterioration of soil quality and water scarcity resulting from groundwater depletion. Therefore, management of natural resources like land and water should be accorded topmost priority. Let us hope that formulation and implementation of appropriate strategies for efficient use and conservation of these vital resources will enable India to attain the path of sustainable agricultural development in the years to come.

References

Ahluwalia MS (1996) New economic policy and agriculture: some reflections. Indian J Agric Econ 51(3):412–426

Andrews DWK (1993) Tests for parameter instability and structural change with unknown change point. Econometrica 61(4):821–856

Bai J, Perron P (1998) Estimating and testing linear models with multiple structural changes. Econometrica 66(1):47–78

Bai J, Perron P (2003) Computation and analysis of multiple structural change models. J App Econ 18(1):1–22

Bhalla GS, Singh G (2001) Indian agriculture: four decades of development. Sage Publications, New Delhi

Bhalla GS, Singh Gurmail (2012) Economic liberalisation and Indian agriculture: a district-level study. Sage Publications, New Delhi

Boyce JK (1987) Agrarian impasse in Bengal: agricultural growth in Bangladesh and West Bengal 1949–1980. Oxford University Press, Oxford

Centre for Monitoring Indian Economy (2010) Agriculture, Mumbai, June 2010

Chand R (2009) Agricultural development policies and growth of Indian agriculture. In: Agarwal M (ed) India's economic future: education, technology, energy and environment. Social Science Press, New Delhi, pp 103–127

Chand R, Raju SS, Pandy LM (2007) Growth crisis in agriculture: severity of options at national and state levels. Econ Polit Weekly 42(2007):2528–2533

Chand R, Parappurathu S (2012) Temporal and spatial variations in agricultural growth and its determinants. Econ Polit Weekly 46(26–27):55–64

Chatterjeee S (2006) Financial sector reforms and rural credit in India. J Soc Econ Dev 8(2):121–146

Cristiaensen L, Demery L, Kuhl J (2006) The role of agriculture in poverty reduction: an empirical perspective. The World Bank Policy Research Working Paper # 4013, Washington, DC

Cristiaensen L, Demery L, Kuhl J (2010) The (Evolving) role of agriculture in poverty reduction: an empirical perspective, UNU-WIDER Working Paper # 2010/36, Helsinki

Cristiaensen L, Pan L, Wang S (2010) Drivers of poverty reduction in lagging regions: evidence from rural Western China. UNU-WIDER Working Paper # 2010/35, Helsinki

Janvry A de, Sadoulet E (2009) Agricultural growth and poverty reduction: additional evidence. The World Bank Res Obs 25(1):1–20

Government of India (2008) Report of expert group on financial inclusion, Chairman: Dr. C R Rangarajan, Government of India

Gulati A (1998) Indian agriculture in an open economy. Will it prosper? In: Isher JA, Little IMD (eds) India's economic reforms and development: essays for Manmohan Singh. Oxford University Press, New Delhi, pp 122–146

Gulati A (2009) Emerging trends in Indian agriculture: what can we learn from these? 2nd Dayanath Jha memorial lecture. National Centre for Agricultural Economics and Policy Research, New Delhi

Gulati A, Kelly T (1999) Trade liberalisation and Indian agriculture. Oxford University Press, New Delhi

Joshi V, Little IMD (1996) India's economic reforms. Oxford University Press, New Delhi

Mahendra Dev S (2009) How to revive Indian agriculture? In: Singh S, Ratna Reddy V (eds) Changing contours of Asian agriculture, Academic Foundation, New Delhi, pp 203–237

Planning Commission (2007) Agricultural strategy for the eleventh plan: some critical issues. Government of India, New Delhi

Quandt RE (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes. J Am Stat Assoc 55(290):324–330

Raghavan M (2008) Changing pattern of input use and cost of cultivation. Econ Polit Weekly 43(26 & 27):123–129

Ravallion M (2010) Agriculture for poverty reduction revisited. Paper for the conference on agriculture for development Revisited, UC Berkeley

Shah T, Gulati A, Hemant P, Shreedhar G, Jain RC (2009) Secret of Gujarat's agrarian miracle after 2000. Econ Polit Weekly XLIV:45–55

Singh G (2007) Growth of Indian agriculture: a district level study. Department of Economics, Punjab University, Chandigarh

Singh M (1994) Inaugural address delivered at the 54th Annual Conference of the Indian Society of Agricultural Economics held at Shivaji University, Kolhapur (Maharashtra) on Nov 26, 1994 and published. In: Indian J Agric Econ 50(1) January–March, 1995

Venkateswarlu A (2008) Agricultural development in India: policies and performance. In: Bhaumik SK (ed.) Reforming Indian agriculture: towards employment generation and poverty reduction, Sage publications, New Delhi, pp 25–41

Chapter 7
Under the Shadow: Pricing and Marketing in Indian Agriculture in Globalization

R. S. Deshpande

Introduction

Pursuit of subsistence living has become one of the major struggles of Indian farmers. This struggle is largely dictated by the monsoon on one side and functioning of the markets (land, labour, factor and product) and 'governed' prices on the other. Current market centric economic reforms in the form of liberalisation and globalisation were introduced and agricultural sector as always was taken for granted. In this entire process, participation of the sector in the market was presumed. The entire process of reforms involved liberalisation of domestic controls and opening up of the external markets for the domestic produce. This new approach to agricultural development is market centric and essentially involves good functioning of both domestic and international markets. But the mute question is, whether we are ready for that? Is Indian agriculture market oriented? Do we have a price discovery mechanism which is the base for any globalisation based reforms? Can state withdraw from the sector as quickly as expected? It is well documented that Indian agricultural markets are far from being ready to adapt to the market centric development approach (Dantwala 1996; Deshpande 2003; Bhalla and Singh 2012). Similarly, the agricultural price policy has also undergone quite a few important changes over last six decades and any field researcher will certainly vouch that price policy hardly helped to solve issues confronting the sector for at least in the last four decades. Agricultural markets continued to be imperfect and price discovery mechanism is far from satisfactory. Quite a few questions crop up for a dispassionate analyses and one needs to address to these questions while looking into the current agricultural scenario.

- Will agricultural sector get an advantage by participating in the changing economic reforms and will it sustain the growth with a gradual withdrawal of the State?

R. S. Deshpande (✉)
Institute of Social and Economic Changes, Bangalore 560072, India
e-mail: director@isec.ac.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_7,

- Is the present agricultural policy including price policy competent enough to deal with the new challenges?
- Is our price policy conducive to absorb the market centric reforms?
- Would super-imposition of the globalisation and liberalisation processes further perpetuate the existing level of imperfections in the markets?
- Are our agricultural markets near perfect and free of exploitative intermediaries (traders and input dealers)? Are they phasing out?
- Why the pricing of agricultural crops, mainly commercial (inter alia commercial production of food crops- rice, wheat, maize and sugarcane) though regulated by states to some extent, have not been able to fully garner the benefits? Are these only due to infrastructural and institutional bottlenecks?
- Do we need fundamental changes in marketing and price policy?

State Intervention and Agriculture

Indian agriculture is largely dependent on the State and State policies. Sizeable amount of work in development economics has been devoted to the debate on the role of State as against the role of market as catalytic institutions in the aggregate development process. The debate on the role of the State began with the early theories of economic development, where it was argued that failure of allocation of investment to the deserving developmental sectors created structural imbalances in the process of development (Chenery and Srinivasan 1989). In such situation, State has to intervene as a correcting force. A group of theoreticians blamed it on the failure of the State whereas others held market responsible for that. One common understanding that emerged in this debate was pertaining to the failure of the role of State as against that of the markets. Of course, this needs to be analysed differently in the context of developing and developed nations (Chenery and Srinivasan 1989). Recently, again, the debate came alive in the context of the process of liberalisation, where the minimalist role of the State is prominently argued, so that market forces create a proper atmosphere for economic growth. In this process of thinking, it is believed that rational and most desirable resource allocation takes place through the market forces. The underlying assumption here is of near perfect and fair market related practices functioning in the sector. Besides, an inherent assumption is that the structure as well as quality of growth will be automatically taken care of. Moreover, the aggregate process will help in correcting the distortions. In view of this new thinking in the context of developing countries, it is essential to look into the basic tenets of the role of the State and that of the market in the process of development in India.

Anne Krueger (1990, p 22) while arguing on the platform of the symposium on the State and Economic Development concluded that "*At a general level there are innumerable questions as to how political and economic markets interact. At more mundane levels, there are endless opportunities for research, analysing the*

functioning of technologies, alternative policies and institutions, documenting and hopefully quantifying policy interventions, the response to them, and their evolution over time." But at the same time, we find a strong opposition to this view emerging not only from the Indian academic field but vehemently supported by well-known development economists. In the words of Byres, *"The current orthodoxy, to the effect that 'rolling back the State' and the full blooded operations of markets are appropriate to India's problems, lacks historical warrant and intellectual justification" (1997: p 37).* This signifies that the scene of argument is not very clear in the minds of development economists, especially speaking in the context of India, viz., Pranab Bardhan, Terence Byres, Amit Bhaduri, Mrinal DattaChoudhury, Arvind Pangaria, etc. Their views appear mainly due to the failure to inter-connect of micro realities with macro theories and understanding from the bottom. The ground realities clearly show almost non-existence of any text bookish market and an overshadowing role of the State.

In a systemic analytical approach, State is visualised as the provider of basic public goods (merit goods) and organises the production process through incentives and allocations while operating through various market, monetary and fiscal policies. Thus, coordination becomes an underlined function of the State. Over the decades, State recovered its presence in almost every field and this proliferation caused over dependence. That further created lack of coordination between market and the State. Besides, the failure of such coordination due to the emergence of transaction costs and rent-seeking attitude of the polity and bureaucracy together have undermined and *immensely* distorted the image of the State. More often, the distortions are hurriedly recognised as functional failures on the part of the State or the quality of the State. In this context, the intervention in agricultural policy at different levels by government in India provides enough evidence (Deshpande 2006) and that brought out the poor quality of interventions. This is despite the fact that agriculture is a private activity in its content but manoeuvred through interventions and as a subject it is assigned to the States (provincial controls) for the purpose of policy-making and implementation of programs. Six decades of experience brings out weak coordination between the Centre and the States (provinces) to achieve a well-defined set of policy objectives (Deshpande 2006). These have been mutually inconsistent economy-wide macropolicies like foreign trade policies, industrial policies on the one hand and agricultural micropolicies on the other. These have had large and offsetting impacts on agriculture (Srinivasan 2000).

While determining the role of Central government (Union Government of India) in formulating agriculture policy, the views of the leaders at the time of independence were of great importance. In a memorandum to the Cabinet during 1947, it was stated that agricultural production policy, price control of agriculture products, the establishment of the Central higher technical institutions, food distribution, and then setting up of a research Institutions and industrial development are vital problems and the Central government should take steps to correct these (Parliament Debates 1947). It was inherently accepted that Central government should have an authority of its own to coordinate agricultural production in the country as a whole and play an increasingly active role in the development of both

industry and agriculture (Food grains Review Committee 1966). The objective of the framers of constitution was to make the states (provincial governments), primarily and directly responsible of all matters concerning the development of agriculture with a focus on rural poor, and at the same time, the Central government owning the responsibility to oversee such arrangements. The powers over agriculture, and the responsibility for its development, were thus ambiguously divided between the Centre and the State (provinces). This exercise was quite fragile as the Federal inaction in its own spear could render any policy ineffective, either by the States (provinces) or by the Centre. No wonder, neither the States (provinces) nor the Centre could formulate and articulate long-standing agricultural sector policy. Marketing policy as also price policy suffered similar lacunae to large extent.

The Commission on Agricultural Costs and Prices issued the first ever Agricultural Price Policy statement in the mid-1980s but did not bother to revise that in the current circumstances. The fate of marketing policies and directives is not largely different. Presently, the functions directly undertaken by the Central government and the functions that it coordinates are:

- Overall planning and coordination of agricultural development in the country; coordinating State (provinces) agricultural plans; assuring the efficient implementation of development plans; and their evaluation.
- Assisting the states (provinces) in securing the requirements of agricultural inputs such as pure seeds, irrigation, fertilizers, and pesticides in adequate quantity and timely supply.
- Providing credit; assisting the states (provinces) in organising marketing storage and transport facilities;
- Price stabilisation; enforcing incorporation with the State's (provincial) minimum and maximum prices for agriculture produce; regulating interstate trade and movements of the commodities;
- Laying down import and export policies in respect of agricultural requisites and products;
- Providing cooperation to the states (provinces) for betterment of extension services to the formers;
- Coordinating programmes of land use and development; soil conservation and utilisation of water sources;
- The administration of external assistance to the agricultural sector.

These have not changed even a bit in the circumstances under the new economic policy. Of course, there are a few initiatives to incentivise the agricultural sector, but all these remained only on the margin. As a result, the overtly accepted market centric approach under liberalisation has not made any significant dent. Agricultural marketing and the price policy continued in the same garb with superficial tinkering and that had no substantial effect but at the same time new distortions creeped in. With a large part of the agricultural policy still being framed at the provincial level guided by the Centre and as many of the contours of

policy have not changed, agriculture sector remains under the influence of the State. As a result existing inefficiencies continued to dog the sector.

It is well-known and quite often mentioned that the contribution of agricultural sector to overall GDP is going down but the workforce is not reducing with the same rate. Larger share of the workforce still depends on agriculture as livelihood. Because of the sheer size of population, its growth and the critical unskilled workforce engaged in agriculture, the sector will continue to be critical in the overall performance of Indian economy. Our experience over six decades very clearly brings out that whenever agricultural sector growth rates experienced a trough, it was clearly reflected in the overall growth rates of the economy (See Fig. 7.1). Therefore, the volume of workforce, their livelihood security coupled with slow absorption of workforce in non-agricultural sector, essentially demands that the agricultural sector still should be at the centre stage of the policy. As such, the experience of the last two decades has not been gratifying for agricultural sector despite the special focus on agriculture in the meeting of National Development Council (2007) that focused on the stagnation of growth in productivity in the sector. Current economic reforms being market centric and agricultural markets being far from perfect are unlikely to touch the core problems of the sector. More likely the market distortions and market-based discrimination is likely to widen further.

The only agricultural policy document that exists in India is the 1999 'Agricultural Policy' released by the then government. This was followed by the fanfare of the Farmers' Commission led by Dr. M.S. Swaminathan (2004), with a few volumes. These documents as well as sporadic attempts at designing agricultural policies by different states skirted the issue of preparing Indian farm sector to confront the challenge or liberalisation. There are very little efforts to draft a new agricultural price policy or to change the present marketing structure. Direct marketing and the APMC Model Act fell far short at the threshold of implementation.

Agricultural Price Policy

Immediately after independence and after getting the grips of the food insecurity issue, evolving a long-term price policy and creating an incentive structure in the agricultural sector, it was felt necessary in order to direct development of the crop economy (Raj Krishna 1963). In view of this, the Government of India appointed L.K. Jha committee to suggest the required steps towards organising the *agricultural price policy*. Recommendations of the Jha Committee included market interventions for procuring food grains to meet the requirement for distribution as well as a protective price levels to the farmers. Following the Jha Committee report, a series of measures were taken and as a result Agricultural Prices Commission (APC) came into being in January 1965.

Prof. ML Dantwala as Chairman submitted the first report of the APC in August 1965, covering Kharif Season. The preamble states that "The Agricultural Prices

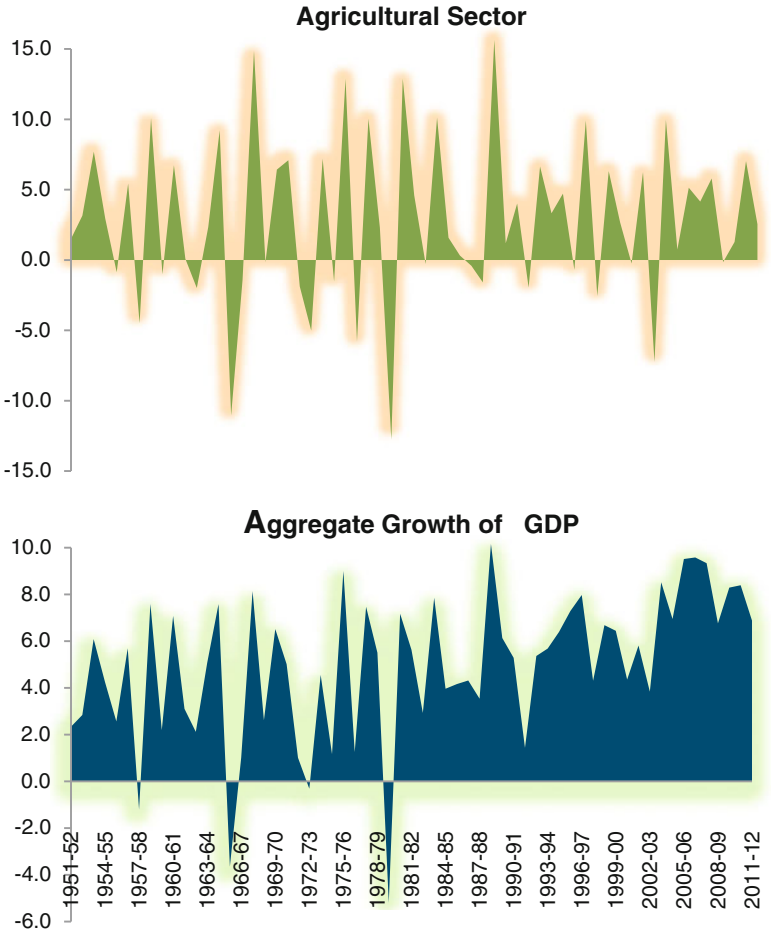


Fig. 7.1 Influence of agricultural growth rates on GDP of India (at 1999–2000 base year prices)

Commission was set up in January 1965 to advise Government *on price policy* for agricultural commodities, with a view to ***evolving a balanced and integrated price structure in the perspective of theoverall needsof the economy*****and with due regard to the interests of the producer and the consumer**" (emphasis added) (Government of India, APC Report 1965). It may be noted that the focus on the '**overall needs**' of the economy was apparent from the very beginning. This marked the foundation of the price intervention scheme and it is in operation since then. An important elimination, however, was the long run 'evolution of the balanced and integrated price structure serving the interest of the producers and the consumers' (from above quote).

It is necessary to remember that initially the price policy was placed in the context of food scarcity and price fluctuations, provoked by drought of mid-1960s. The policy was then framed keeping in view three aspects. (i) Providing food

grains for the Public Distribution System, (ii) Ensuring reasonable (affordable) prices for food grains to the consumers and (iii) Inducing adoption of the new technology. Specifically, it ensured that major economic factors that influence the rate and quality of growth are brought into play and these result in the most desired crop mix. This incidentally ensured allocation of resources, capital formation and inter-sectoral terms of trade. All these together formed a theoretical pedestal for the price policy in the context of the existing economic scenario.

The Agricultural Prices Commission was expected to keep under consideration nine important components while fixing the MSP, levy prices and procurement prices, viz., (i) Cost of production, (ii) Risk in cultivation, (iii) Changes in the input prices, (iv) Trends in the market prices, (v) Demand and supply of the commodities, (vi) Cost of living index and general price index, (vii) Fluctuations of prices in international market, (viii) Price parity between crops input and output across sectors and (ix) Trends in the market prices. There has not been any change announced on these counters ever since inception and even after the significant policy change in favour of globalization. Thus the Price policy was framed keeping in view the farmers' response to prices.

Raj Krishna in his seminal paper emphasised the price response of Indian farmers despite the dominance of subsistence farming (Raj Krishna 1963). A number of price response studies have shown the strong role of price policy to serve as incentive in managing agricultural growth (for a review see Deshpande 1996). The role was endorsed by Professor Acharya, former Chairman of CACP. He noted that "In fact, the instruments of Minimum Support Prices, Food Subsidy and Input Subsidies have played an important role in achieving the objectives of food security and accelerated growth of economy and benefits all the sections of the society" (Acharya 1997). Thus, the contribution of Agricultural Price Policy towards sustaining the tempo created by the technological change of mid-1960s has been widely acknowledged.

During the last two decades, with the forces of liberalisation and globalisation the agricultural price policy connotations have changed significantly. Prices now play much wider and more critical role than just supporting the adoption of technology and serving the food distribution system. It further needs to support the producer and provide the best commodities to the consumers. But the issues confronting the sector do not reflect that change, new issues have rather not emerged significantly. This is more so due to the minimal change of policies in the price policy and the market operations. We are continuing the same game with the outdated rules and well-expected devastating effects.

Major Issues for Changes In Price Policy

Historically, interesting issues emerged in the debate on price policy. These are reviewed succinctly by Raj Krishna (1967), Acharya (1988), Tyagi (1990), Krishnaji (1991), Jharwal (1999), Rao (2001), and GoI (2002). A few important

questions that were discussed in the context of changing price policy over and again but remained unanswered include the relationship between cost of production (CACP) and prices; authenticity and quantum of managerial costs and other imputed input costs, distortions in the price parity across crops; scientifically building and maintaining of the buffer stocks; inefficiency in PDS, targeted versus universal PDS and overall ineffectiveness of the price policy to serve the objectives set forth even in the policy statement of 1986. Beyond that the price policy did not take enough cognizance of the changing trade scenario. Furthermore, efficacy in continuation of the MSP operations in the same format given the changed economic scenario in the context of liberalisation is a question that seeks answer. All these get reflected in the imperfections of agricultural markets and the renewed awareness of these inefficiencies among the farmers. Farmers' groups all over the country have expressed displeasure about the present handling of the policy and their unhappiness may soon take shape of the strong agitation.

The methodology of arriving at the MSP is being questioned by farmers' groups and doubts are raised about the use of data, certain concepts and inclusion/exclusion of imputed cost of various items of farm operations. It was during the late 1970s, that the farmers' organisations emphasised remunerative role of prices and insisted on revisiting the method of arriving at the Minimum Support Prices. In order to reconsider the prevailing structure of the Agricultural Prices Commission and review its methodology, a Committee under the Chairmanship of Dr S.R. Sen was appointed in 1979. The Committee examined the methods of arriving at the cost of cultivation, and suggested required modifications (GoI 1980). Quite a few changes were introduced in the methodology and approach keeping with the Sen Committee report. Following this, the nomenclature as well as the focus of the Agricultural Prices Commission was changed. Subsequently, the Commission was named as Commission on Agricultural Costs and Prices (CACP) with changed terms of reference, but sharpening of the role of price policy remained unattended. A policy document was issued in 1986 under the title '*Agricultural Price Policy: A Long Term Perspective'* officially confirming the redefinition of the objectives of the price policy (GoI, 1986). After that and in the context of liberalisation we are yet to come to terms with an apt Price Policy. Perhaps Delhi will be busy preparing one such document that will remain in the stacks with academicians for the coming decades. Nevertheless, there exists a manual on agricultural prices and markets issued by CSO (October 2010), but that does not qualify as a policy document. It must be noted that the 2000 statement of agricultural policy includes providing 'remunerative prices' to farmers. Whether it is possible to do that is another question but it raised many hopes and farmers groups have deliberated on these issues in many places. Absence of a well-deliberated price policy on one hand and increasing input costs on the other left the farmers with almost stagnant net income from farming in real terms (Sen and Bhatia 2004; Deshpande and Prabhu 2005; Mishra 2006).It is very clear that distress in the farm sector emerged out of the failure to ensure income generation out of farming and that leads us to debate on price policy. This debate raises the most pertinent issue about the competence of our current price policy to accommodate the market centric approach. It is

analytically established that the price policy is neither protective to the farmers nor creates any incentive for them. Rather with the present structure of the price policy, more distortions are creeping in the market and if the efforts are not undertaken to bring in the changes, such aberrations will create significant social stress. Already the sector is reeling under severe agrarian distress and that may get aggravated.

Price Policy and Farmers' Distress

Farmers as an occupational group world over faces high risk and uncertainty in their income flow (Malmberg and Hawton 1999). In the factor market, the farmer has to pay the prices dictated by the suppliers, whereas in the product market, the purchasers determine the prices with farmer as a mute receiver. The role of the farmers in influencing the prices in either market is astonishingly minimal. Thus, a farmer faces not only weather risk along with resource scarcity in terms of availability of inputs (water, fertiliser, pesticides or seeds) but also an inevitable market uncertainty generated through prices. This is compounded by the spurious inputs and lack of information. Further, even though farming is a free enterprise, the state policies largely dictate their course of development.

The studies conducted on farmers' suicides (read acute distress) in India fall in three broad groups. The first group includes the reports Committees on the request of the State governments in order to assist the State in policy formulations. Three such reports are available for Karnataka, Andhra Pradesh, Punjab, Kerala and Maharashtra. In addition to the search for the causes of farmers' suicides, these reports essentially focus on the policy to alleviate the distress in the farming communities. Report of Karnataka state focuses on the 'Farmer Consultation Centre' approach to deal with the 'welfare domain', whereas Andhra Pradesh report highlights the inadequate public investment and need for Agriculture Technology Mission. The second group of studies comprises the Citizens Reports prepared by non-governmental organisations. Largely, these reports concentrate to locate the policy lapses of the State and are also guided sometimes by strong ideological moorings. At times, these reports create sensational news too. The third group of studies highlighted by individual researchers are: indebtedness, borrowing from money lenders, shift towards new technology, commercialisation, crop failures due to spurious seeds and inputs, increased cost of production, over-exploitation of natural resources, absence of safety nets and collapse of village as an institution. Questioning the failure of price policy and intra-market exploitations of the producers did not get into the eyes of researchers. That really says all is not well in the price and market sector and as field workers the farmers are in an agitated mood. We can expect agitations like those during 1980s and 1990s.

During the farmers' agitations during 1980s and early 1990s in Punjab, Karnataka, and Tamil Nadu, the farmer leaders insisted that the method of fixing of MSP be reviewed. The agitations were led by Mr. Sharad Joshi (an Economist with a World Organisation) and Late Prof Nanjundaswamy (a Professor of Law).

The arguments were focused on the method of computation of imputed cost, risk premium, imputed cost of family labour and the terms of trade between agriculture and industry (Nadkarni 1987; Dhanagare 1990). Following this the Government of India appointed another Committee under the Chairmanship of Prof. C.H. Hanumantha Rao to review the methodology of cost of production of crops specifically focusing on valuation of labour, imputed costs of family labour and managerial costs (GoI, 1990). The Committee submitted its report covering these aspects and suggested that actual wages to be taken to value the labour cost and family labour should be valued at the wage rates of casual labour and 10 % managerial cost (Government of India 1990). All these provided so-called 'scientific attire' to the earlier process of arriving at the cost of production. The Committee, however, did not consider the futuristic view for the price policy, where agricultural markets should function efficiently.

The current debate among the farm lobby is about the dwindling Terms of Trade (ToT) between agriculture and non-agricultural sectors. Farmers are almost unanimous in comparing the hiatus in price and income situations between sectors, even though not in very sophisticated terms, but to bring forth the disparity. The Terms of Trade debate almost emerged with a consensus that the ToT are going against agriculture till 1990s (See Fig. 7.2) (GoI,1995) and at the same time capital formation in the agricultural sector was also declining over years (Chand and Kumar 2004). While these debates were on the agricultural sector underwent substantial changes in the wake of liberalisation. We have now opened up the domestic markets for the world trade and that will exert significant pressure on the market situations. Now in this context, questions are being raised about the efficacy and effectiveness of the instruments of price policy specifically the Minimum Support Prices.

The price parity between sectors reflected through ToT was first brought forth in Thamarajakshi's seminal paper on ToT (Thamarajakshi 1969). This was followed by the work of Dhar (1968), Dantwala (1981), Kahlon and Tyagi (1980), Venkataramanan and Prahladachar (1978) and Nadkarni (1987). During the mid-1970s, the debate on ToT between agriculture and non-agricultural sector had picked up as the ToT started showing signs of turning against the agricultural sector (See arguments of Sharad Joshi in Dhanagare (1990)). This along with the farmers' movements during that decade spanning across the country, led to a review of the price policy and also the methods of arriving at the MSP. Farmers reacted very sharply to the methodology saying that CACP never releases the method it uses to derive the magical numbers each year. "The labour costs, or the

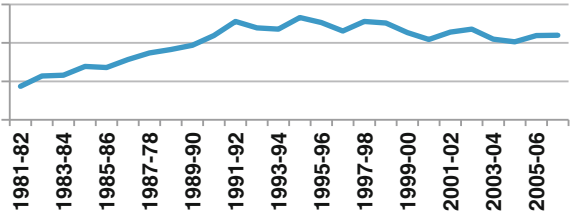


Fig. 7.2 Behaviour of changing terms of trade index

land costs are never revealed,'' says Vijay Jawandhia, of Shetkari Sanghatna of Vidarbha, an organisation for farmer rights. Once again we are likely to face the question of 'remunerative prices'. Now that the concept has been recognised in the Agricultural Policy document of Government of India (GoI, 2000) the issue will feature more prominently among farmers. Farmer leaders are already arguing for providing such remunerative prices. The question became sharper now in the context of withdrawal of subsidies on inputs (fertilisers, water, credit and power) as well as increasing demand for consumer durables, and consequently the changing relative prices with the non-agricultural sector. Similarly, the price wedge between goods produced in urban sector as against the farm products has given rise to the necessity of looking afresh into this issue of price intervention.

Viewed from this angle, the effectiveness of MSP now assumes a totally different context. In a nutshell, the debate pointed out that ToT went against agricultural sector till the mid-1960s and slightly became favourable to agricultural sector for a short while in the late 1960s and early 1970s to revert back against agriculture during late 1970s and early 1980s. It is only in the 1990s that the ToT are turning in favour of the agricultural sector (Government of India 1995). The computations of the ToT largely rest on the data from National Accounts Statistics and hence there are a good number of corrections that are required in the data in the framework of National Accounts Statistics. Even with these corrections, it was observed in one of my earlier studies that ToT in the recent past is slightly turning in favour of Agriculture, but needs to be watched carefully (Thippaiah and Deshpande 1998).

The ToT has been fluctuating in favour of agriculture and against the agriculture alternatively. During 1990s, the ToT was improving in favour of the agricultural sector but not as much as required for the growth of the farm sector. One of the important reasons for such behaviour can be located in depressing trends in the relative prices between agriculture and non-agricultural goods. The prices of agricultural commodities have not been rising at the pace at which prices of non-agricultural commodities are rising. *Therefore, the emphasis on MSPthat seems to be only providing the psychological support to the farmers,continues.*

Need for Long-Term Policy

The price policy related concerns also featured prominently in the Report of the High Level Committee on Long-Term Grain Policy (GoI, 2002). The Committee elaborately discussed the question of Minimum Support Prices and the operations in view of the effectiveness of the scheme, decentralisation of price support and procurement and alternatives to MSP. The Committee recommended that MSP should be continued incorporating corrections which include: "(i) the CACP should be made a statutory body; (ii) CACP should act directly on the basis of C_2 cost of production; (iii) CACP should also indicate a system of imputing family labour cost; (iv) CACP should recommend uniform price for Paddy for the entire

country; (v) All the procurement agencies and Public Grain Management Institutions should be legally bound by the MSP Policy; (vi) Central government should under-write open purchase of grains under MSP; (vii) FCI should be the buyer of the last resort. FCI should withdraw from a few states and concentrate on other states" (based on the detailed recommendations given by the Committee, GoI 2002, pp 9–10). The Committee also looked into the possibility of decentralised procurement scheme where it recommended that the grain procured under decentralised scheme must be treated as part of the central pool with FCI. It further stated that based on the guarantee of central purchase there should be an open-ended bank credit on the lines of FCI provision to the states involving decentralised procurement. These are some of the far reaching recommendations to keep pace with the changes in the context of liberalisation.

There is also a strong view that the entire MSP declaration has political overtone. It is reported that the prices recommended by CACP are often modified by the Government of India (Rao 2001) with the intervention of the political representatives, and therefore fixation of the prices with an elaborate structure and mechanism remained only an academic exercise. The political interventions occurred selectively across crops depending on the crop region and the active lobby. Therefore, some crops received better deal while a few other crops suffered a relative neglect. This hampered the price parity across crops. It also created distortions between the trends in factor prices and product prices for a few selected crops and farmers. Thus, a need to have a fresh look at the MSP and such review must consider its operational efficiency as the vital objective (Narayanamoorthy 2012).

MSP covered a large number of crops across the country. Over years, CACP has added quite a few crops to this list without providing rationale and probably not reviewing the list of the crops on the basis of effectiveness in overall price policy as well as operations of the scheme over years. This also smells intervention of different lobbies. In the process, a good number of crops were added vitiating the inter-crop price parity and the policy became instrumental to such process of deliberate policy neglect, indirectly discouraging certain crops and crop-groups through relative prices. The cropping pattern acted as a conduit to transfer this effect of inequality across regions and farmer groups. *Largely the crops that received raw deal in terms of relative prices were those grown by resource poor farmersand in slow growth regions such as millets and underutilized crops*. The inter-crop price parity of a few crops with wheat and paddy (which benefited the most from MSP), was worked out and it clearly shows that parity has not been followed (Deshpande and Naika 2004).

The relative MSP of the crops indicates a continuously downward trend for coarse cereals and sugarcane. For cotton, however, there are some fluctuations. The relative price series in terms of Wholesale Prices do not show a similar trend as that of the relative MSP. It is quite intriguing that even though similar inputs are used in all these crops and the cost of cultivation of all the crops is increasing almost in a similar pattern, the trends in MSP are dissimilar (Various Reports of Commission on Agricultural Costs and Prices). Invariably, the relative MSP of these crops shows declining trends thereby indicating built in price differential in

the policy, even though the cost of cultivation has similar trends. That probably brings out the policy bias against these crops. In addition, it is also difficult to appreciate how the MSP for dry land or rainfed crops is on par or higher than irrigated crops. This shows that water is totally discounted and undervalued in the cost of cultivation as well as in MSP.

There are hardly attempts to check the authenticity of the data collected under cost of cultivation schemes for different crops, and that is often challenged by the farmers. In addition, CACP is translucent since it does not allow the state units to deliberate with the academicians and farmers regarding reliability of the data collected by the Cost of Cultivation Schemes at state level. Unless this is initiated, the CCSs will continue to be unaccountable to the public with regard to quality of data, let apart the democratic process which ultimately super-imposes the CACP recommendations. CACP arrives at MSP after long deliberations and based on the cost of production data collected from the country-wide centres. In order to check the relationship between of Cost of Production (A_2) and (C_2) with the declared MSP, three time series are plotted for major crops. It was evident that MSP broadly covers Cost C_2 in the case of paddy, wheat, groundnut and gram and, there is some consistent trend in these series. But for jowar, bajra and ragi, throughout the decade of 1990s, MSP had been lower than Cost C_2. This is also true for cotton. These facts (extracted from CACP reports) provide a clear evidence of the relative neglect of jowar, bajra, ragi and cotton.

MSP as a Price Policy Tool: A Cross-Section View

In order to obtain a complete view of the effectiveness of Minimum Support Price across states included in the study, certain parameters determining effectiveness have been extracted from the reports of the study team. A countrywide study was undertaken to review the effectiveness of MSP across states coordinated by the author. The reports revealed quite a few intricacies (Shroff 2003; Ghosh 2001; Ratnam and Rao 2002; Athavale 2002; Singh et al. 2002; Sinha 2002; Singh 2002; Swaminathan 2002). The effectiveness is graded into five levels starting with 'very effective' to the 'ineffective' level as these studies have utilised different ways of analysing the empirical results. Therefore, the levels here are arrived at after completely studying the reports and eliciting the required information on a common scale. Annexure Table A.1 presents the picture at a glance. It can be seen that the process of implementation is carried out with two purposes in view. First, it is for the purpose of procurement and the second for providing a 'cushion' to the farmers against violent downward price fluctuations in the market. The first objective is to mop the available marketable surplus in the food surplus regions to feed the Public Distribution System, whereas the second objective aims at providing support against the income loss due to price collapse. We find that the procurement objective has been largely successful during 1990s only in the states of Punjab and Haryana in the first place and Uttar Pradesh, Tamil Nadu, Madhya

Pradesh and Bihar at the second level. *But the objective of providing cushion against the price fluctuations has not been successful in most of the states. A large number of farmersare not even aware of MSPand the awareness is confined to 4–5 states and even within these states the awareness is largely confined only to the commercial crop belt.* It was observed that majority of the farmers do not get the expected prices for their products and their expectations are not unfounded. The relationship between Wholesale Price and Farm Harvest Price on one hand and Minimum Support Price on the other hand seem to be largely dictated, and all the three prices move in the same direction over time. There are two different views coming out of this relationship viz., (i) Minimum Support Prices keep pace with the Wholesale Prices and Farm Harvest Prices due to probable transmission of price signals. This is more due to the behaviour of Minimum Support Price to stick around the Wholesale Prices and Farm Harvest Prices and (ii) The Wholesale Prices and Farm Harvest Prices help in deciding the trends in the Minimum Support Price in most of the states under review. The Wholesale Prices and the Farm Harvest Prices were related with Minimum Support Price. It was observed that, in the states where the MSP was below the FHP and WSP, procurement has been undertaken and the largest procurement is effected only from these states. *This defies the very objective of MSPas a protective policy.*

MSP has been introduced as one of the important policy interventions that helps in dictating the levels of input use (as reflected by improvement in yield level), adoption of technology and as an incentive to increase the capital formation. It was noted that except four states, MSP could not live upto this expectation. The role of MSP as an important variable in the decision-making process of the farmer also gets confined only to a few states and regions within (Punjab, Haryana, Tamil Nadu and Andhra Pradesh). *Thus, the Minimum Support Price has not been an effective policy tool as a variable in the process of decision-making; as a lever to absorb the market fluctuations; as an incentive to adopt the new technologyand application of new inputs; as a leading price to dictate market prices and Wholesale Prices and finally as a cushion to the farmersto protect from the market imperfections. That calls for a relook in the policy.*

Globalisation and Need for Revisiting MSP

MSP is now viewed as a market intervention on the part of the State and also as one of the supportive measures (safety nets) to agricultural producers. Even though this is perfectly WTO compatible, there are some doubts about its continuance and effectiveness to deal with the objectives set by the architects of the reforms. The issues that dominated the current debate include reasons for continuation of the price support scheme; its effectiveness in terms of the objectives as set forth in the 1986 document and support price vis-à-vis remunerative price approach. We must now realise that the context of price policy has changed substantially over the years and so also the direction and effectiveness of price policy as a tool to

influence the agricultural economy. This provoked many social scientists to argue for a fresh look at MSP as an instrument for interacting with some parameters of the agricultural economy.

Initially, its role was perceived from the viewpoint of incentivising farmers to adopt the new seed-water-fertilizer technology. The initial role of MSP as an incentive to adopt technology is apparent in the writings of Professor Dantwala, the founding architect of India's price policy.: "Though no rigid formula has been accepted to determine the levels of floor prices, the criterion followed is that *progressive farmersshould find these levels adequate to encourage enterprise and investment to augment production through the adoption of improved technologywith all its risk and uncertainty* (emphasis added)" (Dantwala 1996, pp. 213 originally published in 1967). Prof Dantwala wrote again during early 1990s: "Likewise, intervention has to be selective. *Its need must be clearly established and its effectiveness should be constantly under review* (emphasis added). The real problem is not simply to establish the legitimacy of intervention, but that of ensuring its effective and judicious implementation" (Dantwala 1996, pp. 292, originally published in 1993). Following the logic of Prof Dantwala it is time to get at a serious review of MSP as a policy to render it more meaningful in the current context.

With the changing scenario of agricultural sector under liberalisation, the price and market intervention schemes may require significant changes. We find two opposite views expressed by academics. First group believes in fully revamping the price policy in the context of liberalisation (Ashok Gulati, Sharad Joshi). The second group suggests retaining the schemes but changing the structure to suit the present needs (Bhalla (1994) and Sen Committee Report, GoI 2002). A clear analysis of this question requires a review of the scenario of agricultural price policy in the post 1991. The current trends due to liberalisation are expected to induce competition in the factor and product markets. Primary signs of this could be visualised in the new market-oriented changes in the cropping pattern and availability of seeds, pesticides and fertilisers. Quite a few changes are taking place in the product markets too but these are sporadic in nature. Agricultural marketing being the State subject, changes are not uniform within and across states and have not been planned with any focussed theme. A few States have been taken initiatives to provide the farmers with updated market information, through electronic media on daily basis of the main/major market and in a few other states on weekly basis. Farmers' markets (*RayatSanthe*) have been established over-shadowing the earlier process of marketing dominated by middlemen. The removal of the restriction on the interstate movement on the agricultural commodities has also contributed in effecting some changes in the markets.

Agricultural Marketing Scenario

Agricultural marketing in India has changed through different phases of agricultural development but mainly through State regulations than market signals. The interventions and provision of legal framework were essential as the existing mechanism

was more truncated against the sellers and favoured unhealthy practices. It was during the First plan period that the Planning Commission had directed the state governments to bring agricultural marketing under specifically enacted Regulated Market Act (called as Agricultural Produce Marketing Act) which most of the states complied with, rather at slow pace. The problems of marketing failures were highlighted and it was expected that the legal framework would take care of the following problems: Undercover Sale; Removal of large samples; Unwarranted trade allowances; Heavy market charges; Unauthorised deductions; Incorrect weighing and multiple units of measurement; Absence of grading; Presence of touts. My 2 years stint with Agricultural Price Commission of Karnataka during 2004–2006, and the field visits of the Commission clearly revealed that all these problems still persist in the regulated markets prominently and sometimes with deeper intensity.

Historically, the problems compounded the market imperfections and therefore a series of steps were taken to deal with these. In view of the imperfections, agricultural marketing institutions after independence have undergone changes. The Regulated Market Acts enacted by various State governments incorporated a legal framework to deal with some of the issues (see Table 7.1). These changes could be broadly put into four broad groups, namely: (i) Introduction of Agricultural Marketing Institutions in India, (ii) Creation of marketing infrastructure, (iii) Co-operative Marketing as an alternative to protect weaker participants in the market and (iv) Emergence of the New Regulated Market Act and subsequent modifications in that. It is apparent that these changes occurred due to the prevailing circumstances and the ongoing reorganisation in the agricultural sector during these early decades. Four components predominated the policy interventions viz.: (i) Putting in place regulation of marketing functions and removing imperfections; (ii) Creating infrastructure to facilitate the process of marketing; (iii) Introduction of Price intervention schemes; (iv) Procurement and distribution of essential commodities. All these measures were operating simultaneously and therefore had an overlapping effect on marketing sector.

The Agricultural Marketing Boards in most of the states and each of them function under Agricultural Produce Marketing Committee (APMC) Act of the respective states. A quick review of the regulations across states reveals that there is an urgent need to bring more uniformity in powers and functions and demarcations of activities between the Directorate of Marketing and State Agricultural Marketing Boards. Even the model APMC act did not achieve this simple review for reasons beyond comprehension. This can facilitate proper regulation of marketing practices as well as building more infrastructure facilities so as to expedite growth towards competition.

Regulated Markets: Regulating Irregularities

The agricultural market infrastructure in India has been inadequate to handle the situation. Despite increase in the number of regulated and wholesale markets, the

spread of these markets across the states has been uneven. The average area served by each regulated market also varied considerably among the states of India, from 103 Km^2 per market in Punjab, 1185 in Himachal Pradesh. If we consider other north Indian states then this variation will even wider. Area coverage based on geographical area may not reflect true requirement of agricultural markets. If we take gross cropped area as an alternative indicator, the scenario does not show many changes. West Bengal, Andhra Pradesh, Tamil Nadu and Haryana are in quite better position as compared to other states. As can be seen from the Annexure Table A.2, the density, measured as number of regulated markets per lakh hectare of GCA, was also insufficient. At all India level there were 3.3 markets per lakh hectare of GCA existed in 1991 has marginally increased to 3.7 in 2010 (See Figs. 7.3, 7.4 and 7.5). This indicates that even the increment in the density is modest. It can also be observed that ratio of markets per lakh hectare of GCA was lower than the national average for states like Uttar Pradesh, Orissa, Rajasthan and Gujarat (See Fig. 7.6) but was above the national average for other states. Research studies revealed that farmers on an average get 8–10 % higher price and higher share in the consumer's rupee by selling their produce in the regulated markets compared to rural, village and unregulated wholesale markets. The benefits obtained by the farmers by sale of agricultural produce in the regulated market varies from area to area because of the variation in the spread of regulated markets over the regions and the existence of necessary infrastructural amenities/facilities in these regulated markets (Singh 2005; Jairath 2010). The National Commission on Agriculture (1976) and National Commission on Farmers (2004) have recommended that the facility of regulated market should be available to the

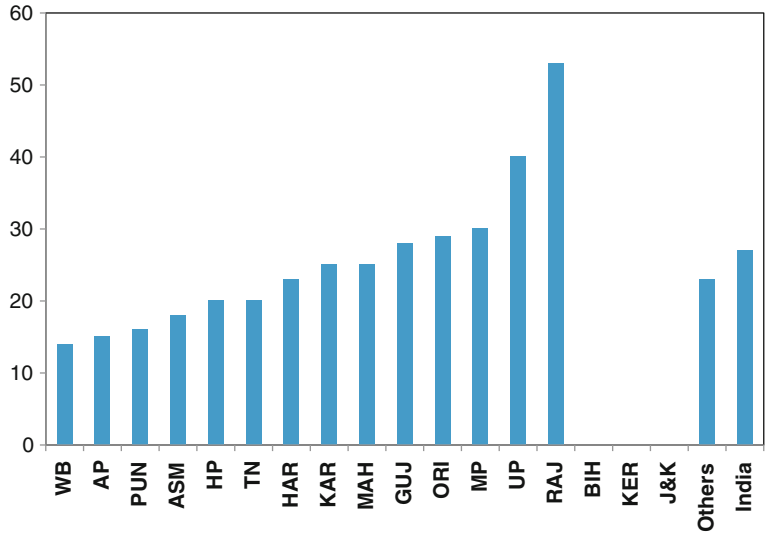


Fig. 7.3 Area (GCA) served by per regulated market—2010 (Area in 000 ha. Data for Bihar Kerala and J&k are not comparable)

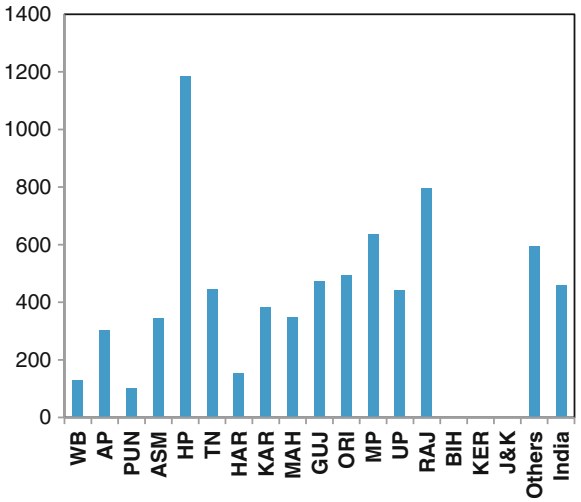


Fig. 7.4 Geographical area served by per regulated market—2010 (Area in Sq. Km)

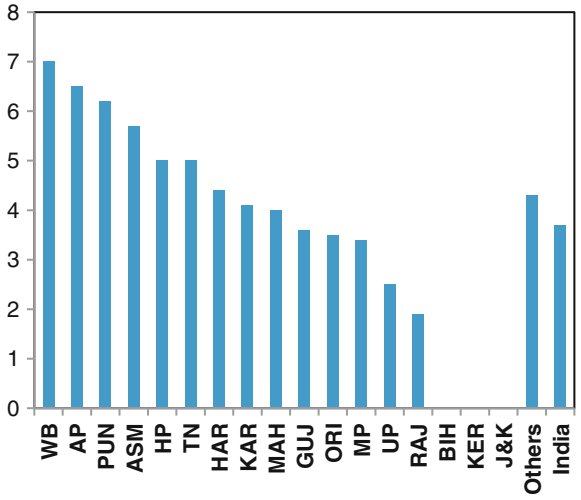


Fig. 7.5 Market (Regulated) density per Lakh GCA (2010)

farmers within a radius of 5 km. If this is considered a benchmark, the command area of a market should not exceed 80 Km^2.

The facilities created in market yards are inadequate. The cleaning, grading and packaging of agricultural produce before sale by the farmers have not been popularised by the market committees on a sufficient scale. Even facilities for these have not been created in most of the market yards. Marketing Board is almost the same in all the States where Statutory Boards exist; a broad variation has been observed in their composition/constitution and functioning. It is necessary to bring more uniformity in powers and functions of Boards and demarcations of activities between the Directorate of Marketing and State Agricultural Marketing Boards. This can

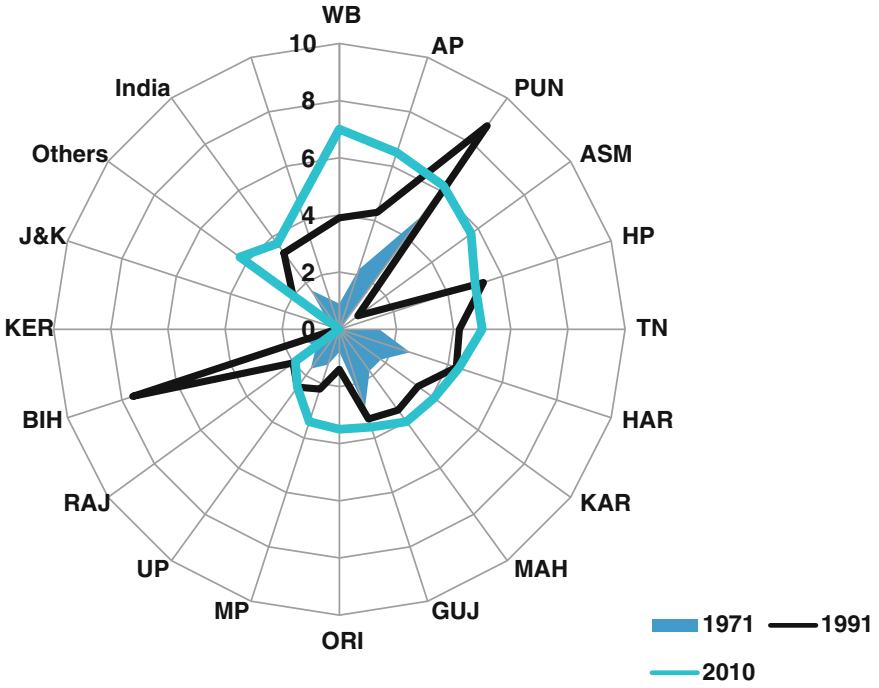


Fig. 7.6 Market (Regulated) density per Lakh GCA

facilitate proper regulation of marketing practices as well as building more infrastructure facilities so as to achieve a faster growth and better private participation. The benefits available to the farmers from regulated markets depend on the facilities/amenities available rather than the number of regulated markets in the area. Both covered and open auction platforms exist in two-thirds of the regulated markets. One-fourth of the markets have common drying yards. Traders modules viz shop, godown and platform in front of shop exist in 63 % of the markets (Jairath 2010).

The cold storage units exist in only 9 % of the markets and grading facilities exist in less than one-third of the markets. The basic facilities viz., internal roads, boundary walls, electric light, loading and unloading facilities and weighing equipments are available in more than 80 % of the markets. Most of the regulated markets at present still awfully lack facilities for handling produce as less space is available for auction platform, there are inadequate number of shops and godowns in the premises etc.; this reduces effective competition (Jairath 2010). The food grains storage capacities across states vary considerably on a comparative scale (See Fig. 7.3). Absence of storage godowns at market level further perpetuates the problems of traders in general and continuous movement of goods in particular. Various State governments recently initiated a process of direct marketing by producers to the consumers in the country by initiating the concept of *ApniMandi* (Punjab), *Rythu Bazar* (Andhra Pradesh), *UzahaverShandies* (T.N.) and Shetkooribazers in Maharashtra. But these

markets have been promoted so far only at the State headquarter and some district headquarters adjoining to the state. Rural periodic markets/*haats/shandies* are the first contact point for producer—sellers for en-cashing agricultural produce. There are about 27,294 rural periodic markets in the country. The minimum necessary infrastructural facilities do not exist in these markets.

Agricultural Marketing Act

The Agricultural Produce Marketing Act made provisions for the regulation, establishment and administration of markets for agricultural produce. It constituted the Market Committees in these regulated markets and entrusted them with the formulation of rules in the market yards and monitoring the functioning and conduct of business. But, in the process the APMCs are emerging as monopolistic centres and the market system is left with major inadequacies like lack of accountability of the full volume of business transacted in the regulated market with the connivance of traders, lack of proper infrastructure (including density) to satisfactorily provide clear access to markets for all the farmers, difficulties in the mode of payment as well as storage and transport facilities in the market yards, domination of 'under-cover' Commission Agents in the operations inside the market yard and inadequate dissemination of market information to the farmers. Hence, measures have to be taken for proper monitoring of the implementation process of the provisions of the Act in letter and spirit.

For implementation of Model Act on agricultural marketing, the central government has taken different steps like consultations, persuasions and subsidy incentives on infrastructure projects, to bring changes in APMC of State governments on the lines of the Model Act. So far progress made on this account is shown in Table 7.1.

Market intervention through the State agencies has been at a low key so also the market reforms. It is taken up sporadically and at times under pressure from farm lobby. Market intervention scheme, actually, requires arrangements for a permanent institution that does not exist. Further, the time lag between the signals originating from the distress caused due to price and the initiative of action, defeats the very purpose of the action. More than the financial constraints, the MIS operations require adequate autonomy to the agencies and a business culture like that of NAFED. *It should also be ensured that these agencies are located in the market yards and APMC yards and should be ready to gear up to face contingencies of deficit or glut in supply.* There has always been asymmetry in the dissemination of knowledge about prices, the markets and various schemes, among the farmers, traders and other actors in the market. Awareness is not created about the latest market developments among the farmers and stakeholders to sell their produce such that they reap the maximum benefits out of their sale.

In an overall view, market imperfections are an important factor that obstructs the path to a developed marketing system. Farmers do not have the necessary

Table 7.1 Progress of reforms in agricultural markets (APMC Act) as on 30.11.2009

Sl.No.	Stage of reforms	Name of states/ Union territories
1.	States/ UTs where reforms to APMC Act has been done for Direct Marketing; Contract Farming and Markets in Private/Coop Sectors	Andhra Pradesh, Arunachal Pradesh, Assam, Chhattisgarh, Goa, Gujarat, Himachal Pradesh, Jharkhand, Karnataka, Madhya Pradesh, Maharashtra, Nagaland, Orissa, Rajasthan, Sikkim and Tripura.
2.	States/ UTs where reforms to APMC Act has been done partially	(a) Direct Marketing: NCT of Delhi. (b) Contract Farming: Haryana, Punjab and Chandigarh. (c) Private markets: Punjab and Chandigarh
3.	States/ UTs where there is no APMC Act and hence not requiring reforms	Bihar[a], Kerala, Manipur, Andaman & Nicobar Islands, Dadra & Nagar Haveli, Daman & Diu, and Lakshadweep.
4.	States/ UTs where APMC Act already provides for the reforms	Tamil Nadu
5.	States/ UTs where administrative action is initiated for the reforms	Mizoram, Meghalaya, Haryana, J&K, Uttrakhand, West Bengal, Puducherry, NCT of Delhi and Uttar Pradesh.

[a] APMC Act is repealed w.e.f. 1.9.2006

*Source*http://agmarknet.nic.in/amrscheme/apmcstatus08.htm

shield against these imperfections and safeguard their interests. Insurance has no way protected the farmers. Establishment of separate Commodity Boards for specific agricultural commodities can effectively circumvent the market imperfections but such experiments are sporadic. It is quite evident that in the process of economic reforms, agricultural domestic market reforms have been totally neglected. The process of globalisation is handled only at the level of international trade with domestic markets confronting problems due to imperfections. Attempts to bring reforms in the domestic agricultural commodity markets have failed and in fact helped to perpetuate imperfections. We have neither attended to correct the distortions in price policy nor helped to correct market imperfections. The policy thus will introduce new forces of inequality.

Concluding Remarks

The process of globalisation was expected to release new market forces in the domestic market environment. Even though the domestic markets are not fully integrated with the world market, the pressure of international prices is a potential threat to domestic consumers as well as producers. Similarly, the liberalisation in other sectors of the economy has exerted significant pressures on the role and functions of the domestic markets. The inefficient functioning of the institutions governing markets and prices have contributed to the already existing market imperfections and market failures. This has not only resulted in high food inflation on one side, but also farmers' distress and lopsided share of market intermediaries on other. The situation of distress in agricultural sector has been analysed and the conclusions begin with indebtedness, low income, technology failure and other many components. However, my fieldwork and research in the three sectors namely the price policy, market inefficiencies and agrarian distress brought me to the conclusion that we have failed at the threshold of price policy and in institutionalising agricultural markets (Deshpande 2003, 2006, 2008; Deshpande and Naika 2004; Deshpande and Prachitha 2005). In the present era of reforms wherein, we are trying to reduce the role of the state and allow markets to function freely, it is not understandable that still the role of the State (not being able to be absolutely efficient as it is) is not reduced. Therefore, two of the most important contested issues, the price policy and dealing with the market imperfections remain untouched in the reforms.

The farm sector distress in most of the states has been analysed from different angles and the substantial literature is available on the subject. One of the major points that missed the analysts is the inability to generate higher net income in the farm sector. This happened both due to stagnation in productivity on one hand, increasing prices in the farm inputs with the product prices not keeping pace with the increased cost of cultivation, on the other. It is not that the farmers do not intend to pay back the dues to the banks or moneylenders but most of them are not in a position to generate that income required due to failure of the price policy and market imperfections. The mood in the farm sector is now a matter of concern and

farm leaders like Raju Shetty, Manveer Singh Tevatia, Y. Sivaji, Vijay Javandiya, M.R. Sivaswamy, Puttannaiah and many others have given strong warnings about the price policy as well as market imperfections. The writings on the walls are very clear and bold that the farm sector is agitated and the situation may aggravate any time. The major issues spoken of by the farmers lobby at micro level include the promise given in the 2000 Agricultural Policy about remunerative prices and maintenance of the price parity. The declaration of support prices and the method of computation of MSP has always been one of the points in all these agitations across the country. Absence of the price policy and market reforms is the major concerns now. The situation is quite critical and may surface seriously soon.

Long-Term Grain Policy deliberated the problems confronted by the Minimum Support Prices and the operations to analyse its effectiveness. The committee favoured decentralisation of price support and procurement as an alternative to MSP, an argument put forward earlier by Rao and Deshpande in 2002. Failure of the MSP policy is traced to faulty data, failure to adhere to the guidelines, political interventions, imputation methodology and failure to maintain across crop parity. The Committee recommended that MSP should be continued albeit in a new form. The suggestions included making the CACP an empowered statutory body and thus the intervention by the government will be minimised. But this also needs transparency in the operations of CACP specifically about the methodology of arriving at MSP. All the computational methods should be clearly stated along with the basis of the assumptions about all costs including the imputed costs. All the procurement agencies and Public Grain Management Institutions need to be legally bound by the Policy. The committee also recommended that the Central government should under-write open purchase of grains under MSP and FCI should become the buyer of last resort. FCI can gradually focus on other states where they missed the rural bus.

Main problem in the policy being ineffective is the involvement of the state bureaucratic procedures in the entire operations of MSP and its failure as a tool of price policy is largely to inefficient administration. The time lag between the incidence of distress price and procurement operation is the challenge. Therefore, procurement centres should be located in the market yards where the farmer should have a choice to exit if the product is being auctioned at a lower price. Farmers group along with the APMC should be assigned the responsibility to procure and disperse the procured grains. This could be achieved on a decentralised level. It is better that the stakeholders manage their own issues with the state support rather than state taking the responsibility and then take the blame of mismanagement.

The preambles of the APMC acts invariably state that the Act was brought into plug the inefficiencies prevailing in the market but the result seems to be regularising the irregularities prevailing then. The market infrastructure is painfully inadequate and each market serves large areas economically making impossible for farmers to take advantage of the institution. The density of markets in many of the states is so thin that farmer prefers to sell at the village and to the middlemen. The analyses of the various APMC acts show that largely the Acts are focused on the operations of the APM Committees and a large number of references to

litigations in elections. The APMC yards are more dominated by political interests rather than the interests of the farmers. Besides, the increased density of small and marginal farmers has caused lower marketable surplus available per farmer reaching the market yard and the farmer also finds it easy to sell the surplus to a middleman rather than taking the small lots to the market. If markets have to become efficient, it will be necessary to reduce the role of the state, middlemen and the farmer should feel free to sell the product at his/her will. It will be essential that the farmers themselves manage the markets through direct sale, through group marketing or other modes where producer share of the consumer rupee is appreciable. For this, the new model act will facilitate but has not come under implementation in many states.

Globalisation process has opened up the gates for international trade and liberalised a few barriers. Small ripples did get reflected in the agricultural trade but the domestic market imperfections and the price policy continued to suffer under imperfections. Precious little was done on this policy front. Agricultural markets and price policy continues with the distortions that entered into this arena during seventies and eighties. The resultant of this neglect can be many folds such as the warning sounded by Bhaduri (2008) about increasing inequality and State growing hostile towards welfare. Deepak Nayyar pointed to the real failure in the second half of the twentieth century as our inability to transform growth into development (2006). Agricultural sector with the major constraints of imperfect market and distorted price policy cannot expect to reach the desired goal of development under the market centric globalisation policy, unless quick policy amends are undertaken. Even now the situation is that if a farmer enters the regulated market yard she/he faces all the irregularities and cannot come out without disposing off the produce at the will of a few operators and prices that do not cover even his

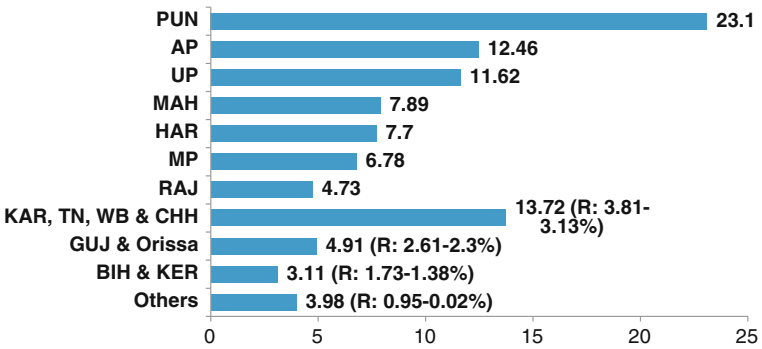


Fig. 7.7 Percentage distribution of overall covered foodgrains storage capacity across Indian states (2011) (*NoteR* indicates Range of the values; *Source* Annexure Table A.3)

cost. This must change (Fig. 7.7).

Table A.1 A comparative picture across states of effectiveness of MSPMSP

Sl No	States	Relationship of MSP with		Impact on			Role as decision making variable	Area response	Process of implementation		Major crops	Per cent of farmers who	
		WSP	FHP	Yield	Technology	Incentives			Procurement	Cushioning		Are not aware (%)	Did not get expected prices (%)
1	PUN	****[a]	****[a]	*****[b]	*****[b]	****[a]	****[a]	****[a]	*****[b]	**[d]	Wheat, Paddy, Cotton	30–90	6–20
2	HAR	****[a]	****[a]	*****[b]	****[a]	****[a]	****[a]	****[a]	*****[b]	***[c]	Wheat, Paddy, Cotton	90	5–20
3	UP	***[c]	***[c]	***[c]	****[a]	****[a]	****[a]	****[a]	****[a]	***[c]	Wheat, Gram, Sugarcane	20–80	10–15
4	KAR	***[c]	***[c]	**[d]	***[c]	**[d]	**[d]	**[d]	**[d]	**[d]	Paddy, JowarRagi, Tur, G.Nut	2–58	20–45
5	MAH								**[d]	**[d]	Jowar, G.Nut, Soyabean	1–50	10–25
6	GUJ	****[a]	****[a]	***[c]	***[c]	**[d]	***[c]	***[c]	***[c]	**[d]	Bajra, Maize, Wheat, Cotton G.Nut	80–90	20–40
7	AP	****[a]	****[a]	****[a]	***[c]	**[d]	**[d]	**[d]	***[c]	***[c]	Paddy, Jowar, G.Nut, Sugarcane	12–100	12–60
8	TN	****[a]	****[a]	***[c]	***[c]	****[a]	***[c]	***[c]	****[a]	***[c]	Paddy, Sugarcane, G.Nut, Pulses	36	25

(continued)

Table A.1 (continued)

Sl No	States	Relationship of MSP with		Impact on			Role as decision making variable	Area response	Process of implementation		Major crops	Per cent of farmers who	
		WSP	FHP	Yield	Technology	Incentives			Procurement	Cushioning		Are not aware (%)	Did not get expected prices (%)
9	MP	***[c]	***[c]	****[a]	***[c]	***[c]	****[a]	***[c]	****[a]	***[c]	Wheat, Jowar, Gram, Soyabean, G.Nut	40	48
10	BIH	***[c]	***[c]	***[c]	****[a]	***[c]	***[c]	**[d]	****[a]	***[c]	Paddy, Wheat, Jute	27	40
11	WB	****[a]	***[c]	***[c]	***[c]	***[c]	**[d]	*[e]	***[c]	***[c]	Paddy, Wheat, Jute, Mustard	16–80	60

Note[a] Very Effective; [b] Effective; [c] Average; [d] Less Effective; [e] Ineffective

Table A.2 State-wise total number of regulated markets and area served by them

States	Geographical area served by per regulated market (Area in Sq. Km)					Area (GCA) served by per regulated market (Area in 000 ha)					Market (Regulated) density per Lakh GCA				
	1971	1981	1991	2001	2010	1971	1981	1991	2001	2010	1971	1981	1991	2001	2010
WB	1305	303	263	160	129	105	26	26	16	14	0.9	3.8	3.9	6.1	7.0
AP	929	510	484	319	305	45	23	23	16	15	2.2	4.4	4.3	6.4	6.5
PUN	182	145	76	75	103	20	19	11	12	16	4.9	5.1	8.8	8.5	6.2
ASS	–	4902	2451	2241	347	–	215	119	117	18	–	0.5	0.8	0.9	5.7
HP	–	1591	1071	1591	1185	–	27	19	27	20	–	3.7	5.3	3.7	5.0
TN	1227	531	471	482	445	70	26	24	23	20	1.4	3.8	4.2	4.3	5.0
HAR	345	248	172	156	156	39	31	23	22	23	2.6	3.3	4.3	4.6	4.4
Others	28350	9021	4510	3422	596	191	44	49	176	23	0.5	2.3	2.0	0.6	4.3
KAR	969	601	483	405	383	55	33	30	26	25	1.8	3.0	3.4	3.9	4.1
MAH	900	567	398	359	350	55	37	28	26	25	1.8	2.7	3.5	3.9	4.0
GUJ	641	660	575	495	473	34	36	30	26	28	2.9	2.8	3.3	3.8	3.6
ORI	2831	2290	1198	1081	496	123	129	74	55	29	0.8	0.8	1.4	1.8	3.5
MP	1680	1381	834	720	636	78	67	45	29	30	1.3	1.5	2.2	3.4	3.4
UP	745	477	464	456	444	59	40	40	39	40	1.7	2.5	2.5	2.5	2.5
RAJ	2113	1076	903	835	796	103	55	51	47	53	1.0	1.8	2.0	2.1	1.9
BIH	1207	392	218	214	–	77	25	13	10	–	1.3	4.0	7.6	10.2	–
KER	6477	9716	9716	–	–	489	716	755	–	–	0.2	0.1	0.1	–	–
J&K	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
India	1194	714	529	461	459	60	37	30	26	27	1.7	2.7	3.3	3.8	3.7

NoteGCA Gross Cropped Area

Source Based on data on DM&I and compendium of selected economic indicators, CSO 2011

Table A.3 State-wise covered storage capacity—FCI, CWC and SWCs in India as on 1.1.2011

State	Covered storage capacity (In Lakh MT)				Percentage share of state in all India			
	FCI	CWC	SWCs	Total	FCI	CWC	SWCs	Total
Punjab	78.77	6.9	58.05	143.72	25.74	6.78	27.07	23.10
Andhra Pradesh	42.21	13.57	21.74	77.52	13.79	13.33	10.14	12.46
Uttar Pradesh	30.77	11.63	29.89	72.29	10.06	11.42	13.94	11.62
Maharashtra	21.19	15.95	11.97	49.11	6.93	15.66	5.58	7.89
Haryana	25.87	5.31	16.74	47.92	8.45	5.21	7.81	7.70
Madhya Pradesh	8.36	5.17	28.64	42.17	2.73	5.08	13.36	6.78
Sub total—(A)	207.17	58.53	167.03	432.73	67.70	57.48	77.90	69.55
Rajasthan	17.71	4.02	7.71	29.44	5.79	3.95	3.60	4.73
Karnataka	8.38	5.42	9.93	23.73	2.74	5.32	4.63	3.81
Tamil Nadu	9.85	6.33	6.36	22.54	3.22	6.22	2.97	3.62
West Bengal	11.01	6.5	2.16	19.67	3.60	6.38	1.01	3.16
Chhattisgarh	8.26	2.75	8.47	19.48	2.70	2.70	3.95	3.13
Gujarat	7.02	7.72	1.49	16.23	2.29	7.58	0.69	2.61
Orissa	6.44	3.73	4.14	14.31	2.10	3.66	1.93	2.30
Bihar	6.98	1.31	2.47	10.76	2.28	1.29	1.15	1.73
Kerala	5.37	1.22	1.98	8.57	1.75	1.20	0.92	1.38
Sub total—(B)	81.02	39.00	44.71	164.73	26.48	38.30	20.85	26.47
Assam	2.75	0.65	2.53	5.93	0.90	0.64	1.18	0.95
Delhi	3.67	1.51	–	5.18	1.20	1.48	–	0.83
Chandigarh	3.60	0.13	–	3.73	1.18	0.13	–	0.60
Uttarakhand	2.34	0.71	–	3.05	0.76	0.70	–	0.49
Jharkhand	1.29	0.35	–	1.64	0.42	0.34	–	0.26
Jammu & Kashmir	1.31	–	–	1.31	0.43	0.00	–	0.21
Tripura	0.52	0.24	–	0.76	0.17	0.24	–	0.12
Goa	0.15	0.41	–	0.56	0.05	0.40	–	0.09
Meghalaya	0.26	–	0.14	0.4	0.08	–	0.07	0.06
Himachal Pradesh	0.26	0.07	–	0.33	0.08	0.07	–	0.05
Nagaland	0.20	0.13	–	0.33	0.07	0.13	–	0.05
Mizoram	0.23	–	–	0.23	0.08	–	–	0.04
Arunachal Pradesh	0.22	–	–	0.22	0.07	–	–	0.04
Manipur	0.20	–	–	0.20	0.07	–	–	0.03
Sikkim	0.11	–	–	0.11	0.04	–	–	0.02
Sub total—(C)	17.80	4.30	2.67	24.77	5.82	4.22	1.25	3.98
All India (A + B + C)	305.99	101.83	214.41	622.23	100.00	100.00	100.00	100.00

Source Based on Annual Report 2010–2011, Department of Food and Public Distribution, Ministry of Consumer Affairs, Government of India

Table A.4 Covered storage/warehousing capacity available with the FCI/CWC/SWCs for foodgrains in India (in Lakh MT)

Years	Total storage capacity				Percentage share of owned capacity to total			
	FCI	CWC	SWCs	Total	FCI	CWC	SWCs	Total
1990	175.88	63.36	88.80	328.04	67.86	72.10	72.67	69.98
1991	195.92	64.77	91.79	352.48	61.23	72.90	71.88	66.15
1992	183.14	64.28	92.03	339.45	65.43	74.28	73.02	69.16
1993	180.18	64.02	90.74	334.94	67.57	75.77	76.46	71.54
1994	209.64	63.73	95.58	368.95	58.43	77.66	73.71	65.71
1995	226.79	69.13	100.04	395.96	54.22	73.28	72.59	62.19
1996	206.66	69.75	108.31	384.72	60.07	73.55	68.29	64.83
1997	194.76	71.19	109.50	375.45	63.79	72.95	70.29	67.42
1998	191.60	72.28	107.80	371.68	64.90	72.34	75.70	69.48
1999	191.56	73.48	113.89	378.93	65.27	73.01	69.88	68.16
2000	254.08	74.79	123.74	452.61	58.06	72.83	69.30	63.57
2001	314.46	83.91	149.05	547.42	47.34	66.88	57.71	53.16
2002	279.01	89.17	185.49	553.67	45.67	76.76	54.87	53.76
2003	265.87	91.14	199.31	556.32	48.21	83.51	76.04	63.96
2004	236.65	93.59	206.81	537.05	54.15	86.28	76.42	68.33
2005	233.70	101.87	195.20	530.77	55.24	82.83	66.00	64.49
2006	228.36	100.38	197.05	525.79	56.63	79.92	64.78	64.13
2007	222.83	102.20	192.20	517.23	58.08	79.35	66.02	65.23
2008	216.61	98.78	188.31	503.70	59.78	68.47	65.79	63.73
2009	230.91	105.25	190.91	527.07	56.16	64.23	64.12	60.65

Note FCI Food Cooperation of India, *CWC* Central Warehousing Cooperation, *SWCs* State Warehousing Co operations

Source Various Annual Reports of Department of Food and Public Distribution, Ministry of Consumer Affairs, Government of India

Table A.5 Farmers' awareness of MSP and procurement agencies

Sl No	States	Percent of farmers aware/not aware of MSP		Those aware of procurement agency
		Aware	Not aware	
1	Punjab	63.2	36.8	53.0
2	Haryana	66.7	33.3	43.6
3	Uttar Pradesh	33.3	66.7	21.5
4	Karnataka	29.2	70.8	23.4
5	Maharashtra	27.8	72.2	21.1
6	Gujarat	26.7	73.3	14.3
7	Andhra Pradesh	29.4	70.6	16.7
8	Tamil Nadu	49.9	50.1	39.0
9	Madhya Pradesh	29.4	70.6	19.7
10	Bihar	19.5	80.5	8.8
11	West Bengal	30.3	69.7	13.1
12	All-India	29.6	70.4	19.0

Source NSSO (2005), Some aspects of farming, 59th round, NSSO, P.A1

Table A.6 Statewise other marketing infrastructure in India

States	No. of grading labs	No. of grading labs per 1000		Other marketing facilities				Agmark nodes			Cooperatives	
		Km^2	MT	Food parks	Agro-exp zones	CE	Pack house	NO. Aug 2010	Per 1,000 Km^2	Per 1000 MT	No. of APMCS	No. of societies/ 1,000 Km^2 area
PUN	62	1.23	1.92	1	3	1	1	199	3.95	6.17	116	2.30
HAR	82	1.85	3.93	2	–	–	–	150	3.39	7.20	103	2.33
KER	52	1.34	7.80	4	2	2	–	92	2.37	13.80	548	14.10
GUJ	43	0.22	1.75	–	3	4	6	319	1.63	13.00	1717	8.76
TN	47	0.36	1.86	2	4	–	1	190	1.46	7.54	114	0.88
AP	44	0.17	1.12	1	5	–	6	334	1.28	8.50	383	1.47
MAH	68	0.22	2.17	7	8	4	89	346	1.12	11.06	1426	4.63
UP	153	0.64	2.09	5	4	4	–	257	1.08	3.51	258	1.08
KAR	44	0.23	1.62	4	4	1	4	171	0.89	6.29	505	2.63
MP	46	0.15	1.62	6	5	2	–	267	0.87	9.40	1097	3.56
HP	10	0.18	3.28	–	1	–	1	39	0.70	12.79	184	3.31
WB	24	0.27	0.58	8	6	1	–	56	0.63	1.34	287	3.23
BIH	12	0.13	0.39	1	3	–	2	58	0.62	1.91	376	3.99
ORI	6	0.04	0.33	1	1	–	–	91	0.58	4.99	171	1.10
CHH	2	0.01	0.20	1	–	–	–	73	0.54	7.18	5	0.04
RAJ	209	0.61	8.67	4	2	1	–	166	0.49	6.89	208	0.61
UTK	5	0.09	1.39	–	4	–	–	21	0.39	5.84	10	0.19
JHA	1	0.01	0.12	–	1	–	–	26	0.33	3.03	–	–
ASS	–	–	–	1	1	–	–	23	0.29	1.93	26	0.33
J&K	46	0.21	9.82	3	2	–	1	41	0.18	8.75	110	0.49
GOA	–	–	–	–	–	–	–	10	2.70	31.37	9	2.43
TRI	–	–	–	1	1	–	–	21	2.00	14.9	14	1.33

(continued)

Table A.6 (continued)

States	No. of grading labs	No. of grading labs per 1000		Other marketing facilities				Agmark nodes			Cooperatives	
		Km^2	MT	Food parks	Agro-exp zones	CE	Pack house	NO. Aug 2010	Per 1,000 Km^2	Per 1000 MT	No. of APMCS	No. of societies/ 1,000 Km^2 area
SIK	–	–	–	–	2	–	–	7	0.99	27.82	46	6.48
NAG	–	–	–	1	–	–	–	14	0.84	17.17	32	1.93
MEG	–	–	–	–	–	–	–	11	0.49	11.53	22	0.98
MIZ	–	–	–	1	–	–	–	9	0.43	24.34	4	0.19
MAN	–	–	–	2	–	–	–	5	0.22	5.37	17	0.76
ARP	–	–	–	–	–	–	–	15	0.18	29.74	4	0.05
All India		956	0.29	2.03	56	62	20	111	3011	0.92	6.4	7792
2.39												

NoteCE indicates Commodity Exchange

Source Compiled from the data obtained from DMI, MOA, GOI, Faridabad; and Jairath (2012)

A.1 Annexures

See Tables A.1, A.2, A.3, A.4, A.5, A.6.

References

Acharya SS (1988) Agricultural production, marketing and price policy in India. Mittal Publications, New Delhi
Acharya SS (1997) Agricultural price policy and development: some facts and emerging issues. Indian J Agric Econ 52(1):1–47
Athavale MC (2002) Impact of minimum support prices on agricultural economy in Madhya Pradesh. Agro Economic Research Centre, J.N.K.V.V, Jabalpur
Bhaduri A (2008) Predatory growth. Econ Political Weekly 43(16):10–14
Bhalla GS (1994) Economic liberalisation and Indian agriculture. ISID, New Delhi
Bhalla GS, Singh Gurmail (2012) Economic liberalisation and Indian agriculture—a district-level study. Sage Publications, New Delhi
Chenery H, Srinivasan TN (eds.) (1989) Handbook of development Economics. Elsevier Science Publishers, New York
Chand Ramesh, Kumar Parmod (2004) Determinants of capital Formation And agricultural growth: some new explorations. Econ Political Weekly 39(52):5611–5616
Dhanagare DN (1990) ShetkariSanghatana: the farmers' movement in Maharashtra—background and ideology. Soc Action 40(4):347–369
Dantwala ML (1981) Agricultural price policy: facts and issues. The Economic Times, 7 Feb 1981 (Reprinted in Dantwala (1996))
Dantwala ML (1996) Dilemmas of growth—the Indian experience. Sage Publications, New Delhi
Dar AK (1968) Domestic terms of trade and Economic development of India, 1952–1953 to 1964–1965 Cornell International agricultural development Bulletin 12. Itacha, New York
Deshpande RS (1996) Demand and supply of agricultural commodities—a review. Indian J Agric Econ 51(1&2):270–287
Deshpande RS (2003) Liberalisation, domestic price policy and agricultural growth. Indian J Agric Econ 58(3):636–648
Deshpande RS (2006) Agricultural policy in India: a policy matrix in a federal System. Research report, Agricultural Development and Rural Transformation Unit, Institute for Social and Economic Change, Bangalore
Deshpande RS (2008) Price policy and minimum support prices in a changing agricultural economy. In: Bhaumik Shankar (ed) Reforming Indian agriculture. Sage Publication, New Delhi
Deshpande RS, Nagesh Prabhu (2005) Farmers' distress: a proof beyond doubt. Econ Political Weekly XL(44–45):4663–4665
Deshpande RS, Prachitha J (2005) Building up of an efficient marketing system to obviate the need for a large scale state intervention. Research report, Agricultural Development and Rural Transformation Unit, Institute for Social and Economic Change, Bangalore
Deshpande RS, Raveendra Naik (2004) Moon in the mirror: farmers and minimum support prices in Karnataka, Monograph Series 7. Institute for Social and Economic Change, Bangalore
Ghosh Jiban Kumar (2001) Impact of minimum support prices on agricultural economy in West Bengal. Agro Economic Research Centre, Visva-Bharati, Santiniketan
Govt of India (1947-48), Parliamentary Debates, Govt of India, New Delhi.
Government of India (1965) Government resolution on the terms of reference of the agricultural prices. Agricultural Prices Commission, Ministry of Food and Agriculture, New Delhi

Govt of India (1966), Report of the Foodgrains Policy Committee, Ministry of Food, Agriculture, Community Development and Cooperation, Govt of India New Delhi.
Government of India (1976) Report of the national commission on agriculture. Ministry of Food and Agriculture and Irrigation, New Delhi
Government of India (1980) Report of the special expert committee on cost of production estimates, (Chairman: SR Sen). Department of Agriculture and Cooperation, Ministry of Agriculture, New Delhi
Government of India (1986) Agricultural price policy—A long term perspective. Ministry of Agriculture, New Delhi
Government of India (1990) Report of the expert committee for review of methodology of cost of production of crops (Chairman: CH HanumanthaRao). Department of Agriculture and Cooperation, Ministry of Agriculture, New Delhi
Government of India (1995) Report of the task force on terms of trade between agriculture and non-agricultural sectors. Ministry of Agriculture, Directorate of Economic and Statistics, Delhi
Government of India (2000) National agricultural policy. Department of Agriculture and Cooperation, Ministry of Agriculture, New Delhi
Government of India (2002) Report of the high level committee on long-term grain policy, (Chairman: Abhijit Sen) Department of Food and Public Distribution, Ministry of Consumer Affairs, Food and Public Distribution, New Delhi
Government of India (2004) Report of the national commission on farmers, (Chairman: Dr. M.S. Swaminathan). Ministry of Agriculture, New Delhi
Government of India (for various years) Reports of the commission on agricultural costs and prices (Earlier called as Agricultural Prices Commission). Ministry of Agriculture, Government of India, New Delhi
Jairath MS (2010) Agricultural marketing facilities in india- state wise analysis. A paper presented in annual general body meeting and conference on reforms in agricultural marketing and infrastructure, 12–14th September, at Silliguri and also accessed from website of National Council of State Agricultural Marketing Boards (COSAMB). http://www.cosamb.org/downloads.html
Jharwal SM (1999) Public distribution system in India: reassessed. Manak Publications, New Delhi
Kahlon AS, Tyagi DS (1980) Intersectoral terms of trade. Econ Political Weekly 15(52): A173–A184
Krishnaji N (1991) Agricultural price policy in Asia. Indian J Agric Econ XLVI(2):90–98
Krueger AO (1990) Government failures in development. J Econ Perspect 4(3):9–23
Malmberg Sue Simkin, Hawton Keith (1999) Suicides in Farmers. Br J Psychiatry 175:103–105
Mishra Srijit (2006) Suicides of farmers in Maharashtra. Indira Gandhi Institute of Development Research, Mumbai
Nadkarni MV (1987) Farmers' movements in India. Allied Publishes, New Delhi
Narayanamoorthy A (2012) Is the MSP Hike Justified? Financial Express, 3 July 2012
Nayyar D (2006) Economic growth in independent India: lumbering elephant or running Tiger? Econ Political Weekly 41(1):1–34
Raj K (1963) Farm supply response in India-Pakistan: a case study of the Punjab region. Econ J 73:232–236
Raj K (1967) Agricultural price policy and economic development in South-North. In: Southworth HM and Johnston BF (eds) Agricultural development and economic growth. Cornell University Press, Ithaca
Ratnam C, ChinnaRao B (2002) Impact of minimum support prices on agricultural economy in Andhra Pradesh. Agro Economic Research Centre, Andhra University, Visakhapatnam
Rao VM (2001) The making of agricultural price policy: a review of CACP reports. J Indian Sch Political Econ XIII(1):1–28
Rao VM, Deshpande RS (2002) Food security in drought-prone areas: a study in Karnataka. Econ Political Weekly 37(35):3677–3681

Sen, Abhijit and MS Bhatia (2004) State of Indian farmer a millennium study—cost of cultivation and farm income. vol 14 Academic Foundation, New Delhi

Shroff Sangeeta (2003) Impact of minimum support prices on agricultural economy in Maharashtra. Agro Economic Research Centre, Gokhale Institute of Politics and Economics, Pune

Singh JP (2005) Building up of an efficient marketing system to obviate the need for a large scale state intervention-A consolidated report. Agro Economic Research Centre, University of Delhi, Delhi

Singh Karam, Vatta Kamal, Kumar Sanjay (2002) Impact of minimum support prices on agricultural economy in Punjab. Agro Economic Research Centre, Punjab Agricultural University, Ludhiana

Sinha Ranjan Kumar (2002) Impact of minimum support Prices on agricultural economy in Bihar. Agro Economic Research Centre, Bhagalpur University, Bhagalpur

Singh SJ (2002) Impact of minimum support prices on agricultural economy Madhya Pradesh Jabalpur. Agro Economic Research Centre J.N.K.V.V, Jabalpur

Srinivasan TN (2000) Eight lecture on India's economic reforms. Oxford India Paperbacks, New Delhi

Swaminathan R (2002) Impact of minimum support prices on agricultural economy in Tamil Nadu. Agro Economic Research Centre, University of Madras, Chennai

Thamarajakshi R (1969) Intersectoral terms of trade and marketed surplus of agricultural produce, 1950–1951 to 1965–1966. Econ Political Weekly 4(26):

Tyagi DS (1990) Managing India's food economy. Sage Publications, New Delhi

Thippaiah P, Deshpande RS (1998) Analysis of market infrastructure, price and terms of trade: a case study of Karnataka. Indian J Agric Econ 53(3):359–369

Venkataramanan LS, Prahladachar M (1978) Study of cropping pattern changes in Andhra Pradesh during 1950–1975. Indian Econ Rev XIII(2):117–142

Chapter 8
Promoting Entrepreneurship Among Women in Agriculture

Nilabja Ghosh

Introduction

As India began to achieve unprecedented high rates of growth in the wake of liberalization, a concern for inclusiveness of the development process became more palpable. Among other disadvantaged groups, it was felt that women who constituted about half of the nation's population did not find adequate opportunity to participate in the growth process or reap its benefits. With broader concepts of 'well-being' and 'human development' finding their place in the economics literature (Nussbaum and Sen 1993; Dasgupta 2001; Sen 2006a, b), research, and policy started to focus on ways in which India's development can encompass all sections of the populace in coming times. In the later part of the 2000s decade, when the growth process itself began to encounter various challenges both from within the country and outside, the question of how agriculture can be made into an engine of growth also gained importance.

It is now widely recognized that policies which leave out a sector that engages more than half of the population and a large majority of the poor cannot generate growth that is sustainable (Sengupta 2013). The approach to agricultural development therefore broadened from its conventional focus on growth of output toward diversification and value addition to agricultural products. Agriculture began to be viewed more comprehensively as a source of income and purchasing power rather than merely a direct source of food security (Gulati et al. 2011). Economic policy in India favored liberalization of agricultural markets and for promoting agro-processing, food processing in particular. The inclusiveness implied in these approaches to development has become a major source of controversy in Indian political economy today as it is felt that only the more affluent sections and multinational companies will be benefited by this form of reform. Agro-processing is a large sector in Indian agriculture, but as of now a large part of

N. Ghosh (✉)
Institute of Economic Growth, Delhi 110007, India
e-mail: nilabjag@yahoo.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_8, © Springer India 2013

it is undertaken within the unorganized sector and any attempt to shift this activity toward the powerful market entities could hardly benefit the disadvantaged people in the country and may be far from equitable.

In this context, there is a case for looking at alternative ways in which the potentials of agriculture and the human resources contained in the agriculture sector can be utilized through the method of value addition in a broadened domain of agriculture so that smaller farms and the women in farm households can contribute to as well as benefit from the development process. In other word, a policy toward integrating agriculture with mainstream industry and the final consumer may be achieved by promoting a more holistic agriculture within which a part of the processing is also embedded with, and exploits the indigenous skills, modern training facilities, and entrepreneurial abilities of women in farm households. This approach would strengthen the linkages between India's growth and the development of human resources.

In this chapter, we address the issue of promoting entrepreneurship among women in rural India by opening up opportunities of processing agro-products and agro-inputs. Enterprise in this context deviates from conventional practices of self-employment among rural populace in that it connects with newer products and activities, involves risk and negotiation and it is perceived to be 'difficult'.[1] The merit of this approach lies in the easier availability of inputs within the sector and in the traditional skills that the women often possess in these arts. We assess the socio-economic situation of the farm women and review India's policy directions toward the economic empowerment of women in farm households. We also take a view of the status of entrepreneurship among farm women using unit level secondary data collected by National Sample Survey Organization (NSSO). The chapter is organized in the following way. The socio-economic stature of women in Indian economy is outlined in Sect. 2, followed by a historical account of Indian policy approach toward the development of women resources in Sect. 3. The role of agricultural extension and an assessment of a recent flagship program National Agricultural Technology Progress (NATP) are especially addressed. Sections 4 and 5 provide the analysis of household data to unveil the economic reality of the livelihoods and entrepreneurship of women especially belonging to farm households. Section 6 ends with a few concluding remarks.

Economic Stature of Rural Women in India

India is one of the few countries with 'missing women' syndrome (Klasen and Wink 2006) and a declining trend of the sex ratio of not only the total population

[1] An entrepreneur is a 'person who organizes and operates a business or businesses, taking on greater than normal financial risks in order to do so' is the definition given in oxford dictionary. Further, Oxford American dictionary describes enterprise as an undertaking that is 'difficult' and also defines an entrepreneur as a person who undertakes a 'commercial risk for profit'.

Table 8.1 Work participation rates (%) of rural male and female from census and NSSO data

	Census (Total)		Census (Main)			NSSO (PS + SS)		NSSO (PS)	
Years	Male	Female	Male	Female	Years	Male	Female	Male	Female
1981	53.8	23.2	52.6	16.0	1987–1988	53.9	32.3	51.7	24.5
1991	52.5	26.7	51.8	18.6	1993–1994	55.3	32.8	53.8	23.4
2001	52.1	30.8	44.3	16.6	1999–2000	53.1	29.9	52.2	23.1
2011					2004–2005	54.6	32.7	53.5	24.2
					2009–2010	54.7	26.1	53.7	20.2

Source Computed from Census 1981, 1991, 2001, 2011, Primary Census Abstract and NSSO, 2009–2010, NSSO, 1999–2000, NSSO 2001, Employment and Unemployment Situation in India
Note Census categorizes employment as 'main' and 'Subsidiary' status and NSSO by 'Usual Principal

but also of the under 6 years age group of population. Gender disparity is also manifest in educational attainments. Property right norms show a bias against females depriving them of all means of sustenance in case of death of or desertion by spouse (Agarwal 1994). With agriculture, having become a relatively low income yielding activity, off-farm opportunities are drawing rural men out of the villages, leaving women to undertake the burden of cultivation, and leading to the so-called 'feminization' of agriculture.

Work participation of women in India as recorded both by the National Sample Survey Office (NSSO 2010) and Government of India (1981, 1991, 2001, 2005, 2008), despite mutual discrepancies, is decidedly lower than that for men (Table 8.1) and remains confined to a few specific occupations some of which are extension of their domestic work. The fact that a correct picture of the extent and nature of the women's economic participation has remained elusive to data protocols (Krishnaraj 2005) has been even more disturbing. The evolution of many social ills like dowry can be directly associated with the degeneration of economic roles of women, their forced dependence on men and a lack of recognition of their role in household welfare and caregiving (Boserup 1970).

Table 8.2 however suggests that rural women are far more participative in economic life than urban women. In 1999–2000 female work participation rate (WPR) was nearly 30 % in rural India compared to 14 % in urban areas.

Table 8.2 Employment (%) in usual status of women in the two sectors in India

Workers Time	Rural 1987–1988	Urban	Rural 1993–1994	Urban	Rural 1999–2000	Urban	Rural 2004–2005	Urban	Rural 2009–2010	Urban
Workers (PS + SS)	32.3	15.2	32.8	15.5	29.9	13.9	32.7	16.6	26.1	13.8
Workers (PS)	24.5	11.8	23.4	12.1	23.1	11.7	24.2	13.5	20.2	11.9
In agriculture (PS)	82.5	21.8	84.7	19.3	84.1	14.6	81.4	14.7	78.9	11.8
Self employed(PS)	54.9	39.3	51.3	37.2	50.0	38.4	56.4	40.4	50.3	35.4
Casual labor(PS)	40.2	26.5	45.3	27.3	46.1	23.1	38.9	17.4	44.2	20.2

Note PS = Usual Principal status, SS = Usual Subsidiary status
Source NSSO 2001 (Employment and unemployment situation in India 1999–2000, 2004–2005, 2009–2010)

A significant part of the participation was in the subsidiary status and a large majority of about 84 % of the women workers were occupied in agriculture. Over time the WPR as well as the women's engagement in agriculture declined, but higher WPR of rural women over urban women is evident even in the 2009–2010 survey. Rural women are known to withdraw from not only the workforce but even from the labor force when seasonal scarcity of work sets in but in reality, they are known to face underemployment. The decline in women's WPR in the recent period is attributed to various social and cultural reasons including increased engagement in higher level education but poor quality and inadequacy of acceptable work opportunities need not be wished away.

Quality of Employment

Not only is the rural woman's work profile highly confined to farming, even for the women working within agriculture, it is severely limited to specific operations only. The lack of diversity of women's occupation is compounded by the fact that such operations are found to be difficult, monotonous, and repetitive and are described as drudgery. The All India Coordinated Research Project (AICRP) 1999 of the Ministry of Agriculture identified transplanting, weeding, cutting and uprooting, sowing and storage as common female activities of farm job. While rating the operations by their 'difficulty scores', these operations were assigned relatively high values. The activities are generally manual and many of them require awkward postures that produce body pain and injury that are sometimes known to cause reproductive problems. They are also not conducive to human resource development since scientific skills and decision-making have a small role in such jobs. Farm women have little association with negotiations in market and communication in public life and possibly there is little incentive for even education and skill.

Mechanization which took place during the period of the green revolution has made farm jobs easier but it is argued that the favorably affected jobs were mostly in the male domains. Although analyses also show that mechanization was gender neutral and created additional demand for labor through increased intensity of cultivation (Agarwal 1994), a biased societal reaction and a misplaced belief that women could not handle machines[2] could have been responsible for jobs even customarily done by women to pass into male domain when they became mechanized. The government's extension mechanism was largely answerable for this failure and in response to such criticisms, the *National Agricultural Technology Program* (NATP) and the Central sector scheme *Women in Agriculture* discussed in foregoing sections were designed to address this anomaly.

[2] This notion was shown to be untrue in other societies where women do operate and maintain machines (Wasnik 2005).

Table 8.3 Gender disparities in wage rates in agricultural operations usually done by women (Rs.)

Operation	Male		Female		Ratio (F/M) %	
	2004–2005	2007–2008	2004–2005	2007–2008	2004–2005	2007–2008
Sowing	66	79	46	57	69.7	72.2
Weeding	58	70	48	58	82.7	82.9
Transplanting	62	74	52	62	83.9	83.8
Harvesting	65	75	54	62	83.1	82.7
Winnowing	61	71	47	56	77.0	78.9
Threshing	64	73	50	59	78.0	80.8
Picking	62	72	45	58	72.6	80.6
Casual labor*	45	101	29	69	64.4	68.3
Public works*	48	98	38	86	79.2	87.8

Note *Wage rates in casual labor and public works are taken from NSSO surveys 1998 and 2010 respectively. Further wage rate in MGNRGEA public works are also reported by NSSO for 2010 as Rs. 90.93 for male and Rs. 87.20 for female

Source Employment and unemployment situation in India 1999–2000, 2009–2010, Indian Labor Journal Vol. 46 No 11 2005, Labor Bureau Government of India (2007–2008)

The structure of employment has visibly moved away from a self-employment base, (mostly family helpers) toward wage labor between 1993–1994 and 1999–2000 and in the most recent period 2004–2010. The shifts were accompanied by increased 'casualization', more patently in the recent period. As a result, wages earned by workers became a crucial determinant of their economic well-being (Hirway and Roy 1999) and movements of wages became important for women's economic uplift. It comes out uniformly that men's earnings have been consistently above those of women's in agriculture where women earned about 50 % or at the most 80 % of men's wage in 2004–2005 (Table 8.3) depending on the work and region.[3] Micro studies also bring out similar disparity in wage rate (Kaur and Goyal 1996; Sudha Rani et al. 1990). The gap has narrowed remarkably since then although the disparity is still marked.

[3] The gender disparity observed in farm wages is often attributed to the fact that men and women perform different kinds of operations. While it is true that jobs done solely by men such as plowing, irrigation, leveling, and transporting invariably command higher wages, it has been pointed out that even in those jobs that are done by both men and women or perhaps better done by women such as picking and transplanting, the wage rates given for women are less than men. The supply factor in the rural labor market even when viewed as moderately competitive could also be playing a significant part in creating the gulf. Household responsibilities make it difficult for women to migrate creating abundance in the rural labor market. The concern for immediate household sustenance sometimes makes their participation appear as distress labor.

Shifts in India's Policy on Women's Development

Articles 14, 16, and 39 of the Indian Constitution made provision for gender equality in all spheres of life and Article 15 even empowers the State to make official discrimination in favor of women. Although India's Five-Year plans had a place for the special needs of women, the inadequacy of the initial approach cannot be missed. No notion of development of women resources is visualized in the First Five-Year Plan (1951–1956) except for the formation of the Central Social Welfare Board (CSWB) and the intended *Mahila Mandals* as part of the Community development approach.[4] Clearly, there was no vision of women forming an enlightened workforce of the future. The Second Plan (1956–1961) coincided with serious reforms in agriculture. During that period the approach to rural development was integrated with *Intensive Area Development Program* (IADP); however, it is widely admitted that the subsequently Green Revolution patently suffered from a male bias.

The initial plans addressed women's welfare rather than development and seemed obsessed with issues of family planning. No landmark step was envisaged during the entire period up to the Fifth Plan to alter the gender order in the economy and develop women as human resources, though admittedly, the welfare approach by incorporating education and health of women and child care in its objectives had a degree of positive but indirect implication for the development of women.

The Fifth Plan period was important for two reasons. The submission of the Report of the *Committee on the status of women in India* (CSWI) brought the much needed consciousness of women as inputs for national development rather than as mere targets of welfare. It was realized that the Constitutional guarantees of equality would be meaningless unless women's rights for economic independence was acknowledged and their training in skills as contributors to the family and national economy improved. The new understanding was of momentous importance. The Fifth Plan also coincided with the International Women's Decade organized under the UN's behest.[5] In 1976, a National Plan of Action was prepared providing guidelines based on UN's World Plan. A Working Group on Employment of Women was setup as an exercise of the Sixth Plan.

Influenced by CSWI Report and two other Reports on Women in Agriculture and Rural development, the Sixth Plan (1980–1985), devoting a whole chapter on

[4] Today the precise role of the CSWB is under evaluation and there are proposals of restructuring the Board.

[5] The twentieth century has been witnessing landmark actions at the international level, gradually and steadily changing the gender order in society, bringing women toward the mainstream society and economy. Beginning with the UN's first World Conference on Women in Mexico in 1975 there were subsequent conferences in Copenhagen, Nairobi and Beijing. The UN decade for women (1976–1985) began an epoch of self-exploration among the nations on their internal policies. In 1990, the millennium development goals (MDG) further reinforced the resolve of including components relating to women's welfare and development.

women and development, for the first time shifted the approach toward women from welfare to development. Women's 'opportunities for independent employment and income' finally found a place in the Plans. The Seventh Plan (1985–1990) also made qualitative emphasis on inculcating confidence among women and generating awareness of their rights and privileges and deliberated on training them for economic activity. Other gender sensitive issues such as women's land and property rights (joint *pattas*), credit, marketing, and technology received consideration. Women's role in home and child care and in collecting fuel, fodder, and water was recognized. The Plan also identified beneficiary oriented programs IRDP, TRYSEM, DWCRA, and continued the externally funded program for Women in Agriculture initiated during the previous Plan. Domestic violence also received attention in the context of women's social position.

In 1985 Government of India constituted a separate Department in the Ministry of Human Resource Development to monitor and execute programs for women in other Ministries and Departments and a perspective Plan for Women (1988–2000) was made to provide a long term overall policy for Indian women. The National Education Policy gave a specific direction for the education of the girl child.

The Eighth Plan (1992–1997) was the first plan in post-reform era. Marking a departure from the past, this plan was indicative only as the role of the Planning Commission was subject to redefinition. The Eighth Plan focused on human resources and worked out an employment strategy based on opportunities of self-employment, seeking to improve the bargaining power of producer groups in the unorganized sectors. Vocational training for self-employment was an important plank in the strategy.

The Plan viewed women as equal partners in development and demonstrated gender sensitivity so that development programs did not bypass women. Training of women in soil conservation, dairy, social forestry, sericulture, horticulture, and poultry became popular. Also the formation of the *Self-help Groups* of women and training them for employment were milestone of this plan. The modification of agricultural extension to cover women beneficiaries was also a landmark initiative of the era. Thus, the Eighth Plan was crucial in many ways for setting the stage for future Plans to carry forward the task of empowering women in society and economy.

The Ninth Plan (1997–2002) placed women's empowerment among the primary objectives of national development. A *National Policy for Empowerment of Women* (NPEW) was adopted in 2001 which viewed women as agents of socio-economic change.[6] While the transition is yet in process, as men and women move toward equity, one looks forward to a time when the special emphasis on women

[6] The NPEW seeks to create an enabling environment for women to exercise their rights both within and outside their homes, reserve one-third of seats in Lok Sabha and State Legislative Assemblies, ensure that at least 30 % of funds and benefits flow to women in all development sectors by the *Women Component Plan* (WCP), to organize women into self-help groups and increase credit flow to women.

in particular may not be required.[7] The Tenth Plan also emphasized women' empowerment and employment generation and 30 % allocation for women in beneficiary oriented projects ensued.

The Eleventh Plan

Economic empowerment of women, treated as agents of development, gained considerable importance in the ongoing Eleventh Plan which aimed at improving allocation toward women-oriented schemes and making their implementation more effective. The Plan has a special focus on 'inclusion' which implies that Scheduled Caste and Scheduled Tribe and Muslim women would be the target beneficiaries. Gender budgeting and Gender outcome assessments are given high priority. Education, Training, and credit for women and 'leadership development for life, livelihood, and civic empowerment of minority women' received increased emphasis within the overall aim of gender 'mainstreaming'.

The Place of Farm Women in the Eleventh Plan

'Gender and Agriculture' is an issue especially addressed by the Eleventh Five-Year Plan. The Plan took note of the pivotal role played by women in agriculture as farmers, co-farmers, family labor and managers, and of women's extensive involvement in the production of major grains, millets, in land preparation, seed selection, seedling production, use of fertilizer and pesticide, sowing, manuring, weeding, transplanting, and in post-harvest procedures. Besides, women's intensive participation in livestock rearing and collection of forest produce and water are taken into account.

The policy of gender mainstreaming cutting across a number of Ministries and departments, is the key instrument in addressing women in agriculture in the Plans. Incorporated in the National Agricultural Policy of 2000 this means that women have to be part of all the agenda and can participate and benefit at par with male farmers by setting their own agenda, Empowerment remains the overarching goal with capacity building and input access being the two main instruments. The Eleventh Plan identified the thrust areas for agricultural development such as increase in productivity, regional dimension, livestock, coastal and inland fishery,

[7] The Tenth Plan has also emphasized women' empowerment alongside its focus on employment generation but unfortunately combines women with disadvantaged groups like children and disabled devoting a chapter on 'Women and Children' although it is important to realize that women's disadvantages are more man made than natural or biological and their constraints, concerns, and opportunities of the two groups are vastly different.

horticulture, plantation and nursery, sustainability concerns, water use and management and storage, and emphasized the role of women in all such activities.

Access to land is a further dimension in the Eleventh Plan. The ability of women to function effectively as farmers is constrained by their disadvantage due to the lack of property rights despite Constitutional provisions. Recognition of this right would also enable women to access credit and other resources and better contribute to productivity. Amendments of law made in 2005 effected correction to the biased and unjust land rights of women. Since 78 % of rural households own some land, access to even small plots of land can help women improve their bargaining power and use the same asset for productive purposes. The Plan also dwells on training and skill development of women. Programs organized by Universities at the door step should provide admission regardless of sex, age, and educational qualifications. Strengthening of backward and forward linkages of the agricultural sector with the non-agricultural sectors would provide gainful employment to women workforce. Training in nursery raising, horticulture, techniques in cereals productions, storage technique, seed support, biodiversity preservation, organic farming, and manure preparation are mentioned.

Educating and Training Women Entrepreneurs

Self-employment is a dominant economic occupation of Indian rural women according to employment survey results but rather than enterprise, it mostly signifies engagement of women in family farms. Yet, entrepreneurship for profit is recognized as a promising form of self-employment in India's development policy. Many rural development programs have emphasized the promotion of entrepreneurship among rural men and women.

Several public programs for industrialization and self-employment have been implemented in India for decades with disadvantaged sections, including women intended as priority beneficiaries. The *Integrated Rural Development Program* (IRDP), was possibly the largest and the most outstanding employment program for rural people. This centrally sponsored Scheme operating since 1980 was aimed mostly at poverty alleviation by providing self-employment to the rural poor through acquisition of productive assets or appropriate skills to generate additional income on a sustained basis to enable them to cross the poverty line. Assistance is provided to families below the poverty line, largely the small and marginal farmers, agricultural laborers and rural artisans. Associated with the IRDP were programs like the *Training of Rural Youth for Self-employment* (TRYSEM) for creating entrepreneurial competence and technical skills among the youth. *Development of Women and Children in Rural Areas* (DWCRA) and *Supply of Improved Tool Kits to Rural Artisans* (SITRA) are two other sub-programs associated with the IRDP.

The success of many of these programs has been far from satisfactory. The benefit failed to reach the poorest in the society, viability eluded the enterprises,

investment remained at sub-critical levels and above all, given the low honorariums and the lack of incentive, mobilization of the best training services proved to be challenging. Since the IRDP involved provision of subsidized credit, the lack of viability of the IRDP enterprises also tended to have an impact on the health and incentive of the banking system in the country. The IRDP however exists even today though known by *Swarnjayanti Gram Swarozgar Yojana* (SGSY). Perhaps a more noteworthy direction than the government initiatives was provided by the organization known as SEWA which is a non-government initiative. Acknowledged widely for its contribution to women's empowerment at the ground level (Chen et al. 2005, Chen et al. 2006, Nussbaum 1999, Nussbaum 2000) the Self Employed Women's Association of India (SEWA) is a 'trade union' for poor, self employed women workers in India that was founded[8] in 1972. SEWA members are women who earn a living through their own labor or small businesses. These women, constituting 93 % of the labor force, are unprotected by the common social benefits available to employees in the organized sector. Operating as a non-government agency, SEWA enjoys flexibilities as well as limitations but one innovative and redeeming feature is the presence of a research academy embedded in it that seeks to integrate the self employed women into the mainstream of the knowledge world, to rationally understand the implications of macro policy on welfare of the poor women and mitigate the adversities through grass root level solutions and alternative strategies. The World Bank has treated SEWA as a model to be replicated elsewhere.

Another successful program is the *Khadi and Village Industries Commission* (KVIC). Associated with India's freedom movement, it was created as a statutory body in 1957 to provide support and organization to village artisans. Unlike the IRDP, with its explicit emphasis on poverty alleviation, the KVIC has greater commercial orientation with the economic objective of 'producing saleable articles' finding a place among the social objectives. The definition of village industry restricts an industry to be located only in rural areas and to be producing goods and services with or without use of power but with a ceiling on the fixed capital used by an artisan. The products of the KVIC sector are marketed through its own sales outlets. Interestingly, 41 % of the beneficiaries are women. The KVIC organizes Exhibitions, Fairs, Seminars, and Public Education Programs to promote the scheme.

A program known as the *Support for Training and Employment program* (STEP) was launched[9] in 1987 after a situational review was made. The STEP program aims to increase the self-reliance and autonomy of' women by enhancing their productivity and enabling groups of women to take up income generation activities. It provides a package for skill up gradation through training; extension,

[8] A noted Gandhian and civil rights leader Dr Ela Bhatt is associated with the founding of SEWA. SEWAs' main office is located in Ahmedabad, Gujarat but it works in several states of India with a large number of members.

[9] This was made by the National Commission for Self Employed Women and Women in the Informal Sector.

and other inputs and the targets include poor, marginalized, and assetless women. The Scheme covers traditional sectors of employment, viz., Agriculture, Small Animal Husbandry, Dairying, Fisheries, Handlooms, Handicrafts, Khadi and Village Industries, and Sericulture. Two more sectors, namely, Social Forestry and Waste Land Development have been added later.

The Place of Women in India's Agricultural Extension

With the decline of the erstwhile method known as the *Training and Visit* or TAV system of extension of technology to agriculture, consensus moved toward a more holistic form of extension which viewed agriculture in a broader form and unlike in the TAV, gender became an essential component in extension. The Directorate of Extension of the Department of Agriculture and Cooperation (DAC), Ministry of Agriculture, in operation since 1958, is the nodal agency for effecting agricultural extension. Dissemination of technology and training with a focus on skill transfer are key methods in extension.

Today, the gender concern is addressed by mandating that 30 % of the resources on developmental programs and activities be for women farmers and women extension functionaries in major states. More interestingly, the Center also disseminates information on gender friendly tools/technologies in crop production, processing, post-harvest management, and other allied sector activities in a step toward correction of the past bias and flaws in the erstwhile extension system. The *National Gender Resources Centre in Agriculture* (NGRCA) located in the Pusa campus in Delhi under the same Directorate is responsible for adding a gender dimension to agricultural policies and programs and for rendering advisory services for gender mainstreaming in agriculture. Gender budgeting of the DAC is done at the NGRCA. A gender dimension is also added to the performance budget. In fact a separate chapter on gender perspectives in agriculture is inserted in the performance budget of the DAC.

Even when there is no specific beneficiary oriented program or any significant allocation earmarked for women farmers, attempt is made in each scheme implemented by the department to identify and promote the involvement of women in areas having better potential and scope. For example, although the schemes *Technology Mission on Cotton* and on *Farm Water Management* for increasing crop production in Eastern India are neither gender specific nor have any specific allocation for women, the states and the implementing agencies are encouraged to give preference to women farmers in the distribution of inputs, training, and demonstrations.

The Cooperative education and Development program for Women is a specific attempt at engineering the agricultural development to promote literacy, awareness, and for educating women to be organized into cooperatives. A number of agricultural implements and hand tools suitable for women farmers have been developed by the ICAR and in the Central Sector Scheme, the *Promotion and*

Strengthening of Agricultural Mechanization through Training, and 10 % of the funds were earmarked for women. A number of women were trained at both institutional and on-site venues and women farmers have actively participated in training programs.

Projects for Women in Agriculture

A few programs also explicitly address the women belonging to farm family. A majority of these were financed by external agencies including foreign countries. The first of all these was the DANIDA which was a Danish supported farm women program that began with the signing of *Women Youth Training and Extension Project* (WYTAP) in 1982 between the Government of India, the Government of Karnataka, and the Royal Danish Government. It worked in different phases up to the year 2005. The longest duration project in it was in Karnataka in which all districts except one were covered gradually. The *Tamil Nadu Women in Agriculture* (TANWA) concluded in 2003, the *Training and Extension for women in Agriculture* (TEWA) in Orissa concluded in 2003, and the *Madhya Pradesh Women in Agriculture* (MAPWA) concluded in 2005 were other parts of this holistic program that aimed at empowering women with better access to knowledge. Village-based training conferences, link workers training, inter-block and inter-district study tours for exposure, formation of '*Mahila Goshthi*' groups, and creation of drudgery saving technology were activities in the project.

The Dutch supported projects included training of *Women in Agriculture Andhra Pradesh* (ANTWA) between 1993 and 2007 and the training of Women in Agriculture Gujarat between 1997 and 2003. Signed between the government of India and Government of Netherlands, the projects promoted holistic development of small and marginal farmers and agricultural laborers by exposing them to relevant technologies and entrepreneurial opportunities.

The Government of India and UNDP signed an agreement in 1998 for an umbrella *Food Security Program* with six sub-programs. Under this, three women-specific sub-programs were implemented with a focus on drought prone and desertified agriculture in a holistic and sustainable way. The aims of the program included providing women's groups with access to crop land, encouraging tree-development in fallow land, storage and distribution of locally grown subsistence crops. Processing of food for easy cooking and marketing are also built into these sub-programs. Clearly, the provisions underline an integration of crop cultivation and entrepreneurship within the broad ambit of agriculture. Technology dissemination and exposure of women to best practices were also parts of the aim. State Departments of Agriculture, State Agricultural Universities, State Department of Women and Child Development, and NGOs as well as numerous Self-help groups were involved.

A central scheme of *Women in Agriculture* financed by the Government of India was launched on a pilot basis during the Eighth plan in one district each of

seven selected states Punjab, Haryana, Uttar Pradesh, Himachal Pradesh, Maharashtra, Kerala, and Rajasthan. It was extended to Ninth Plan and Tenth plan and the North East hill states were later included. The scheme aimed at motivating, mobilizing, and organizing women farmers to form groups so that agricultural support services, such as inputs, extension, and credit could be chanalized though the networks. A composite package of recurrent skill-based training was provided to enable them to adopt new technology. The women farmers also got training in managerial, entrepreneurial, and decision-making skills.

All these programs are reported to be highly successful. Income levels are found to have markedly increased, dependence on village money lenders reduced, and general empowerment effects significantly attained. Exposures to both intrastate and interstate practices and experiences in setting up stalls to sell vermin, compost, and food products were confidence inducing. The enterprises covered related directly to agriculture such as storage by operating 'Grain Bin', producing solar food processing, beekeeping diary, poultry. The main limitation was the inadequacy of implementation.

National Agricultural Technology Project

The limitation of the erstwhile extension system could be addressed by a bottom-up approach. The *National Agricultural Technology Project* (NATP), a World Bank-aided project, implemented by the ICAR and the Department of Agriculture and Co-operation (DAC) between November 1998 and 2005 has taken a group approach toward farmers sensitive to location-specific requirements. The institutional mechanism of NATP was the historic creation of autonomous, multi-partite, and participatory bodies known as *Agricultural Technology Management Agency* (ATMA). A major thrust was the technological empowerment of rural women for skill development and income generation, the maximum number of trainings being planned on post-harvest activities, and value addition mainly on fruits and vegetables, animal husbandry, and fishery under the NATP.

A country wide evaluation of programs (AERC, 2009–2010) for promoting women's entrepreneurship was conducted under the overall guidance and coordination of the present author based on survey results generated in the same project by a number of *Agro-economic research centres* (AERC) located in different Universities in the country. The states covered under the study were Andhra Pradesh, Punjab, Himachal Pradesh, Haryana, Uttaranchal (now Uttarakhand), Rajasthan, Bihar, Assam, and Uttar Pradesh. Nearly all the ventures studied were a result of the NATP program though not all of them were financially supported. The enterprises identified to be common were (i) nursery raising, growing fruits, vegetable, and pharmaceutical products using organic methods, poultry, bee keeping, and fishery under the primary activities though dairy was more common, (ii) Fruit and vegetable preservation including production of pickles, jam, jelly, *murabba,* and potato chips was an important activity under food processing along

with making of *papad and bari* from pulses, spice, and milk processing and making of health products like *dalia*, (iii) other processing activities included were quilt making, coir products, jute handicrafts, plate making, and weaving of baskets or plates and making *agarbatti* out of leaves, twigs, and other forest products and (iv) newer activities involved producing certain inputs for agriculture with environmental value such as vermi-composting based on locally collected worms, preparation of nutritious animal feed, mushroom culture, and crops production using use of bio-pesticides. Although the study covered only products related with agriculture, in practice NATP had a broader coverage of enterprises like doll making, candle making, and shampoo making.

The primary information-based study (Ghosh 2010a, b) found all indications of the economic potentials of the promoting enterprise among the farm women. Enterprises like making of *papad and badi* in Punjab were found extremely lucrative so much so that the diseconomy of scale became a serious limitation. A key strength of the enterprises studied was the cheap and easy availability of inputs that brings down cost considerably. Organizational innovations like cooperatives in dairy and more remarkably, working in groups to reduce fixed and overhead costs, were features that drew attention. Trainings, mostly arranged by *State Agricultural Universities* (SAU) and *Krishi Vigyan Kendras* (KVK), were useful especially in the new and eco-friendly ventures but were deemed redundant in other activities of traditional practice like dairy and food processing. Finance was sometimes availed from the NATP itself but otherwise from private sources only; there was no role of institutional credit. Access to land was an important factor in adoption of many enterprises. A serious limitation to the ventures was noted in the inadequacy of marketing channels. Enterprise was found not only to generate 'visible' monetary earnings of women, but also to help women acquire greater confidence, recognition, and awareness in society (Ghosh 2010a, b).

Livelihoods of the Farm Women

The livelihood of farm women and their status of entrepreneurship is studied by analyzing the household level data collected by NSSO 54th Round in the years 1998 (NSSO, 1999), extracting the data on farm households (cultivators and agricultural labor) only. The work force is assessed in relation to the female population of age above 15 years reported as 'adults' by NSSO. A 'self employed in own enterprises' group[10] (SEOE) of farm women is categorized on the basis of

[10] This distinction is made from the NSSO defined 'Self-employed' that includes also helpers in household enterprise who 'assisted the related person living in the household in running the enterprise'. This assures the inclusion of only individuals who have 'autonomy' and 'economic independence' for running the enterprise and receive 'remunerations determined by the sales or profits of goods and services produced'. For a unpaid family helper these specifications need not be satisfied.

Table 8.4 Participation in work of (adult) farm women in India

Principal status workforce (% population)	40.3
Usual (including subsidiary) status workforce (% population)	52.5
Employed in agriculture and forestry (% principal status workforce)	95.5
Casual labor and public works (% principal status workforce)	51.1
Unpaid family worker (% principal status workforce)	37.3
SEOE-own account, employer (% principal status workforce)	9.7
Subsidiary status SEOE (% subsidiary status workforce)	19.3
Entrepreneur (%Principal status Workforce)	3.5
Entrepreneur adjusted for tobacco workers (% principal status workforce)	3.0
Not beyond primary (% population)	86.8

Note SEOE = Self employed in own enterprises. Underemployed are employed by usual status but unemployed by weekly status
Source Computed

information provided on the status of employment to delineate women who are self employed but not as family help. Further the nature of enterprise is identified using five digit coding of the *National industrial Classification* (NIC) used in the reports. In particular, our focus is on the group of 'Entrepreneurs' who are specified as women in the SEOE category but who engage in activities other than the conventional cultivation of food crops. This group also includes tobacco workers but due to reasons discussed below entrepreneurs working in diversified activities excluding tobacco receive special attention in our study.

Farm women constituted about 47.9 % of Indian women and 48.9 % of farm population.[11] The work participation of farm household women is considerably more than for women from other rural households.[12] A large 95.5 % of the workforce is employed in agriculture and forestry in various capacities. Unpaid family workers constituted 37 % of workers and a larger share 51 % of the workers are employed as casual labor (Table 8.4). The SEOE constitute a considerable section of the subsidiary status workers but less than 10 % of those in the principal status. The education level of the farm woman was extremely low.

[11] This is an approximate figure only. A population of 230 million (1999–2000) is worked out based on NSSO data. There is usually a discrepancy between NSSO projected and Census population figures so that scholars usually project Census data to the survey period and make adjustment. This is not attempted here.

[12] It is also seen that even among non-workers 41% perform jobs that could be potentially priced if market is created.

Entrepreneurship

The Box below gives a summary picture of the diversified activities of enterprises in which the farm women's presence is observed.[13] The category 'processing'[14] covers a large number of activities all of which convey value addition to agriculture. Tobacco enterprise covers two different activities of *'bidi*' production or the making of the *'bidi'* and collection of the raw material, i.e., '*tendu*' leaves from forests. In particular, although the size of this sector is equivalent to processing sector taken together, our interest lies in activities of enterprise that exclude tobacco, because the viability of this historic self-employment base of women in India is questionable in the current context.[15]

[13] Tobacco, perfumery (including agarbatti) and craft are the activities that are found to have high share of women entrepreneurs (96, 100 and 70% respectively).

[14] The processing group includes (i) *food processing* activities like Flour milling, Dal milling, Rice milling, Processing and grinding grains, vegetable milling, flour of meal of dried leguminous vegetables of roots and tubers of edible nuts, breakfast foods by swelling and roasting cereal grains, starch and sago products, glucose syrup, gluten, corn oil, animal feed, bakery products, sugar, gur, khandsari, cocoa, chocolate, confectionery, sweetmeats, macaroni, noodles, tea, coffee, edible nuts, malted and infant foods, spices, papads, vitaminized high protein flour, dried dal and cereals, Manufacture of Beverages (non-alcoholic); (ii) *Textile-based processing* such as Spinning and weaving of cotton fiber, of silk fiber, wool, other animal hair, manmade fibers, Durries, druggets, rugs, carpets, rugs, coverings of jute, mesta, coir, other floor coverings of textiles, sunhemp, cordage, rope, tine, netting, thread, jute rope and cordage, coir rope and cordage, mesta, nets, Embroidery, laces, fringes, zari, ornamental trimmings, linolium, mantles, canvas goods, sanitary towels and tampons, metalized yarns, gimped yarn, rubber, water proof textiles, knitted, crocheted fabrics of cotton, woolen and synthetic substances, textile garments and accessories, raincoats, sheetings, hats, caps of waterproof textiles, leather apparels and (iii) *Crafts* like Basketry, grain bins, bamboo and reeds works, wooden containers, canes, rattan, bamboo willow, grass, leaves, wooden, industrial goods, cork products, bamboo and cane articles, and fixtures thatching from reeds, grass, broom sticks, wooden agricultural implement, shopping bags and ornamental boxes, costume articles, trays, table-lamps, fancy baskets, table mats, vessel holders.

[15] Bidi industry spanning most states in India may be as old as 119 years (since 1887) in India. A number of activities such as procuring Tendu leaves, cutting leaves to specific sizes, filling with tobacco, rolling and tying with yarns and curing in oven, and selling make up the major functions in the industry. Factory production of bidi is known for its sub-human conditions but processing mostly has a home-based and unorganized character managed by contractors. The flexibility of a home-based avenue of earning possibly attracts women (and even children) constituting an estimated 68% of total workers encompassing the whole industry. The pitiable conditions of working in this industry, (SEWA 2005, Varma and Rehman 2005) as also the decreasing production and consumption of beedi/tobacco, known to cause 3 million deaths per years in the world and facing curbs from government policies raise the need for suitable alternative livelihoods for women.

Box: Enterprises with Women's presence

Agricultural (Primary): Growing roots, tubers, cones, forage plants, etc., fruits and vegetables; seeds, nursery products, spices, nuts, animals, and fishery; agro-services, Forestry, forest services and collection. These exclude conventional crop-based activities.

Share in workforce: 1.9 %

Processing (PRC): Food processing; Textiles; Leather; Wood; Paper. Share of workforce: 0.4 %

Tobacco: Tobacco products (Bidi, cigar, etc.) manufacture and collecting Tendu leaves.

Share of workforce: 0.5 %

Trade: Retail Trade

Share of workforce: 0.2 %

Services: Hotel; health; recreation; education; washing, hairdressing, etc., finance and real Estate.

Share of workforce: 0.2 %

All Enterprises (Entrep)

Share of workforce: 3.5

All enterprises (Entrep-adj) except Tobacco

Share of workforce: 3.0 %

The gender share is cent percent in *agarbatti* (perfumery), i.e., nearly all entrepreneurs are women, particularly high in tobacco and lowest in food processing. Tobacco processing is an important constituent of women's enterprises but although categorized as such in this study it is useful to recognize that the nature of the business lies somewhere between entrepreneurship and wage labor as the work is assigned by the purchaser so that the element of risk is reduced. The flexibility of a home-based avenue of earning along with assured income possibly attracts women (and even children) constituting an estimated 68 % of total workers encompassing the whole industry. The Textile and Crafts activities among the entrepreneurs usually involve manual work. These jobs were also traditionally done by women in household often as domestic chore but may have fairly good market prospects.

Socio-Economic Profiles of Entrepreneurs

The socio-economic profiles of selected groups of households by women's entrepreneurship are presented in Table 8.5. Households with women in primary activities (growing commercial crops, raising animals, etc.) turn out to be socially

Table 8.5 Socio-economic profile of women in entrepreneur from farm households

Attributes	Entrepreneur	Tobacco	Primary	Processing	Trade and services
Household characteristics (% of sample households)					
Average age of entrepreneur (Years)	38.2	29.3	41.1	36.5	41.6
Land possessed (ha)	1.08	0.40	1.40	0.82	0.89
SC/ST (% Households)	31.0	33.5	28.0	41.3	36.5
Female headed (%)	15.4	9.8	20.5	9.0	10.7
Member not beyond primary education (%)	80.3	87.5	92.8	86.1	90.0

Notes Entrepreneur excludes tobacco workers

and economically relatively privileged. The average farm size at 1.4 hectares is the largest and the proportion of disadvantaged sections (scheduled castes and scheduled tribes) is least in this group but these women have less education and are more likely to belong to female-headed households. In contrast, average holding of the women in processing falls short of a hectare and it is worse among tobacco workers. The proportion of backward sections is also relatively higher in these groups. The processing entrepreneurs followed by the tobacco workers however seem to be more educated than the others which perhaps reflect their probable links with the formal sector in the economy.[16] The average age of the entrepreneur is over 38 years. This is more than the average of a typical farm woman worker estimated at 35.9 years. It is also worth noting that entrepreneurs in *beedi* making and in processing activities are relatively younger compared to their counterparts in primary activities and in trade. Perhaps the predominantly home-bound nature of these activities draw the younger women despite their greater household responsibility while the relatively senior women indulge in outdoor jobs like farm work and selling products in the market, but given the relatively weaker status of the households observed in these specific groups, their compulsion to earn may also be stronger.

Structure of the Enterprise

Table 8.6 presents fairly archaic organizational profiles of different enterprises including two major sub-sectors of processing enterprises, namely food processing and textiles. Less than 7 % of the enterprises have any system of written accounts. Only tobacco enterprises, usually tied up with more organized units, record a limited extent of accounting practice. Most of the enterprises run primarily on family labor

[16] Outsourcing of work by formal processors is not uncommon and contracts are most common in *beedi making*.

Table 8.6 Modernity of enterprises of farmwomen

Modernization indicators	Entrepreneur	Tobacco	Processing (total)	Food processing	Textiles
Employer (%)	0.7	0	0.8	0	0
Worker < 6 persons	88.6	95.3	83.1	80	80.4
Keep accounts	6.7	13.6	0.1	0	0.3
Works from home	68.6	88.9	73.4	61.7	71.9
No fixed place of work	1.2	0	0	0	0
Electricity use	–	–	2.5	9.7	0.8

Note Electricity use is for manufacturing only
Entrepreneur excludes tobacco workers

with no hiring practice. Also nearly 70 % of the enterprises operate from home with no separate office or work space. The picture is especially dark for manufacturing particularly so for food processing and textile enterprises where family members are widely found to participate. The processing organization, which is most akin to manufacturing in the organized sector, is far from modern among the women entrepreneurs. Majority of the groups reported in the tables however work at fixed places though the fixed place of work is generally their home.

Spatial Distribution of Entrepreneurship

Although intuition suggests that economic development would draw more women to participate in workforce by creating employment, the indications observed are quite the reverse. That affluence reinforces patriarchy, a hypothesis that is often used to explain differences in social norms that relate women's economic status with family honor among various classes, castes, and most notably between northern states and other states in India (Chen 1995). The lack of education, property rights, and the role of the plow culture (Agarwal 1994, Boserup 1970) are other possible but related explanations why rural women are drawn to the labor market by sheer distress rather than the urge for development.

The analysis of the data on farm households show that the agriculturally more developed north Indian states Punjab and Haryana record the lowest WPR of farm women. The estimated WPR is as low as 4 % in Punjab and 5.5 % in Haryana which is far less than the estimated 40 % at the all India level seen in Table 8.4. However, the same states rank high in women entrepreneurship measured by the ratio of entrepreneurs to women workers (Fig. 8.1a). In contrast, southern states Andhra Pradesh and Karnataka with WPR at 68 and 57 % lead in women's work participation but trail far behind in entrepreneurship. Northeastern states show strong tendencies for women entrepreneurship. States like Tamilnadu in the south and West Bengal in the east have moderately high tendency for entrepreneurship at around 5 % compared to the estimated 3 % at the all India level (see Box) but as seen in Fig. 8.1b the activity of enterprise is different from the northern states.

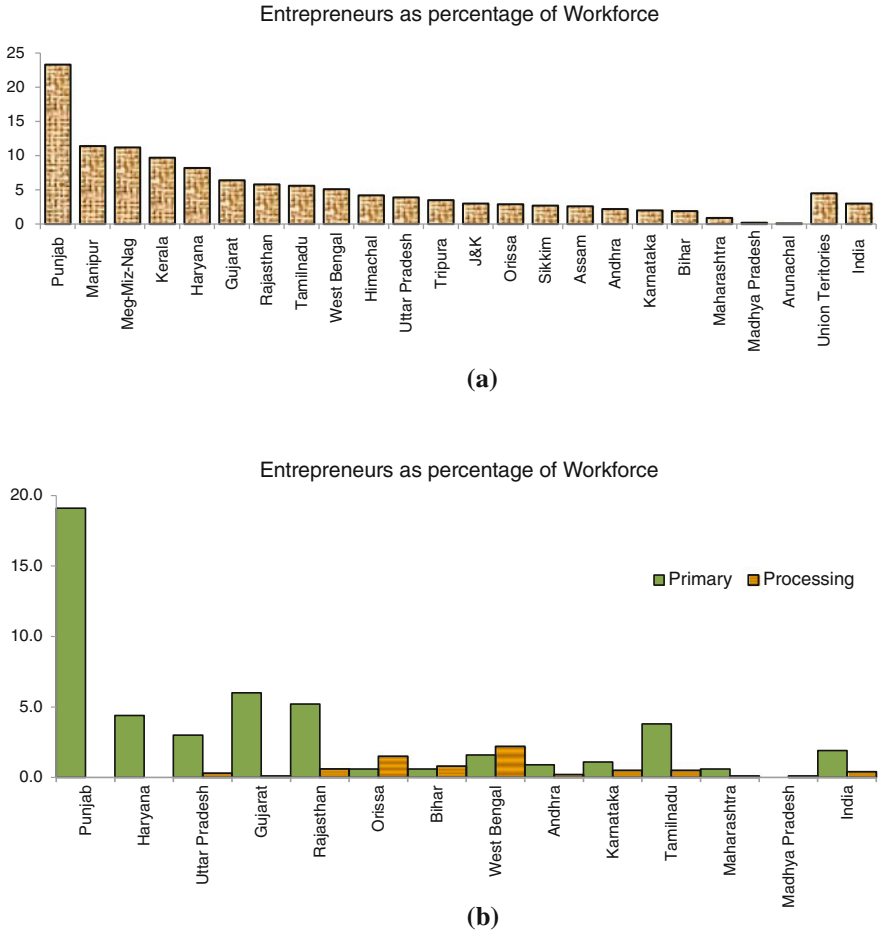


Fig. 8.1 **a** Tendency for entrepreneurship of farm women in Indian states. **b** Tendency for entrepreneurship in primary and processing activities of farm women in Indian states

The activity-wise spatial representation of entrepreneurship (Fig. 8.1b) signifies specific preference among the farm women in the two north Indian states for primary-based enterprises. Other neighboring states like Uttar Pradesh, Gujarat, and Rajasthan too share the same trait. The observation can be related to many factors like the economic strength of the region in agriculture and allied service, nature of extension services availed, easier access to land, and credit and family specialization in farming. The eastern states of West Bengal, Bihar, Orissa, Assam, and Tripura and to an extent the southern states too are in higher positions when ranked by processing. The active presence and priorities of self-help groups and supporting government and non-government agencies may be possibly associated with this specificity but more evidence is required for a firm conclusion.

In industrialized states like Maharashtra and Gujarat the tendency toward processing based entrepreneurship is absent and work participation as well as entrepreneurship tendency is less visible in Madhya Pradesh.

Constraints Due to Under-Development

Low quality of employment, discrimination, poor wages, and immobility imposed by urban housing inadequacies may be deterring rural women from joining the economic workforce in India's growth process. Women in the Indian society are also said to prefer 'home-based activities' and a role of household 'care givers'. The rigidity in the nature of their work participation and household ties is however reinforced by the antiquated technology with which domestic chores are conducted, the lack of adequate power supply, lack of good transportation, and assurance of safety that limit the mobility of women. Meager access to land is also a severe limitation for farm women to indulge in enterprise.

Agriculture, viewed merely as growing of crops is becoming un-remunerative. However, considered holistically to encompass diversification, processing, preservation, and other agro-services, it holds enormous potential for not only empowering women in their households and communities but also in giving a significant place to women in development and growth. A more balanced approach for policy makers could be to promote policies for industrialization such that part of the agro-processing activities is passed on to farm women by encouraging outsourcing through fiscal incentives, mandates, and certification.

References

Agarwal B (1994) A field of one's own. Cambridge University Press, Cambridge

Agro-economic Research Centres (Reports submitted by AERCs to Ministry of Agriculture, Government of India 2009–2010) Viable Entrepreneurial Trades of Women in Agriculture (Andhra Pradesh, Bihar, Haryana, Rajasthan, Punjab, Assam, Himachal Pradesh)

Boserup E (1970) Women's role in economic development. George Allen and Unwin, London

Chen M (1995) A Matter of Survival in women, culture and development. In: Nussbaum M and Glover J, Oxford

Chen MA, Khurana R, Mirani N (2005) Towards economic freedom: the impact of SEWA. Sewa Academy of Shree SEWA Mahila Trust, Krishna Bhuvan, Ellisbridge, Ahmedabad

Chen MA, Mirani N, Parikh M (2006) Self-employed women: a profile of SEWA's membership. SEWA Publication, Ahmedabad

Dasgupta P (2001) Human well-being and the natural environment. Oxford University Press, Oxford

Sen A (2006a) Development as capability expansion. In: Agarwal B, Humphries J, Robeyns I (eds) Capabilities freedom and equality. Oxford University Press, Oxford

Sen A (2006b) Gender and cooperative conflicts. In: Agarwal B, Humphries J, Robeyns I (eds) Capabilities freedom and equality. Oxford University Press, Oxford

Ghosh N (2010a) Economic empowerment of rural farm women through Agro-based entrepreneurial ventures. Report submitted to the Ministry of Agriculture, Government of India from Institute of Economic Growth, Delhi

Ghosh N (2010b) Organisational forms and market intermediations: a study on rural women entrepreneurs in India. Indian J Labor Econ Q J Indian Soc Labor Econ 53(2):381–395. New Delhi, ISSN 0971-7927, ZDB-ID 860762x

Gulati A, Ganguli K, Shreedhar G (2011) Food and nutritional security in India. International Food Policy Research Institute, New Delhi

Nussbaum M (1999) In defense of universal values, occasional paper series, women and human development: the fifth annual hesburgh lectures on ethics and public policy. University of Notredame

Nussbaum M (2000) Women's capabilities and social justice. J Hum Dev 1(2)

Nussbaum M, Sen A (1993) The quality of life. Clarendon Press, Oxford

Sengupta A (2013) Mainstreaming small farms: a new approach needed. In: Ghosh N, Sekhar CSC (eds) The future of Indian agriculture: technology and institutions. Academic Foundation, New Delhi

Klasen S, Wink C (2006) Missing women: revisiting the debate. In: Agarwal B, Humphries J, Robeyns I (eds) Capabilities freedom and equality. Oxford University Press, Oxford

Krishnaraj M (2005) Food security: how and for whom? Econ Polit Wkly 40(25):2508–2512 (18–24 Jun 2005)

Hirway I, Roy AK (1999) Women in Rural Economy: the case of India. Indian J Agric Econ 54(3):251–271

Kaur S, Goyal SK (1996) A study of wage discrimination against Women Agricultural Laborers. Agric Situat India 53(4):225–228

Government of India (GOI, 2005, 2008) Employment and unemployment situation in India. National Sample Survey Organization

Government of India (GOI, 1981, 1991, 2001) Census of India. http://censusindia.gov.in/

Government of India (GOI, 2008) Labor Bureau. http://laborbureau.nic.in/WR%20Rural%20India%202k7-8%20Contents.htm

NSSO (2010) Employment and unemployment in India, report no. 531 (64/10). National Sample Survey Office, Ministry of Statistics & Programme Implementation, Government of India

SEWA (2005) Hands that roll bidi: a socio-economic survey of bidi rollers from Murshidabad district. Sewa Academy, Shree SEWA Mahila Trust, Krishna Bhuvan, Ellisbridge, Ahmedabad

Sudha Rani P, Raju VT, Raghuram P, Naidu GM (1990) Wage differentials and factors governing employment of women in agriculture. Agric Situat India 45(4):249–252

Varma U, Rehman MM (2005) Tobacco, tendu leaf and beedi workers in India. Shipra Publications, Delhi

Wasnik KP (2005) Women in agriculture : strategy for socio-economic empowerment. Kalpaz Publications

Chapter 9
Construction and Engineering Industry in India in PPP Regime with a Special Focus on Airport Development

Bhaskar Chatterjee and Sarmila Banerjee

Introduction

Till the early 1990s, the general understanding that guided economic policy framework and planning in India hinged on the paradigm that the Government is mostly responsible for the provision of public goods/services which are by nature indivisible and non-excludable and thus cannot be priced by the market mechanism. The liberalization and globalization process that was initiated in the early 1990s catalyzed a paradigm shift in this perception. Rapid growth brought forth a manifold increase in demand for social/civil infrastructure. The Government's financial status and delivery capabilities were grossly inadequate to meet this demand. Moreover, the delivery mechanism itself was functioning quite inefficiently in the hand of the public sector authorities. So, from the demand side there was sudden expansion in scale and from the supply side the prevailing institutional arrangement was proved to be insufficient. It was in this backdrop that the government invited private participation in the public domain of social infrastructure. The financial as well as circumstantial rationale extended to justify such a change in paradigm in India has basically hinged on three crucial premises: (i) given the limited wherewithal of the Government in the face of burgeoning budget deficits, the rapid demand for social/civil infrastructure of a growing economy can only be met through active investment and participation of the private sector; (ii) with the involvement of private capital under global influences there would be constant monitoring of sovereign credit worthiness by international financial institutions and global credit rating agencies; and (iii) PPP model would raise efficiency of both resource usage as well as the speed of delivery helping the economy to

B. Chatterjee
Larsen & Tubro Limited, Chennai, India
e-mail: bhasx68@gmail.com

S. Banerjee (✉)
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: sarmila_b@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_9,

sustain rapid growth. The optimism regarding this arrangement derived support from the successful adoption of this model in Europe in the 1960s, in Latin America in the mid-1980s, and more recently in Africa and East Asia.

In 1998, under the National Highway Development Program (NHDP), the Government of India has initiated the public–private-partnership (PPP) approach which was aggressively extended later to other sectors. The first major diversification out of the road sector was in the airport sector with privatization of a number of major airports, viz., Bangalore, Hyderabad, Delhi, Mumbai involving around Rs.19,000 crores worth of investments. The airport initiative was followed by a deluge of PPP projects in power and urban infrastructure by the Central Government. The State Governments subsequently entered the fray and extended the PPP initiative to social sectors like health, education, etc. Available data till end December 2011 show that Rs.645,000 crore of investments have been done in India through the PPP initiative, taking into consideration both State as well as Central Government projects (Ministry of Finance, Government of India).

The PPP initiative brought in a new set of stakeholders and led to a significant change in the structure of the delivery market. Traditionally, the Indian construction market was dominated by a few strong contractors doing traditional engineering, procurement, and construction (EPC) business on a turnkey[1] basis. With the advent of the PPP model a new set of market participants emerged while the existing ones evolved into new forms. First, pure developers/concessionaires[2] entered the markets; second, many of the traditional EPC firms entered into the concession business looking for leveraging the advantages of the concession–construction integration. Third, the consortia approach was increasingly adopted and a number of global players entered the scene. Now the infrastructure project is conceived as an integrated project that would provide utility against infrastructure support and simultaneously provide some other direct utility bearing market goods as a part of a package deal. For example, in providing airport infrastructure, private players are enjoying the rent from duty free shops, land development, and provision of different hospitality services. Infrastructure is considered as a bundle of assorted services where one has to pay for enjoying the benefit of each service. Thus, with the adoption of PPP model, the non-divisible and non-excludable properties of traditionally defined *public goods* broke down and the role of Government changed from a direct provider to a *facilitator/regulator* for ensuring adequate supply. Services from projects developed under the PPP model are priced according to the market principle and are excludable in nature. Such projects are developed by the private sector with the objective of maximizing return on capital employed over a given horizon rather than the traditional norms of social welfare maximization. However, the projects

[1] In a turnkey project a private contractor is selected through a bidding process. The contractor designs and builds a facility for a fixed fee (rate or total cost) which is one of the key criteria in selecting the winning bid. The private party bears the risks involved in design and construction phases, his scale of involvement is low and for a short period of time.

[2] Firms that obtain PPP projects get it financed and constructed and enjoy profits from expected income over the period of the concession.

would still be under public ownership and there would be a compatible mixing of social cost–benefit analysis with private [3] analysis. It is expected that this new arrangement would relax pressure on Government budget and release resources for other Government directed programs, would raise efficiency of resource usage and delivery, connect the future of the construction and engineering activities with financial stability of the economy through developers/concessionaire's channel, and help the economy of India to earn her demographic dividend from the world market (Bloom and Williamson 1998).

Given such presuppositions for adoption of PPP model and the corresponding change in the delivery structure of the private players, successful implementation of such model in India would be contingent on a number of important factors:

(i) Adoption/creation of a balanced regulatory environment for right and equitable sharing of risk/profit amongst the major stakeholders i.e., the Private Sector, the Government, and the ultimate Consumers;
(ii) Arrangement/Availability of finances for such projects;
(iii) Capability of the domestic engineering and construction industry to deliver quality in such large-scale complex projects;
(iv) Long-term viability of PPP as a business model for corporate players given the high degree of leverage and uncertainty associated with the projected income streams;
(v) Availability of feasible *exit routes* for private parties, given their long-term exposure to such investments.

Given this background an attempt is made to analyze the organizational strength of PPP institution along with its associated risks and vulnerability in the light of other developing country experiences and the prevailing situation in India. Section 2 discusses the definition and alternative structures of PPP model, Sect. 3 reports the experience of the developing countries in the face of global integration, Sect. 4 talks about the state of PPP projects in India with a focus on viability in terms of financial, institutional, and regulatory challenges faced by them, Sect. 5 presents a case study of Delhi Airport, Sect. 6 assesses the regulatory practices, and finally, Sect. 7 presents an overall assessment. D="Q2" Text="Please check that the Sect. 7 is correct, and provide link amend if necessary ."

PPP: Definition, Structure, and Operation

The term public–private partnership (PPP) does not have a legal meaning and can be used to describe a wide variety of arrangements where the two parties share rights and responsibilities during the duration of the contract, i.e., distribution/sharing of risk and ownership (Edward et al. 2011). PPP is not privatization. There

[3] Strength-Weakness-Opportunities-Threat (SWOT).

is a clear difference between these two forms of private sector engagement: privatization involves the permanent transfer of a previously publicly owned asset to the private sector, whereas a PPP necessarily involves a continuing role for the public sector as a "partner" in an ongoing relationship with the private sector. Most importantly, under a PPP, accountability for provision of the service is clearly in the hands of the public sector, while under privatization, immediate accountability for providing the service lies with the private provider. In a PPP arrangement, the private sector party usually agrees to undertake the following:

(i) Design and build, expand, or upgrade the public sector infrastructure;
(ii) Assume substantial financial, technical, and operational risks;
(iii) Receive a financial return through payments over the life of the contract either from the users or from the public sector or from both;
(iv) Usually return the infrastructure to public sector ownership at the end of the contract.

The nature of the PPP is determined by what rights, obligations, and risks are assumed by the public or private parties within the partnership (Delmon 2010). In accordance with the share of responsibility of the private partner different contractual arrangements like O&M contract, lease/affermage contract, concession, joint venture and divestiture are observed in practice (Fig. 9.1).

(i) *O&M Contract*. This is mostly known as the management contract. Here the public authority finances and builds a project and the private entities provide operation and maintenance services against a fee for a specified period of time. The renewal of the contract is contingent on the quality of service provided. Popular examples in India are the management of municipal solid waste, that of Sulabh Complex, etc.
(ii) *Lease/Affermage*. The private entity builds/refurbishes and operates a service where the public authority or grantor finances any major capital expenditure. The private entity generally collects tariff from the consumers directly. The difference between an affermage and a lease contract is rather technical. Under a lease, the operator retains revenue collected from customers/users of the facility and makes a specified lease fee payment to the contracting authority.

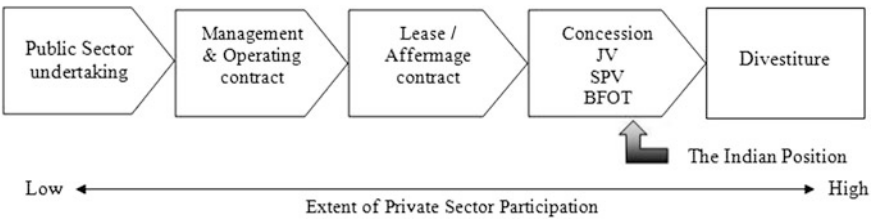


Fig. 9.1 The private ownership continuum model of PPP

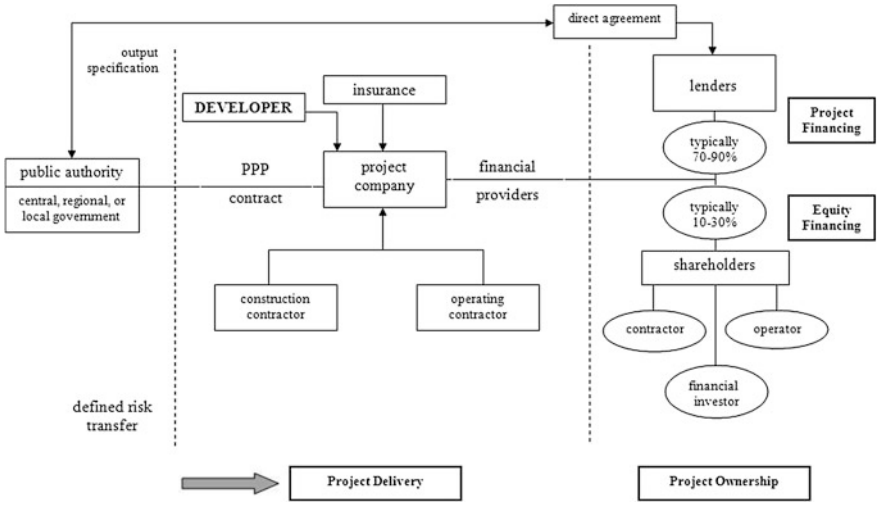


Fig. 9.2 Typical delivery structure of a DBFOT Project (*Source* National Highway Authority of India (NHAI))

Under an affermage, the operator and the contracting authority share revenue on a pre-specified ratio. So, in affermage contract the risk is shared between the contracting parties while the lease contract involves a fixed fee and, therefore, no risk sharing (Fig. 9.2).

Concession. In this form of PPP, the government defines and grants specific rights to an entity to build and operate a facility for a fixed period of time after which the ownership reverts back to the public sector. Here the private partner finances, builds, operates, and delivers a service in a leased-in-Brownfield[4]/Greenfield and collects user-fee directly from the consumers.[5] Terms such as BOT (build, operate, and transfer), BROT (build, rehabilitate, operate, and transfer), DBFOT (design, build, finance, operate, and transfer), DCMFT (design, construct, manage, finance, and transfer) are often used to describe such schemes. In concessions, payments can take place both ways: concessionaire pays to government for the concession rights and the government may also pay the concessionaire to meet certain specified conditions laid down in the agreement to make projects commercially viable and/or to reduce the level of commercial risk undertaken by the private entity, e.g., capital grants or other financial incentives. Here the risk share and the volume of investment of the private partner are large. Often the concessionaire is required to establish a special purpose vehicle (SPV) for

[4] BROWNFIELD Project refers to the modernization, extension, expansion, and remodeling of the existing projects; i.e., to start a project based on prior work or to rebuild (engineer) a product from an existing one. In contrast a GREENFIELD project means to start a project without the need to consider any prior work.

[5] This mode of cost realization is known as the User-fee-based approach.

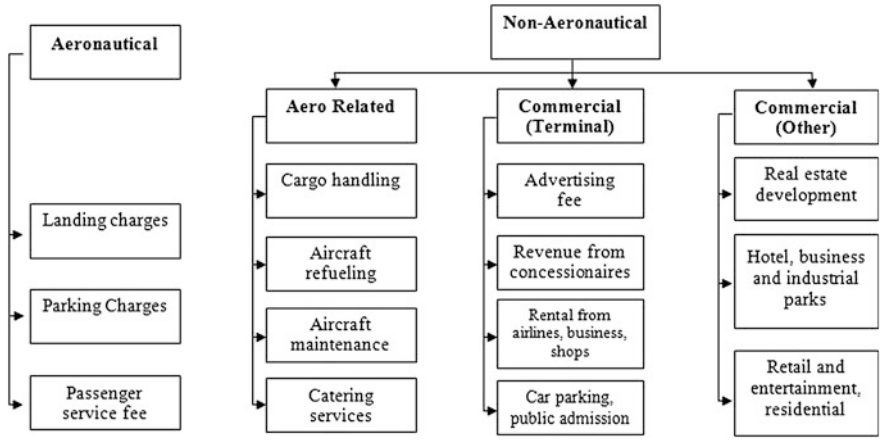


Fig. 9.3 Aeronautical and Non-aeronautical services in IGI airport, Delhi (*Source* GMR (2006))

implementing and operating the project. The SPV may be formed as a joint venture company with equity participation from multiple private sector parties and the public sector where a project or a set of projects is treated as a separate entity from the sponsors. Funds are then borrowed solely based on the project's cash flows and the equity in the SPV itself. This independence allows the project package to be separated from the equity investors' balance sheet. Commonly observed models of concession contracts in South Asia are toll roads, sea-ports, airports, etc.

Divestiture. When the underlying asset is not returned to the public sector ever, i.e., the public sector relinquishes the right of ownership of assets to the private sector, it is sometimes referred to as a BOO (build, own, and operate) contract also. Since the public sector is still a party to this project, hence, it is not the case of pure privatization.

For the construction and development of road, seaport, airport, and other infrastructure in India the concession model of DBFOT-type has been taken up quite frequently. Figure 9.1 shows the constituents of a project company as a joint venture involving the public authority, developer, construction contractor, operating contractor, and the financial investor to define the ownership and risk sharing pattern of a PPP project. The sponsor/developer bids and gets a DBFOT Project based on their assumptions of long-term projected cash flow from the project given certain long-term assumptions regarding the state of the (macro) economy and its structural features, sectoral and local realities, and regulatory arrangements. A separate independent SPV is floated. It gets equity from developer and other stakeholders and raises funds from the market (Project Finance) which is typically 70–90 % of the total investment. The project risk borne by the developer is limited to his share of equity participation; the delivery compulsion, and the debt remains in the books of SPV only. Since the ownership of the asset lies with the public authority, the cash flow from the project is the only source of return for both debt

and equity investors, with debt being the senior claimant. So, the financial institutions (as the supplier of the debt finance) have a big say in the process and they employ agents like Consultants, Merchant Bankers, and/or Insurer to monitor the transparency and internal efficiency of the system.

Thus, the transaction heavily relies on contractual commitments between the project participants and, therefore, the regulatory and supervision capacity of governments is crucial to the success of these transactions. However, project finance is more complex than traditional corporate or public financing, typically involving many more parties and resulting in significantly higher transaction costs. It is not unusual for this transaction cost to exceed the magnitude of the equity finance. The total transaction costs exceeded 7–10 % of the total project value in a number of projects in Latin America (Guasch 2004). In that event the developer may like to go for re-negotiation.

PPP Initiatives in Developing Countries

Chile first embarked upon its PPP project in 1974 and with a 15-year lag Argentina adopted this model in 1989. At around the same time some of the most economically aggressive East Asian countries started to rely on creative project finance for many of their large infrastructure projects followed by Arica and South Asia. The 1990s were characterized by a massive policy redirection toward private participation in infrastructure development in developing economies where the volume of such investment grew more or less steadily till 1997–1998 in Latin America and till 2007 in the other part of the globe. Between 1984 and 2003, PPP generated investment commitments of about USD790 billion. This reflected, in part, the disappointment with ineffective state-operated utilities, the promise of private funding, and the opportunity offered by technological, regulatory, and political changes (Estache 2005). Table 9.1 reports the share of private participation in different public utilities across different regions of the World by the year 2002. It is observed that both electricity generation and telecommunication are more strongly dominated by private investments than electricity distribution and in all these services the share of participation is increasing with the increase in the level of development. In high income countries, on an average 70 % of electricity generation and 83 % of telecommunication services are dependent on private capital. In terms of regional distribution of the adoption of PPP, among the developing countries, the situation is most discouraging in the Middle East and North Africa and most encouraging in Latin America and Caribbean. In fact, for the latter the share of private participation in electricity distribution (57 %) exceeds that in the developed countries (43 %).

According to the information provided in the database of Public–Private Investment Advisory Facility (PPIAF) of the World Bank it is observed that energy, transport, and telecommunication have attracted private investment more successfully than water supply or sewerage management. Since the latter involves

Table 9.1 Countries with private capital as of 2002 (% of sample)

Number of countries in sample	155 Electricity Generation	155 Electricity Distribution	155 Telecom
By income level			
Low income developing	33 %	26 %	37 %
Low-Middle Income developing	39 %	31 %	51 %
Upper-Middle Income developing	58 %	39 %	66 %
High Income (Developed)	70 %	43 %	83 %
By World Bank-Region Classification			
Sub-Saharan Africa	33 %	24 %	41 %
East Asia and Pacific	60 %	20 %	38 %
Europe and Central Asia	41 %	37 %	58 %
Latin America and Caribbean	68 %	57 %	67 %
Middle East and North Africa	7 %	6 %	23 %
South Asia	38 %	13 %	50 %
High Income (Developed)	70 %	43 %	83 %
Total	46 %	32 %	55 %

Source Estache (op.cit.)

direct tariff collection from the consumers for a common facility, the private investors may be extra careful in getting involved in such projects due to greater chance of political as well as regulatory interventions. In Latin American countries transport service is a popular PPP avenue, whereas in Asian countries road, seaport, airport, etc., are increasingly attracting private investment. Investors in general tended to favor Greenfield projects over projects that rehabilitate existing assets. However, concession and joint venture (BOT contract) are most visible in case of transport services and complete divestiture in case of energy sector.

PPP was initiated in Latin America in an environment in which the cost of private funds was at a historical low level. Indeed, there was an unusually high excess supply of funds on the international capital markets favoring FDI and other forms of international capital flows. This eased PPP-type investment in politically committed countries. The 1990s also saw a wave of ideological changes favoring market-oriented reforms among leaders of all political sides. It may also be useful to point out that in almost all countries, privatization and infrastructure reforms were actually part of a wider reform agenda.

However, in Latin America private sector participation has started losing its popular acceptance since the early years of the present century and investors' appetites have gradually waned. About 40 % of the population expressed discontent with private sector participation in 1998 and in 2007 it was closer to 75 %. Thus, public opposition has become a real constraint on Private-Public Investment (PPI) in some countries, both politically and operationally. The anecdotal and more formal evidence available suggests that in most countries, the direct pressure on the government budget was relaxed in the short run but most of these projects

were not commercially viable as the indirect pressure on government subsidy in the form of OPEX and CAPEX concessions were exceeding the revenue earning from tax and non-tax sources (Campos et al. 2003). The longer run story is however much more subtle. In many regions, the concern for affordability of public services was so clearly at the top of the politicians' priorities that subsidies are required to pacify sentiments. Even for regions where initially the reforms had introduced major tariff rebalancing and subsidy cuts to ensure increased sector autonomy, the political pressure eventually forced them to return to the subsidization of the sectors (Chong and Lopez-de-Silanes 2003). There are thus plausible reasons for tensions between public authority, operators, and users leading to some of the divorces in PPP relationships.

Besides this political flexibility in the regulatory commitment, with the onset of global recession in 2008 private investment in infrastructure witnessed a sharp retreat in developing economies with less developed credit markets. Thus, the overall flow turned out to be cycle sensitive. This lays credence to the fact that emerging markets with less developed domestic sources of long-term credit may suffer disproportionately in times of global crunch. All these factors culminated into a high rate of request for separation or divorce from this policy.

The Indian Experience

India is still a relatively new entrant into the PPP market as compared to the more matured European and Latin American markets. The Indian bonhomie with PPP initiative started tentatively with the adoption of the National Highway Development Program in 1998. The Golden Quadrilateral Project required enormous capital investments and delivery capabilities. Initially the projects were given out on a contractual basis following the usual tendering process. The established contractors executed such projects based on the timeline set by the NHAI. However, this traditional model of infrastructure delivery suffered from two major drawbacks. First, the huge capital required for such infrastructure projects were to be financed by the Government's budgetary resources. Second, the delivery excellence of the contractors, in terms of time and quality, were in many instances not up to the mark. As a result, the Government decided to go for the PPP model of infrastructure delivery with the private sector bringing in the investment funds and optimizing its Return of Capital Employed (ROCE) through a partially regulated pricing system. Initially, NHAI went for a mix of user fee-based and availability-based return[6] with a guaranteed return to the developer with contractually agreed timeline for revision of such returns over the lifetime of the investment. The

[6] Just like user-fee based PPP the availability-based PPP also involves a private party designing, financing, building, and operating an infrastructure but instead of collecting the fee directly from the end-user here the public authority makes the payment.

average project sizes were also small with a few big projects entailing an investment of greater than Rs.500 crores.

The success of the NHAI initiative, both from the point of view of increasing efficiency in delivery and mitigating the fiscal constraints of the Government brought forth a deluge of PPP projects encompassing almost all the sectors of the economy. The first major diversification out of the road sector was in the airport sector with the privatization of four major airports, viz., Delhi, Mumbai, Hyderabad, and Bangalore on a concession basis (Table 9.2). The cases of Delhi and Mumbai International Airports were instances of Brownfield projects with Lease-Develop-Operate-Transfer (LDOT) type concession contracts and those of Hyderabad and Bangalore airports were fresh Greenfield projects with Built-Own-Operate-Transfer (BOOT)-type agreements (Table 9.2).

It has already been noted that these airport initiatives were followed by other PPP projects in power, urban infrastructure development, etc. by the Central Government. The State Governments subsequently entered the picture and extended the PPP initiative to social sectors like health, education, etc. Available data till 2011 show that till date, Rs.645,000 crores of investments have been done in India through the PPP initiative, both at the State as well as at the Central Government levels. A preliminary study of the data (Table 9.3) highlights a few important features. First, though late starters, the total investment under the State Government sponsored projects has overtaken the total investments under the Central Government sponsored projects. Second, for both center and state, road projects are enjoying the maximum investment; however, the State Governments are increasingly using the PPP model for delivery of social infrastructure projects like schools, hospitals, etc., along with traditional areas like roads, ports, power, and urban infrastructure.

State-wise breakup of the PPP projects in infrastructure development is reported in Table 9.4. Mostly these are state-level projects; however, some inter-state projects have been taken up by Kerala, Tamil Nadu, and Puducherry. A wide range of variations is observed in terms of both number of projects and the total cost of projects. States like Maharashtra, AP, and Karnataka are taking up a large number of projects followed by Rajasthan, Gujarat, and Madhya Pradesh; among the

Table 9.2 Major Airports developed under PPP in India

Project name	Cost (Rs in crores) [Initial estimate]	Type
Modernization of Delhi International Airport	8,600	Brownfield, LDOT
Modernization of Mumbai International Airport	5,800	Brownfield, LDOT
Hyderabad International Airport	2,478	Greenfield, BOOT
Bangalore International Airport	1,930	Greenfield, BOOT

Source Ministry of Finance, GoI (website)

Table 9.3 PPP Projects taken up by the Central and State Governments in India

Project Type → Project Sponsor ↓		Roads	Airports	Ports	Railways[a]	Urban dev.	Industries	Total cost (Rs. Cr)
Central Government	No.	161	5	37	7			
	Cost (Rs. Cr)	98034	19111	17690	5724			140559
State Government	No.	412	6	71	6	314	17	
	Cost (Rs. Cr)	134701	1630	89855	24448	109699	71626	

Project Type →		Power	Tourism	IT	Education	Health	Others	
State Government	No.	47	83	16	27	16	10	
	Cost (Rs. Cr)	59093	4478	3566	2618	1821	765	504300
Total Cost (Rs. Cr)								644859

[a] Including metro railways

Source Ministry of Finance, GoI (website)

Table 9.4 State-wise distribution of PPP projects (2011)

State	Number of projects (average cost)	Cost (Rs. in cr)	State	Number of projects (average cost)	Cost (Rs. in cr)
AP	178 (359)	63934	KER_TN	1 (99)	99
Assam	17 (88)	1495	MP	63 (183.5)	11562
Bihar	1 (4)	4	Maharashtra	250 (248)	61992
Chandigarh	2 (37.5)	75	Orissa	22 (456)	10030
Chhattisgarh	9 (181)	1628	Puducherry	2 (1683.5)	3367
Delhi	9 (223)	2006	PUDCH_TN	1 (61)	61
Goa	1 (30)	30	Punjab	37 (74)	2756
Gujarat	78 (695)	54202	Rajasthan	96 (199)	19104
Haryana	25 (2563)	64071	Sikkim	24 (713)	17111
Jharkhand	6 (113.5)	681	Tamil Nadu	27 (447)	12069
Karnataka	117 (797)	93206	UP	11 (6466)	71130
Kerala	18 (490)	8816	WB	30 (162)	4870
Total				**1025 (492)**	**504300**

Source Ministry of Finance, GoI (website)

remaining states nearly two-thirds have less than 20 projects. Total private investment is the highest in Karnataka and the lowest in Bihar. The average cost of project is the highest in UP and the lowest in Bihar. In fact, Bihar has taken up only one project of Rs.4 crores under PPP scheme. If PPP arrangement is

recognized by the state as an attractive market-based alternative then the number of projects would be large and PPP option would be taken up even for low cost projects. However, when PPP is opted to ease pressure on the government exchequer then only large projects would be funded through this arrangement leaving the smaller ones for direct government funding. A careful scrutiny of Table 9.4 puts Maharashtra in the first group and Uttar Pradesh in the second group. Hence, coincidentally, the economically better performing states are the ones who have been the leaders in the adoption of the PPP model.

The PPP initiative in India brought in a new set of stakeholders. Traditionally, the Indian construction market was dominated by a few strong contractors. The structure was like a pyramid. At the top of the pyramid were a few (around 8–10) large established listed companies with reasonable barriers to entry, pan India presence, and delivery capability across the EPC (engineering-procurement-construction) value chain across business segments. A thin middle layer comprised upcoming firms (around 20) with weak barriers to entry, strong regional presence, limited delivery capability with strong emphasis on JV/Consortium (mainly with global firms) approach, and limited capability across business segments. Finally, a multitude of small/unorganized players were at the grass-root level with no entry barriers and strong local presence doing construction work with low value added. They mostly worked as sub-contractors to the first two categories. However, with the advent of the PPP model a new set of market participants developed while the existing ones evolved into new forms. There is increasing entry of developers like GMR, GVK PIL, IRB Infra, Lanco Infratech, etc., who are taking up PPP projects as concessionaires. On the other hand, many existing construction companies started integrating the concession-construction business to ensure steady and predictable income flows that the pure construction business would fail to guarantee.[7] While some of the construction companies floated independent infrastructure firms to manage the concession business, others developed new business verticals within the existing company for PPP projects. Quite a large number of new and smaller firms also entered the business by adopting the consortia approach. Since global firms are sometimes members of these consortia, there is a consequent linkup between small domestic firms and the giant multinational firms through these PPP projects.

The main stakeholders in the Construction and Engineering business of India under the PPP regime can be identified as (a) the creditors, (b) the promoters, (c) the regulators, and (d) end users. The creditors are mostly public sector banks and only recently 100 % FDI has also been allowed in these PPP projects; the promoters are of two varieties, developers like GMR, GVK, etc., and newly evolved

[7] In pure construction business the cash-flow depends on a number of statutory obligations and clearances from the administrative office and generally, the time lag between completion of job and receipt of the payment is unusually high, leading to a cost escalation due to unavoidable transaction cost. However, in this market-linked concessionaire-contractor setup the cash flow is directly tagged with the financial market, and therefore, efficient performance gets always rewarded by timely payments.

contractor developers like L&T, Gammon India, IVRCL Infra, Nagarjuna Construction, Simplex Infrastructure, etc. For central projects regulators are specially established institutions like the National Highway Authority of India (NHAI), Airports Economic Regulatory Authority of India (AERAI), Telecommunication Regulatory Authority of India (TRAI), etc., sanctioned by the Cabinet Committee on Economic Affairs and for state projects the regulators are the respective state ministries.

The Indian Planning Commission is talking of making USD 1 trillion investments (approximately Rs.5 million crore) in infrastructure for the 12th plan period (much of which is to come from the private sector under the PPP approach). Add to it planned CAPEX of industries like steel, cement etc., and real estate demands (both commercial and residential) for 2012–2017 and we would arrive at a mind boggling figure where the present gap between capability vis-à-vis demand is glaring.[8] This raises two issues: first, how to tag up the finances and second, how to guarantee the physical creation of infrastructure, i.e., the delivery issue. Both physical and financial capacity of the economy needs to be assessed in terms of resource constraint with respect to available/expected supply of key resources like Men (workmen), Project Management Skills (supply of technically trained manpower), Material (commodity supply), Machinery (Technology, automation/mechanisation), and Money (finances). As the domestic availability of these key factors is not adequate to satisfy the requirements, hence, in a globally connected world, PPP-led infrastructure development turns out to be the only viable option. So, the point is whether the existing regulatory framework and the stakeholders has the capability to deliver on three fronts, viz., (a) ensure adequate return on capital employed by the developers, (b) ensure delivery excellence in terms of quality and timeliness for large complex infrastructure projects, and (c) ensure low user costs.

In light of the foregoing discussion, to examine the adequacy, feasibility, and sustainability of the institutional arrangement of the prevailing practice in India, a case study of pure developers is presented in the following section on the basis of the specific experience of Delhi International Airport Limited (DIAL), which is the largest Brownfield airport development project completed under the PPP initiative since 1998.

[8] Assuming that the average construction component of the infrastructure projects is around 50 %, implies that the construction component of infrastructure projects that are to be executed in the next 5 years would be around US Rs.2.5 million crore, which implies an average order execution of Rs.0.5 million crore per year. Even assuming it 50 % success as the base case, it comes to Rs.0.25 million crore worth of construction delivery per year. Now, as at end March 2012 (2011–2012), the cumulative sales of the major E&C added up to only around Rs.0.125–0.130 million crore. Even if we account for the fact that the E&C industry in India is highly fragmented the gap is glaring.

Delhi International Airport Limited: A Case Study[9]

Background

The aviation industry in India has experienced remarkable transformation in the last 20 years. As the economy experienced changes in policies from a highly regulated regime to a more market-oriented structure with rising per capita income, greater global connectivity, and stronger outward orientation there was significant increase in the demand for mobility to meet business and personal needs. Thus, there was an increase in demand for physical infrastructure in general and transportation in particular. Air travel is now becoming an affordable mode of transportation for a much broader segment of the population than before. In order to meet the growing demand for air transport the Ministry of Civil Aviation (MoCA) in 2006 projected a requirement of an additional Rs.405 billion to augment and modernize existing airports as also to construct new Greenfield airports. The revenue surplus generated by Airports Authority of India (AAI) was found to be grossly inadequate to meet this requirement. Thus, the Planning Commission of India set a target of investing Rs.300 billion in airport infrastructure during the 11th Five Year Plan (2007–2012) through a combination of public and private investment. The public sector is modernizing Chennai and Kolkata airport along with 35 non-metro airports while two major metro airports, Delhi and Mumbai, have been attempted to be modernized as Brownfield projects in the PPP mode. Hyderabad, Bengaluru, and Cochin airports also offer good examples of Greenfield PPP initiatives.

Formation of DIAL

In September 2003, Cabinet approved the proposal of the MoCA that restructuring of the Delhi airport may be undertaken by the JV route by formation of a separate company between AAI and the JV partners. Cabinet also directed the formation of an Empowered Group of Ministers (EGoM) comprising Ministers of Law and Justice, Finance, Disinvestments, and Civil Aviation to decide on the detailed parameters for the initiative. In Feb 2005, the EGoM approved all the key principles for the Request for Proposal (RFP) document along with the draft transaction documents, viz., (i) the Operations, Maintenance, Development Agreement (OMDA); (ii) State Support Agreement (SSA); (iii) Shareholder Agreement (SHA); (iv) The Lease Agreement; (v) CNS-ATM[10] Agreement; and (v) the State

[9] The analysis largely depends on source materials like CAG Report on DIAL (2012), NCAER report on Economic Impact Assessment of Delhi International Airport (2012), Airports Economic Regulatory Authority (AERA) website, and Draft National PPP Policy (2011).

[10] Communication, Navigation, Surveillance (CNS) and Air Traffic Management (ATM).

Government Support Agreement (SGSA) document. Consequently, technical and financial bids were called for from interested JV partners and after evaluation of the technical and financial bids the EGoM forwarded their recommendation to the cabinet which approved the proposal on Jan 1, 2006.

For Indira Gandhi International (IGI) Airport, the JV partner was identified as M/s GMR consortium (6 partners). After the technical bid, the most important criterion used for the financial bid was to select a JV partner which gave the best revenue share to the AAI. GMR consortium was the highest bidder offering 45.99 % of the gross revenue share to the AAI. The Airport Authority Limited (AAI) incorporated on March 1, 2006 a subsidiary company as an SPV named M/s Delhi International Airport Limited (DIAL). Consequently the OMDA was signed on April 4, 2006 and 74 % of the equity share of DIAL was sold to GMR consortium in accordance with the Shareholder's Agreement. In terms of the Agreement, issued share capital of Rs.200 crore[11] was jointly held by AAI (26 %); GMR Infrastructure Ltd (31.1 %); GMR Energy Ltd. (10.0 %); Fraport AG Frankfurt Airport Services Worldwide (10.0 %); Malaysia Airports (Mauritius) Private Limited (10.0 %); GVL Investments Private Ltd (9.0 %), and India Development Fund (3.9 %). Subsequently, with the signing of other agreements, viz., SSA, SHA, SGSA, Lease Agreement, and CNS-ATM Agreement, the IGI Airport was handed over[12] to DIAL with effect from May 3, 2006. The initial concession period is 30 years, extendable for another 30 years, subject to satisfactory performance. The total project cost was initially estimated at Rs.8,600 crore and subsequently revised at Rs.8,975 crore.

To understand the nature of financing as well as risk and obligation sharing through these agreements the scope of each agreement should be considered in delivering the required services.

The OMDA Agreement

The OMDA lays down the rights and obligations of both the parties, the terms of revenue sharing and the duration of the concession, definition of aeronautical and non-aeronautical services, fixation of tariff for aeronautical and non-aeronautical services, conditions for outsourcing of services, mechanism for dispute resolution, and finally the conditions and modalities of asset transfer in the present and the future.

---

[11] As on March 2012, the equity capital of DIAL was Rs.2536 crore.

[12] When the OMDA was signed between AAI & GMR consortium there was no single airport regulatory authority in India. The Airport Economic Regulatory Authority (AERA) Act was passed in December 2008 and the Act came into force in January 2009 and AERAI came into force in September 2009.

State Support Agreement

This is complementary to the OMDA and lays down the roles and responsibilities of the Government of India and DIAL. Furthermore, it also lays down the principles of tariff fixation for aeronautical services.

State Government Support Agreement

Signed between the Government of the National Capital Territory of Delhi and DIAL, it provides the State government's support in matters relating to removal of encroachment, procurement of extra land, removal of obstruction outside airport boundary, provision of all utilities on a payment basis to DIAL, and to provide assistance in procuring various clearances which are required under law for implementation of the OMDA.

Lease Deed Agreement

The lease deed agreement was signed to lease the *demised premises* (as defined in OMDA) on *as is where in* basis on an annual lease rental of Rs.100 for a period of 30 years extendable by another 30 years by virtue of extension of the concession period.

CNS/ATM Agreement

To provide air traffic support at the airport since only AAI is authorized to provide necessary air traffic services within the Indian airspace.

Shareholders' Agreement

Recorded the terms and conditions to govern the relationship in their mutual capacity as the shareholders of the JVC.

Airport Operator Agreement

According to the OMDA, DIAL is required to enter into an airport operator agreement with the Airport Operator (AO) who is a member of the consortium. The agreement contractually sets out the role, responsibilities, accountabilities, and financial arrangements between DIAL and AO and exit clauses. Accordingly,

the AO agreement was signed between DIAL and Fraport AG Frankfurt Airport Services Worldwide.

Performance

DIAL commissioned a new runway 11–29 at IGI Airport on September 25, 2008, and also inaugurated the new domestic departure terminal 1D (T1D) on 26 February, 2009. In March 2010, DIAL completed the construction of an integrated passenger terminal (Terminal 3) completing the first phase of the airport. Thus, the Concessionaires completed Phase I of the construction of IGI airport as specified in the Master Plan in stipulated time and according to the scheduled plan and quality standards. Today, IGI Airport, Delhi is one of the busiest and fastest growing airports in the country and in the Asia Pacific region. It is acknowledged in the CAG Report (2012) that there have been significant improvements in services at the airport for the traveling public. It is ranked the second best airport worldwide in the 25–40 million passengers category for Airport Services Quality (Airport Council International 2011). Delhi airport connects 51 international and 41 domestic routes. During 2010–2011 this airport handled nearly 30 million passengers where the domestic: international break up was 69:31. A major portion of Delhi airport's revenue comes from non-aeronautical services (45 %) followed by aeronautical services (36 %) and the rest comes from services like cargo, commercial and professional development, etc.[13]

In aeronautical revenue, airline landing fee contributed 70 % followed by passenger fee (26 %); the remaining 4 % comes from aircraft parking charges. In the non-aeronautical revenue, the highest percentage share comes from rental and services followed by retail and concession. On the expenditure side the share of operating cost is the highest (49 %) followed by staff costs (28 %) and administrative and other costs (23 %). Though over time the revenue per passenger is marginally increasing, the revenue to operating cost ratio remained stagnant within the range of 1.2–1.3 since 2008–2009 (NCAER 2012). However, the financial performance of DIAL has not been encouraging. Although the basic forecast of AREA was that the DIAL would break even from 2012 to 2013, available data shows that DIAL has been making losses quarter over quarter including the first quarter of 2012–2013. Lower utilization of existing capacity was recorded for both aeronautical as well as non-aeronautical services leading to revenue gap. Some cost escalation was also noted by CAG mostly due to expansion of ground floor area beyond the limit sanctioned in the Major Development Plan, and, therefore a consequent proportional increase in other costs. Instead of the estimated cost of Rs.8,975 crore in January 2008, by July 2010 it had gone up to Rs.12,857 crore. Hence, besides revenue gap there is a huge burden of debt. To make this situation

[13] For a classification of aeronautical and non-aeronautical services see Fig. 9.3.

commercially viable the Ministry of Civil Aviation promised a number of additional incentives to DIAL. A critical evaluation of this regulatory environment will help us to assess the sustainability of these PPP initiatives in the Indian context.

Regulation and Its Loopholes

The first issue relates to capital expenditure and financing. As has already been noted, the initial estimated project cost as approved at financial closure of IGIA was Rs.8,975 crore in Jan 2008. As per the original estimates and the OMDA, the entire funding was proposed to be financed through equity, debt, security deposits, and internal accruals. However, the total project cost as on July 2010 as claimed by DIAL was Rs.12,857 crore of which AERA approved Rs.12,502.86 crore as the Regulatory Asset Base. Thus, there was a huge funding gap. Based on an earlier representation by DIAL (and its promoters), the Ministry of Civil Aviation in December 2010 allowed DIAL to levy a development fee (DF) at the IGI Airport for a period of 18 months to bridge this funding gap. The total accruals under the Development Fee route were targeted at Rs.3,619.5 crore. This changed the total financing pattern of the project which was not envisaged under the original OMDA (Table 9.5). Because now, the development fee would cover more than 27 % of the project cost. It is opined that this sanction to impose development fee would lead to an alteration of the original basis on which the concession was competitively bid out and would affect the competitive fairness of the award. On the other hand, without this provision the airport itself was not a feasible commercial proposal.

Secondly, issues have been raised relating to transfer and use of assets and the period of concession. As per the OMDA, DIAL was initially handed over 4608.9

Table 9.5 Cost escalation and development fee in DIAL

Funding (1)	Cost as per estimates in January 2008 (2)		Final cost in March 2010 adopted by AERA (3)		Cost escalation (Percentage) 3(b)−2(b)
	Amount (a)	% of total cost (b)	Amount (a)	% of total cost (b)	
Equity	1202	13.39	2300	18.40	5.01
Loans	4986	55.55	5266	42.11	−13.44
Security deposits	2738	30.51	1471.51	11.77	−18.74
Internal accruals	49	0.55	50	0.40	−0.15
DF	Not envisaged	Not envisaged	3415.35	27.32	27.32
Total	8975	100	12502.86	100	

Source CAG Report on DIAL (2012)

acres and subsequently 190.19 acres out of the total IGI Airport holding of 5106 acre for an upfront one-time fee of Rs.156.2 crore. This land area is called the demised premises. DIAL can use 5 % of this land as *hospitable land* for commercial purposes (240 acres).[14] Moreover, DIAL has been given the unilateral right to manage the IGIA for 60 years[15] and the State Support Agreement (SSA) allows the Right of First Refusal (ROFR) to DIAL with regard to any second airport planned within 150 km radius of the IGIA. Independent valuation by Merill Lynch in Aug 2011 valued this land at around Rs.100 crore per acre. According to the CAG Report the *hospitable land* has a current valuation of Rs.24,000 crore and a much higher future valuation (based on lease rental for 60 years) which is not reflected in the upfront fee. On the other hand, it should be highlighted that the potential value/income flow from (i) the land that is available implicitly through the contract; (ii) possible extension of concession period, and (iii) the "non-compete" provisions was part of the original valuation considered by the concessionaire when bidding for the project on which the targeted revenue share was calculated by the bidders. Thus, when the original OMDA was drafted by the EGoM (even before the inception of AERAI) and placed before the bidders to find out the best possible private sector partner, these highly predictable future possibilities of rent-seeking and cream-skimming should be taken into consideration. Thus, the issue lies in the structuring of the contract itself and not in any particular slack at the level of implementation.

The third area relates to the revision of airport tariff. According to the OMDA/SSA, DIAL can appeal to the AERA for revision of aeronautical charges to the AREA for the first regulatory period, i.e., for 5 years (01.04.2009–31.03.2014) based among others on an inflation indexation formula, changes in traffic forecasts, changes in cost of debt, and the corresponding weighted average cost of capital (WACC) and a justifiable return on equity for DIAL. DIAL initially submitted a proposal for revision of tariffs for aeronautical services in June 2011 and after a series of correspondences, finally asked for a onetime increase in aeronautical tariff based on the X Factor of 874 %. The AERA fixed the new tariff structure, effective May 15, 2012 with a onetime Escalation determined on an X factor of 345 % based on crucial assumptions relating to traffic growth, the Return on Equity (16 %), and Weighted Average Cost of Capital (10.33 %). DIAL went for further appeal. However, the moot point is that the level and the wide divergence in the extent of revision sought and the revision awarded reflect huge variations in perceptions about returns from the business by the concessionaire and the regulator.

The fourth important concern is related to the social cost of this sharp increase in user-fee over a very short period. The newly introduced development fee will raise the cost of each embarking domestic passenger by Rs.200.00 and for foreign

[14] Special arrangements to mitigate commercial risks.

[15] Initially granted for a period of 30 years and upon satisfactory performance extendable for another 30 years.

passenger by Rs.1,300.00. Thus, the common travelers are bearing the burden of cost escalation which is violating the ability principle of financing public good. Instead, DIAL could have enhanced the user charge for VIP lounge, which is a common practice followed in the UK and other European countries.

The fifth issue relates to changes in regulatory purview. It needs to be highlighted that when the OMDA was signed between AAI and GMR consortium there was no single airport regulatory authority in India. The Airport Economic Regulatory Authority (AERA) Act was passed in December 2008 and the Act came into force in January 2009 and AERA came into force in September 2009 as the sole regulator for the airport sector and tariff fixation. As per AERA Act, the targeted revenue from tariff takes 100 % of aeronautical and only 30 % of non-aeronautical services. However, there are definitional differences as to what are aeronautical and non-aeronautical services between the OMDA and the AREA. Services like Ground Handling, Cargo Handling, and parking are treated as non-aeronautical services in OMDA but are aeronautical services as per AREA. Thus, inclusion of such provisions as non-aeronautical services, in effect, lowers the total targeted revenue base for tariff fixation. It may be highlighted that the concession bid was done on the basis of total revenue sharing and GMR consortium agreed to pay 45.99 % of its revenue to the AAI. Furthermore, the revenue for transfer assets (commercial exploitation of 5 % of demised premises) was not included as a part of the revenue-base for tariff fixation. This ambiguity in crucial definitional issues and the risks associated with the change in regulatory 'goalposts' adversely affects the long-term sustainability of the PPP agreement.

Concluding Concern

This chapter tries to trace the path of the gradual entry of the private sector in provision of infrastructure in a rapidly growing developing economy like India. It started with describing the context in which the economy was almost compelled to go for PPP arrangement in infrastructure development. To help the economy to grow at a faster rate there is a huge requirement for infrastructure support. It was neither financially possible nor technically feasible for the Government of India to make this provision through the public system alone. Involvement of the private sector was expected to ease pressure on government budget, relax the constraints imposed by the limited local availability of trained and skilled manpower, and bring in appropriate technology and better management along with monitoring of the funds through the integration with the global financial market. For each project a Joint Venture Company would be established under public–private-partnership mode following the doctrine of limited liability and both equity and debt financing would be combined to raise the required fund. In case of DIAL it was seen that the share of public investment in the total project cost is only 4.8 %. Hence, there is no doubt that this arrangement helped to ease the burden on the government budget. However, relaxation of the delivery constraint on a long-term basis crucially

depends on the quality of private players entering the game. It is observed quite often that guided by the motive of targeted acquisition and rapid diversification. even established players are venturing outside their core area of competence and forming consortium with players with relatively unknown worth and reputation. In this rather fluid state of everybody doing everything, the regulatory challenges increase manifold. It is a rather tricky job for the regulator to decide when to bail out the joint venture partners from temporary snags and help them to be in the long-run sustainable path through specific incentives and supports including renegotiation of contracts. On one hand there is the question of commercial viability and on the other there is resistance from politicians and the civil society. If the balance cannot be clicked between this commercial motive of efficiency and the social pressure of equity, then the overt subsidy will take a latent form of covert concessions and the eventual pressure on public budget will put the viability of this entire approach under suspicion.

From the analysis of DIAL experiment, it was found that the regulators were not prepared to write a comprehensive contract where all the events at least for the near future would be foreseen and provisions for such contingencies would be kept in the contract. That is why the possibility of cost-escalation was not totally anticipated and called for a re-negotiation of the initial contract to accommodate development fees. Though at the time of writing the contract the regulatory body failed to predict the inflation rate, growth rate of gross domestic product, and other macroeconomic events they have granted 30 years (and with extension of the term 60 years) concession right to the private sector JV-partner. Finally, the regulatory sieve was not fine enough to detect the deviation from the Major Development Plan. Hence, under the new economic order though India may not have much options but to adopt PPP mode for infrastructure development, in the absence of judicious and careful regulation in terms of writing and implementing the contract it will create huge uncertainty in the supply chain and the economy will be thrown from the frying pan into the fire.

References

Bloom DE, Williamson JG (1998) Demographic transitions and economic miracles in emerging Asia. World Bank Econ Rev 12:419–455

Campos J, Estache A, Martin N, Trujillo L (2003) Macroeconomic effects of Private Sector Participation in Infrastructure. In: Easterly W, Serven L (eds) The limits of Stabilization. Stanford University Press, Stanford

Chong A, Lopez-de-Silanes F (2003) The Truth about Privatization in Latin America. In: Inter-American Development Bank Working Paper, October

Delmon J (2010) Understanding Options for Public-Private Partnerships in Infrastructure Sorting out the Forest from the Trees. In: Policy research Working Paper 5173, World Bank

Edward F, de Mästle CT, Yescombe ER, Encinas J (2011) How to Engage with the Private Sector in Public-Private Partnerships in Emerging Markets. PPIAF, World Bank

Estache A (2005) PPP partnerships versus PPI divorces in LDCs. In: World Bank Policy Research Working Paper 3470

Guasch JL (2004) Granting and Renegotiating Infrastructure Concessions: Doing it Right. Development studies. World Bank Institute, Washington
NCAER (2012) Economic impact study of Delhi airport
Report CAG (2012) Implementation of public private partnership: Indira Gandhi International Airport. GoI, Delhi

Chapter 10
Performance Evaluation Techniques: An Application to Indian Garments Industry

Simanti Bandyopadhyay and Subrata Majumder

Introduction

One plausible way to assess empirically the impact of reforms in globalization on an industry is by evaluation of its performance under these reforms in the input–output framework. Whether the reforms have resulted in a change that can be captured in the performance indicators is an interesting empirical question. The main objective of the chapter is to build up alternative empirical frameworks of performance evaluation with an application to analyze the impact of reforms through performance indicators of an industry. In the process, we give an overview of the techniques of performance evaluation applied in the economics and management literature.

One of the industries which experienced dramatic changes after globalization in India is Textiles with its several subsectors, of which 'garments' component benefited the most under the said reforms regimes. The positive impact comes through a prospective rise in exports, which caused a change in domestic consumption pattern backed by globalization in preferences. This consumption reshuffle, along with the structural change, leaves a visible impact on employment and productivity (Ray 2004). Keeping in mind the relevance of the issue, we have chosen the garments sector in India and have applied different techniques to estimate the technical efficiency of the sector.

Production is an act of transforming inputs into outputs. The objective of production is to create value through transformation and, therefore, outputs are in

S. Bandyopadhyay
Indian Council for Research on International Economic Relations, Core-6 A, 4th Floor, India Habitat Centre, Lodi Road, New Delhi 110003, India
e-mail: simantib@gmail.com

S. Majumder (✉)
Commerce Evening Department, St. Xavier's College (Autonomous), Kolkata 700016, India
e-mail: subrata_eco@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_10,

general desirable outcomes. Hence, more output is better. At the same time inputs are valuable resources with alternative uses. The objective of a firm or a decision-making unit in general is either to produce as much output as possible from a specific quantity of inputs or to produce specific quantity of output using as little input as possible. An input–output combination is a feasible production plan if, given the state of technological knowledge, the output quantity can be produced from the associated input quantity or vice versa.

The performance of any production unit can be evaluated in terms of relative productivity or the efficiency of the unit concerned. Productivity is a descriptive measure of performance without any reference to the optimal achievable target, whereas efficiency is a normative measure assessed with reference to the production frontier. If the sector experiences an upgradation in production technology, a constituent firm may experience an increase in total factor productivity. However, this cannot be attributed to any specific factor of production. In fact, a change in total factor productivity refers to the shift in the production frontier.

Efficiency by its simplest definition of the output version refers to the ability of a firm to produce the maximum levels of outputs with a set of inputs. With a change in prices of inputs or a shift in technology or otherwise there can result a change in the input mix used by the firm which in turn affects efficiency. When we refer to the firm's ability to produce as much as it can without taking any possible impact of input-prices, it is called productive or technical efficiency (TE), whereas when the effect of prices of inputs is taken into consideration while measuring efficiency it is termed as price or allocative efficiency. The overall performance of a firm depends on both the components of efficiency.

The main objective of the chapter is to give an overview of the concept of technical efficiency and discuss different methods to measure the same. This measure of efficiency is particularly useful in the absence of market prices for inputs or outputs as in the case of problems in public economics and public policies. Hence, one of the advantages of this tool is that it can be applied to any economic activity in any sector even with constraints in availability of data on market prices. Throughout the discussion we refer mainly to the output version of the measure.

Technical efficiency is an index expressed as the ratio of actual production and the potential productive capacity of a firm using the same amount of resources. There are various ways to measure technical efficiency. Once the decision-making unit in a sector performs an economic activity transforming a set of inputs into a set of outputs and a frontier of production can be conceived of, considering all the decision-making units in the sector, we can apply the concept of technical efficiency to assess the performance of the units. While the basic principle of measurement of technical efficiency is the distance of the point of operation of a decision-making unit from that projected on the frontier, two factors, viz., the way the frontier is constructed and the way the distance is measured, make one method of estimation different from the other. There can be parametric methods in which the frontiers are given functional forms and parameter specifications or nonparametric methods in which behavioral assumptions are used to construct the notional frontier. For each family of parametric or nonparametric specifications, the

estimation can be done through mathematical optimization or econometric techniques. The distance between the point on which a decision-making unit *actually* operates and the point on the frontier on which it *should have* operated can be measured as a radial or a non-radial characterization. In this discussion we would mostly like to base our analysis on radial measures, taking one each from nonparametric and parametric families. We discuss the parametric Stochastic Frontier Analysis (SFA) which is the most widely used econometric technique in estimating technical efficiency. We also discuss the nonparametric method of Data Envelopment Analysis (DEA) which uses a linear programming principle for assessing technical efficiency and is deterministic in nature. The analysis is confined to the basic models on cross-section data. It is to be mentioned that the differences in the principles of constructing the frontier have some impact on the efficiency results, each class of techniques having their advantages and disadvantages which are discussed with the help of an application.

The garments industry in India has been chosen for this purpose and we have applied SFA and DEA to estimate the technical efficiency of the sector. The purpose of this exercise is to spell out elaborately the steps involved in applying each class of technique. This includes procedures for different pre-estimation checks and interpretation of results (post-estimation). Our analysis comprises garment units in eight major states in India. We also analyze the effect of firm-level internal characteristics like size and age on efficiency. Constrained by availability of data in the pre-globalization period in the same classification as the post globalization one, the analysis is carried out for four alternative years starting from 1999–2000 to enquire whether the effect of globalization can be captured through technical efficiency measures in the sector. A comparative assessment of the two methods of estimation is also attempted.

The chapter is organized as follows. Section-2 ("Evolution of Parametric Econometric Approach: Stochastic Frontier Production Function") discusses the parametric econometric approach of SFA; Section-3 ("Evolution of Nonparametric Optimization Approach: Data Envelopment Analysis") elaborates on the nonparametric optimization approach of DEA; Section-4 ("An Application") describes our application of the two techniques to the garments sector and the results; Section-5 ("Conclusions") spells out the conclusions.

Evolution of Parametric Econometric Approach: Stochastic Frontier Production Function

The standard regression technique estimates the conditional average value of the dependent variable (Y) given the set of independent variables (X) [i.e. $E(Y|X)$]. When applied in production theory it is able to estimate somewhat an average production function. However, a production frontier is defined as the maximum value of output of a firm produced by the given level of inputs. Depending upon the level of output produced by a firm with a given set of inputs, a firm can be

technically efficient and can operate on the frontier or it can be technically inefficient and can operate below the frontier. The empirical estimation of a frontier involves tracking the efficient firms from the data on inputs and outputs and tracing the locus of these maximum output levels. The challenge in the econometric approach is two-fold. First, to ensure that the estimated frontier is nothing but a locus of the maximum possible outputs which is also the closest possible envelope of the input–output data; second, to characterize the error term as a combination of statistical noise and inefficiency.

Deterministic Frontier Production Function Model

In case of cross-section data on the quantities of k inputs (denoted by vector X) and the quantity of a single output (denoted by Y) is available for each of N producers (indexed by i), a production frontier model may be written as: $Y_i = f(X_i; \beta)TE_i$ where Y_i is the output of producer i, $f(X_i; \beta)$ is the production frontier, X_i is the vector of k inputs used by producer i and β is the vector of the corresponding technology parameters (which is to be estimated). Now, the (output-oriented) technical efficiency of producer i, TE_i, is defined to be the ratio of the observed output (Yi) to the maximum output feasible at the input quantities used by the producer i, viz., $f(X_i, \beta)$. By definition, TE_i lies between 0 and 1.

The initial discussion on such parametric estimation of a frontier from the observed data points dates back to Winsten (1957). The approach becomes popular after the publication of the seminal work by Aigner and Chu (1968), who estimate technical efficiency considering a deterministic production frontier by using programming techniques. Timmer (1971) developed this procedure further by introducing the probabilistic frontier production model.

Applying a simple mean centric traditional econometric methodology, Winsten (1957) faces the challenge to estimate the production frontier with the help of input–output data. Clusters are formed with firms which are homogeneous in terms of their input use. Then from each cluster, only one firm which is the largest output producer relative to other firms within that group is selected. These firms, which are taken from each cluster, are considered as representative firms for the construction of the frontier. With the help of input–output data of those representative firms, a simple linear regression line is estimated. This regression line, estimated from frontier representative firms, is claimed to be the production frontier by Winsten's methodology and the input–output data of all the firms are superimposed over this frontier to estimate the technical efficiency. This methodology is not successful in constructing the frontier in a way so that all the firms are either on or below it. Some of the firms may lie even above the frontier as the regression line estimated as the frontier follows a mean centric approach which is considered to be a major drawback of the methodology.

Aigner and Chu (1968), offer a refined framework to incorporate technical inefficiency in a deterministic production frontier. If we take a simple log-linear Cobb-Douglas functional form

$$\ln f(X_i;\beta) = \beta_0 + \sum_{n=1}^{k} \beta_n \ln X_{ni}, \tag{10.1}$$

so that the logarithm of the observed output is

$$\ln Y_i = \beta_0 + \sum_{n=1}^{k} \beta_n \ln X_{ni} - u_i \tag{10.2}$$

where it is assumed that $u_i \geq 0$ for each i so as to ensure that $Y_i \leq f(X_i, \beta)$. That is to say for a particular vector of inputs the level of output observed for any firm cannot exceed that achievable at the frontier from the same input vector. The distance, if any, between the frontier output and the observed output would imply technical inefficiency of the firm. Equation (10.2) is a linear regression model with a non-positive disturbance. The objective is to obtain an estimate of the parameter vector β (which describes the structure of the production frontier), as well as an estimate of the disturbance term u_i which yields an estimate of TE_i for each producer i [in view of the fact that $TE_i = \exp(-u_i)$]. This framework has been used in three alternative ways in the literature to design methodologies for technical efficiency estimations which are briefly discussed below.

Corrected Ordinary Least Squares Method

The corrected ordinary least squares (COLS) method was developed by Gabrielsen (1975). Equation (10.2) is estimated following a two-step procedure. First, the ordinary least squares (OLS) method is applied which derives consistent and unbiased estimates of the slope parameters (say $\hat{B}_n$'s) and consistent but biased estimate of the intercept parameter (say $\hat{B}_0$). In the second step this (biased) intercept estimate is shifted up, i.e., corrected in the following manner so as to ensure that the estimated frontier bounds the data from above:

$$\hat{\beta}_0^* = \hat{\beta}_0 + \max_i(\hat{u}_i) \tag{10.3}$$

where $\hat{u}_i$ are the OLS residuals and $\hat{\beta}_0^*$ is the COLS estimate of the intercept. It may be noted that the OLS residual $\hat{u}_i$ satisfies:

$$\begin{aligned}\ln\hat{Y}_i &= \hat{\beta}_0 + \sum_{n=1}^{k} \hat{\beta}_n \ln X_{ni} + \hat{u}_i \\ &= \hat{\beta}_0 + \max_i(\hat{u}_i) + \sum_{n=1}^{k} \hat{\beta}_n \ln X_{ni} + \hat{u}_i - \max_i(\hat{u}_i) \\ &= \hat{\beta}_0^* + \sum_{n=1}^{k} \hat{\beta}_n \ln X_{ni} + \hat{u}_i - \max_i(\hat{u}_i)\end{aligned}$$

Now define $\hat{u}_i^*$ as

$$\begin{aligned}\hat{u}_i^* &= -\left[\ln\hat{Y}_i - \hat{\beta}_0^* - \sum_{n=1}^{k} \hat{\beta}_n \ln X_{nii}\right] \\ &= \max_i(\hat{u}_i) - \hat{u}_i\end{aligned} \tag{10.4}$$

which is non-negative for all i and zero for at least one i. The $\hat{u}_i^*$ is the COLS residual which may be used to obtain consistent estimate of individual firm's *TE* as

$$TE_i^* = \exp\left(-\hat{u}_i^*\right)$$

Applications of this technique include Belbase and Grabowski (1985) which estimates a deterministic Cobb-Douglas production frontier model to assess the efficiency in Nepalese agriculture; Seaver and Triantis (1989) fits a Cobb-Douglas production frontier to obtain plant and process-level technical efficiency of the linerboard manufacturing facilities in the U. S. However, one limitation of this technique is that there is no guarantee that the COLS frontier would be the closest possible boundary which can envelope the input–output data. Since it is required to be parallel to the estimated OLS regression equation this might impose undesirable restrictions on the production technology.[1]

Goal Programming Approach

The Aigner and Chu (1968) formulation shows that the deterministic production frontier can be expressed as a mathematical programming model where the goal would be to find a set of optimal values of the parameters $(\beta_0; \beta_1; \ldots \beta_k)$ such that the observed output of no producer exceeds the maximum output feasible for its input vector.

[1] see Kumbhakar and Lovell (2000), pp. 70–71, for details.

$$\beta_0 + \sum_{n=1}^{k} \beta_n \ln X_{ni} \geq \ln Y_i; \tag{10.5}$$

$$\forall i = 1; 2; \ldots; N$$

As Eq. (10.2) shows, for each firm i, u_i is the difference between its frontier output and observed output (both in logarithms) and hence, is non-negative, by Eq. (10.5). Optimal β's are obtained satisfying the set of constraints Eq. (10.5), by either minimizing the sum of the u_i's (i.e.,$\sum_i u_i$) (in which case the problem is solved by linear programming), or the sum of the squares of the u_i's (i.e. $\sum_i u_i^2$,) (in which case the problem is solved by quadratic programming). Once parameter values are calculated from either model, the TE of each producer can be computed from the (optimal) slack in the corresponding constraint:

$$TE_i = \exp\left(-u_i^*\right) = \exp\left(\beta_0^* + \sum_{n=1}^{k} \beta_n^* \ln X_{ni} - \ln Y_i\right) (i = 1, 2, \ldots, N) \tag{10.6}$$

where an asterisk (*) corresponding to a variable (parameter) indicates the optimal value of the variable (parameter) in question.

Some important applications of this technique include Levin (1974) to account for technical inefficiency in education; Shapiro and Müller (1977) to measure technical efficiency through a deterministic Cobb-Douglas production frontier constructed using linear programming method; Van den Broeck et al. (1980) to solve a nonlinear programming problem to construct the frontier for the data on the general milk processing from 28 individual dairy plants collected from the Swedish Dairy Federation for the period 1964–1973; Bjurek et al. (1990) to estimate a Cobb-Douglas frontier by using both linear and quadratic programming techniques to analyze productive efficiency of about 400 local social insurance offices in Sweden using data for the period 1974–1984 collected from the National Social Insurance Board.

A major drawback of this approach is that the optimal parameter values are *calculated* rather than *estimated* and hence statistical inference concerning the calculated parameter values becomes difficult. Later development of the literature includes the work of Schmidt (1976) who points out that the models could be given a statistical interpretation, if a distributional assumption is imposed on the u_is. In fact, the calculated values can be shown to be the maximum likelihood estimates (MLE) of the parameters for the linear model if the $u_i (\geq 0)$ follow an exponential distribution and, for the quadratic model if the u_is follow a half-normal distribution.

Modified Ordinary Least Squares Method

Afriat (1972) and Richmond (1974) suggest that the parameters of the deterministic production frontier model (Equation-10.2) as well as the estimates of firm-level TEs could be obtained using a variant of the COLS method. The method, widely known as the MOLS method, is a two-step method, with OLS regression in the first step (as in the case of COLS method) and then modifying in the second step the OLS estimates of both intercept $\left(\hat{\beta}_0\right)$ and residuals $(\hat{u}_i)$ in the following way:

$$\hat{\beta}^{**} = \hat{\beta}_0 + E(\hat{u}_i) \tag{10.7}$$

and

$$\hat{u}_i^{**} = E(\hat{u}_i) - \hat{u}_i \tag{10.8}$$

The modification is done, using the *mean* of the estimates of a disturbance term, instead of its *maximum* value as is done in the case of the COLS method. Individual efficiency estimate could then be obtained as $TE_i^{**} = \exp\left(-\hat{u}_i^{**}\right)$.

A few applications include Rossi and Canay (2001) which estimates a frontier model using this method and compares its results with those obtained through alternative methods of estimation. They use four different data sets consisting of information on gas distribution firms in Argentina, water companies in Asia and Pacific region, water companies in England, and electricity distribution firms in South America.

This method also has a major limitation. Using the mean of u_i to modify the intercept of the frontier equation in order to shift it upwards may not ensure that all the observed data points are bounded by the estimated production frontier from above.[2]

Stochastic Frontier Analysis

Apart from the limitations of each individual method spelt out in the above discussion, all the deterministic methods are subject to a serious deficiency. Each measures TE relative to a deterministic production frontier, i.e., a frontier in which the maximum feasible output for a given vector of inputs is not affected by any random factors. Aigner et al. (1977) and Meeusen and van den Broeck (1977) independently proposed a model where an additional random error v is added to the non-negative random variable u to take care of the effect of such unobserved random factors as well as measurement error (if any) in the frontier output

[2] see Olson et al. (1980), for a detailed discussion.

framework. This is known as the Stochastic Frontier Production Function (SFPF) approach, since the proposed method takes the frontier itself to be stochastic, and is subject to random variations beyond the control of the producer.

The stochastic frontier is usually the preferred alternative as it is capable of accommodating random factors which affect output but are beyond a firm's control and possible misspecification owing to the omission of any relevant variables (uncorrelated with the included regressors). However, obtaining firm-level estimates of technical efficiency is an involved task and is not as straightforward as in the case of deterministic frontier. In case of a deterministic frontier such estimates are readily obtained from the estimated values of the residuals. In a stochastic framework, once the functional form is specified for the frontier, the two components of errors and their specifications regarding the behavioral and distributional assumptions play a complicated role in the derivation of the efficiency estimates.

Keeping the above factors in mind, alternative methodologies are proposed in the literature to obtain estimates of firm-level efficiency under SFPF. We summarize such discussions below assuming a flexible translog functional form (the one used most widely in the literature). The method followed is that of maximum likelihood estimation which is distribution sensitive. v_i and u_i are assumed to be distributed independently of each other.

The SFPF model is given by

$$\begin{aligned}\ln Y_i &= \beta_0 + \sum_{n=1}^{k} \beta_n \ln X_{ni} + \frac{1}{2}\sum_{m=1}^{k}\sum_{n=1}^{k} \beta_{mn}(\ln X_{mi})(\ln X_{ni}) + v_i - u_i \\ &= \left\{ \beta_0 + \sum_{n=1}^{k} \beta_n \ln X_{ni} + \frac{1}{2}\sum_{m=1}^{k}\sum_{n=1}^{k} \beta_{mn}(\ln X_{mi})(\ln X_{ni}) + v_i \right\} - u_i \qquad (10.9)\end{aligned}$$

where the expression within the second brackets represents (the logarithmic value of) the frontier output and the coefficients β_{mn}s are assumed to be symmetric, i.e., $\beta_{mn} = \beta_{nm} \forall m \,\&\, n$. Note that the frontier itself is stochastic in nature since it contains a random error term, v_i which is generally assumed to follow a normal distribution with zero mean and constant variance. The estimation of the Eq. (10.9) also requires specific distributional assumption for the one-sided inefficiency error component u_i.

A number of alternative distributions have been suggested for characterizing the u_is in the literature. The non-negativity constraint on u_i restricts the set of distributions that can be assumed to the ones falling in the positive quadrant. These include distributions for which the density of the disturbances has the closest concentration near zero, like a slightly inflexible half-normal distribution (proposed in the original Aigner et al. (1977) model) or an exponential distribution (proposed in the Meeusen and Van den Broeck (1977) model). In contrast, Stevenson (1980) suggested shifting the half-normal distribution to obtain a nonzero mode, producing thereby a general truncated normal distribution. He also proposed a restricted version of Gamma model in his chapter. Greene (1990) had proposed a

two-parameter Gamma distribution where a COLS estimator, obtained by using the method of moment's estimation, had also been presented.

Among the distributions spelt out above, the exponential distribution is the simplest to assume. But since it cannot be approximated as a normal distribution which is useful in the later stage for desirable statistical properties, exponential distribution for the unobserved inefficiency errors might not be a good assumption. The immediate choice would be variants of truncated normal distribution, which is the most general case. One variant is the half-normal distribution. If u_i follows a half-normal distribution, it implies that the mean of the inefficiency errors distribution is assumed to be zero. Now this might be an oversimplification which can be overcome by assuming a truncated normal distribution for which before truncation the mean of the distribution can be nonzero but the truncation can be at the point zero.

With a truncated normal distribution the deficiency is that we cannot deal with inter-firm variations in inefficiency as we have assumed a constant mean for all the firms' inefficiencies. If we include variability in means, that is to say all the firms have different expected values of the mean of their efficiency with a truncated distribution model, there arises an indeterminacy problem due to a number of parameters exceeding the number of equations to be solved. This problem is taken care of in the conditional mean model which suggests accounting for the inter-firm differences through a relationship between mean of each firm's inefficiency and a variable z_i which affects inefficiency. This method suggests a simultaneous determination of the frontier parameters as well as the parameters specified to characterize the firm-level inefficiency through z_i. Generally the inefficiency errors are assumed to be linear in z_i. Hence we have additional parameters in the system which can be solved with a minimal loss of degrees of freedom. This can be applied for models where the inter-firm variation is high. One way to test the existence of high variation in inefficiency errors is to run an OLS and check the residuals. If the properties of these residuals are similar to those of the OLS residuals, the random error part in the model is more prominent and we can assume lesser variation in the inefficiency error component and use the truncated normal model. Otherwise, the conditional mean model can be applied.

Let us refer to some of the important work in the literature in this context. Estimation of observation-specific TE was a difficult task until the early 1980s. Waldman (1982) shows that there can be problems in working with the likelihood function if a truncated normal distribution is assumed for the inefficiency term u_i. Another limitation of this truncated normal distribution model is that it would fail to identify u_i and v_i separately at the firm level. Jondrow et al. (1982) were the first to develop a method of segregating the inefficiency term u_i from the statistical noise term v_i using the distribution of the u_i conditional on the estimate of the composite error term $\varepsilon_i (= v_i - u_i)$ for each individual unit i. An empirical application of this method for the case of 111 privately owned steam electric generating plants in the U. S. is also presented in the chapter. To get a clearer idea about this issue, we discuss below estimation strategies along with the corresponding

likelihood functions and analytical expressions for the observation-specific estimates of TE under a few alternative distributional assumptions.

As far as the conventional random term (v_i) in case of SFPF, it is assumed to be a normal variate, identically and independently distributed with zero mean and constant variance, i.e., $v_i \sim iid - N\left(0, \sigma_v^2\right)$ (for all *i)*, irrespective of the assumptions made about the distribution of the inefficiency variable, u_i. As far as the inefficiency variable (u_i) is concerned, let us first consider the case where u_i follows an *exponential distribution*, being identically and independently distributed with variance σ_u^2. The log likelihood function for a sample of N producers may then be written as

$$\ln L = -N \ln \sigma_u + N\left(\frac{\sigma_v^2}{2\sigma_u^2}\right) + \sum_{i=1}^{N} \ln \Phi(A_i) + \sum_{i-1}^{N} \left(\frac{\varepsilon_i}{\sigma_u}\right) \tag{10.10}$$

where $A_i = \left(\widetilde{\mu_i}/\sigma_v\right)$, $\underset{i}{\widetilde{\mu}} = -\varepsilon_i - \left(\frac{\sigma_v^2}{\sigma_u}\right)$ (for all i) and $\Phi(\cdot)$ is the distribution function of a standard normal variate (i.e., normal variate with zero mean and unit variance). Again, conditional density of the u_i, given the value of ε_i, i.e., $f(u_i|\varepsilon_i)$ can be shown to be the density of a normally distributed variable but truncated below at zero, with (prior to truncation) mean $\underset{i}{\widetilde{\mu}}$ and the variance σ_v^2. A firm specific estimate of TE can then be obtained using the expression given below.

$$TE_i = E[\exp(-u_i)|\varepsilon_i] = \frac{\left[1 - \Phi\left(\sigma_v - \frac{\widetilde{\mu_i}}{\sigma_v}\right)\right]}{\left[1 - \Phi\left(-\frac{\widetilde{\mu_i}}{\sigma_v}\right)\right]} \exp\left(-\widetilde{\mu_i} + \frac{\sigma_v^2}{2}\right) \tag{10.11}$$

To cite a few applications, Van den Broeck et al. (1980) estimate the composite error model for the Swedish dairy industry, assuming an exponentially distributed inefficiency variable. In a later study Cummins and Zi (1998) assume exponential distribution for the inefficiency disturbance term to estimate a cost frontier and then measure cost efficiency of the US life insurance industry using a variety of econometric and mathematical programming techniques.

Consider now the *general* truncated normal model which assumes that for all *i,* $u_i \sim iid - N\left(\mu, \sigma_u^2\right)$. The log likelihood function for a sample of N producers in this case can be written as follows:

$$\ln L = \text{CONSTANT} - \frac{\sum_{i=1}^{N} (\mu + \varepsilon_i)^2}{2\left(\sigma_u^2 + \sigma_v^2\right)} + \sum_{i=1}^{N} \ln\left\{\Phi\left(\frac{\mu_i^*}{\sigma^*}\right)\right\} - \frac{1}{2}\ln\left(\sigma_u^2 + \sigma_v^2\right) - N \ln\left\{\Phi\left(\frac{\mu}{\sigma_u}\right)\right\} \tag{10.12}$$

where $\mu_i^* = \frac{\mu\sigma_v^2 - \varepsilon_i\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$ and $\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma_u^2 + \sigma_v^2}$. The firm-level estimate of *TE* is given by

$$TE_i = E[\exp(-u_i)|\varepsilon_i] = \frac{\left[1 - \Phi\left(\sigma_* - \frac{\mu_i^*}{\sigma^*}\right)\right]}{\left[1 - \Phi\left(-\frac{\mu_i^*}{\sigma^*}\right)\right]} \exp\left(-\mu_i^* + \frac{\sigma_*^2}{2}\right) \tag{10.13}$$

To cite a few applications, a frontier model with composite disturbance structure (as proposed by Aigner et al. 1977) is fitted for earning functions in labor markets by Herzog et al. (1985), Hofler and Polachek (1985) and by Robinson and Wunnava (1989). Assuming a half-normal distribution for the one-sided inefficiency variable Ali and Flinn (1989) measure efficiency among Basmati rice producers of Pakistan Punjab relative to an estimated normalized profit frontier. Using two alternative methods, viz., OLS and MLE Hunt-McCool and Warren (1993) estimated semi-log linear earning frontiers, both deterministic as well as stochastic, to find labor market efficiency. They assumed the one-sided inefficiency variable to follow a gamma distribution in case of a deterministic frontier, but a half-normal distribution in case of a stochastic frontier model. Recently, Fuwa et al. (2007) also used this model to estimate the individual-level technical efficiency of rice farms of Chhotanagpur Plateau in eastern India. Cavalluzzo and Baldwin (1993) estimate stochastic frontier models to obtain productivity differences across union and non-union projects in office building construction, assuming both a half-normal as well as an exponential distribution for the inefficiency variable, u_i.

In the presence of high variation in the technical efficiency amongst firms the immediate question after estimating the technical efficiency scores of the firms is to find the determinants of these inter-firm differences. There have been attempts to explain inter-unit differences in the levels of *TE*s in terms of economic factors. Some early exercises on this include Kalirajan (1981) and Pitt and Lee (1981) who followed a two-stage method to identify the factors responsible for the inter-firm differences in *TE* within an industry. They estimate a stochastic frontier production function and also the firm-specific *TE* scores at the first stage. At the second stage, these efficiency scores are regressed on a number of firm-specific factors like size, age, ownership pattern, level of education, managerial experience of a firm, and so on which are likely to affect its *TE*. Since these firm-specific factors (say, z_i for firm i) that explain technical inefficiency of a firm vary across firms, u_i, the inefficiency error component for the ith firm, should be assumed to be distributed independently *but not* identically across i. Here we assume that the distribution of u_i differs only in its mean, i.e., $u_i \sim iid - N^+(\mu_i, \sigma_u^2)$ but its variance remains constant across firms, as was assumed in the cases discussed earlier. However such a two-stage procedure suffers from serious inconsistency problem, as the model estimated at the first stage is misspecified. The results obtained would therefore be biased.

A one-step procedure based on a correctly specified model for the distribution of *Yi*, given the input vector *Xi* and the vector of firm-specific factors affecting its *TE* has been proposed in the literature. Such a one-step procedure has been proposed by Kumbhakar et al. (1991) and Reifschneider and Stevenson (1991). They

specify stochastic frontier models in which the inefficiency effects are defined to be an explicit function of the z_i. The associated parameters along with those of the frontier production function are then estimated through a single-stage maximum likelihood estimation method. Such a model can be represented by adding the following relation to the Eq. (10.9).

$$\mu_i = z_i\delta \tag{10.14}$$

where δ is the vector of parameters associated with the firm specific factors, z_i. The non-neutral model proposed by Huang and Liu (1994) is applicable to the case when some of the input variables of the model appear in the z_i. The alternative models which have been developed in the literature may be seen as alternative ways of specifying the z_i s. For example, if the first element of z_i is unity with nonzero-associated parameter while the remaining parameters in δ are all zero, the general truncated normal distribution for the u_i proposed by Stevenson (1980) and Battese and Coelli (1988) would be obtained. The half-normal distribution originally proposed by Aigner et al. (1977) would be obtained, if each of the elements of δ is zero. Obviously, *Yi* should not be included in the z_i. For example, if one variable in z_i is a measure of firm size, it may be defined in terms of some variable (say, the amount of an input used) but not in terms of output (see Wang and Schmidt, 2002, p. 130).

With relation (10.14) added to Eq. (10.9), the log likelihood function and estimates of the firm-specific *TE* are as follows[3]:

$$\ln L = \text{CONSTANT} - \frac{\sum_{i=1}^{N}(z_i\delta+\varepsilon_i)^2}{2\left(\sigma_u^2+\sigma_v^2\right)} + \sum_{i=1}^{N}\left\{\Phi\left(\frac{\mu_i^{**}}{\sigma^*}\right)\right\} - \frac{1}{2}\ln\left(\sigma_u^2+\sigma_v^2\right) - \sum_{i=1}^{N}\ln\left\{\Phi\left(\frac{z_i\delta}{\sigma_u}\right)\right\} \tag{10.15}$$

where; $\mu_i^{**} = \frac{z_i\delta\sigma_v^2-\varepsilon_i\sigma_u^2}{\sigma_u^2+\sigma_v^2}$

$$TE_i = E[\exp(-u_i)|\varepsilon_i] = \frac{\left[1-\Phi\left(\sigma_* - \frac{\mu_i^{**}}{\sigma^*}\right)\right]}{\left[1-\Phi\left(-\frac{\mu_i^{**}}{\sigma^*}\right)\right]}\exp\left(-\mu_i^{**}+\frac{\sigma_*^2}{2}\right) \tag{10.16}$$

The similarity among the equations given in Eqs. (10.11), (10.13), and (10.16) for the estimates of firm-level TEs and between Eqs. (10.12) and (10.15) for the likelihood functions may be noted. This model has been applied widely in the recent past in a number of economic activities, e.g., agriculture (Battese and Broca, 1997; Coelli and Battese, 1996; Wilson et al. 2001), marine fishing (Sharma and Leung, 1998), industry (Hjalmarsson et al. 1996; Lundvall and Battese 2000; Bhandari and Maiti 2007), electricity distribution (Hattori, 2002), transport (Coelli

[3] Equations (10.17–10.21) in the Appendix give the details for a half-normal model.

et al. 1999), and so on. Several other studies have also used this model. Nourzad (2002) deserves special mention which finds out the effects of real money balance on production efficiency for ten developed and ten developing countries.

In this connection we may briefly review studies which have been done to estimate levels of technical efficiency (*TE*) for Indian industries. Some of the studies are based on the data collected through surveys specifically designed for this purpose (e.g., Little et al.1987; Page 1984). Many of the studies are concerned with estimating and explaining variations in *TE* in small-scale industries by fitting either a deterministic or a stochastic production frontier (e.g., Bhavani 1991; Goldar 1985; Neogi and Ghosh 1994; Nikaido 2004; Ramaswamy 1994). A review of some other studies in this area may be found in Goldar (1988).

All the studies mentioned above, however, use data relating to years prior to the beginning of economic reforms. For instance, Bhavani (1991) uses the data collected under the first Census of Small Scale Industrial Units, 1973 to estimate *TE* of firms at the four 4-digit level industries of metal product groups by fitting a (deterministic) translog production frontier. Similarly, on the basis of the data collected by the Second All India Census of Small scale Industrial Units, 1987–88, Nikaido (2004) fits a single stochastic production frontier, considering firms under all the (two-digit) industry groups and using intercept dummies to distinguish between different industry groups. Neogi and Ghosh (1994) examine the inter-temporal movement of *TE* using industry-level summary data for the years 1974–1975 to 1987–1988. The studies by Goldar et al. (2004) and Lall and Rodrigo (2001), however, relate to the post-reform era. Using the panel data for 63 firms in the engineering industry for 10 years from 1990–1991 to 1999–2000 drawn from the Prowess database 2001 version of the Centre for Monitoring Indian Economy (CMIE), Goldar et al. (2004) fit a translog stochastic production frontier to estimate firm-level *TE* scores in each year. At the second stage they attempt to explain variation in *TE* in terms of some economic variables like export and import intensity, degree of vertical integration, etc. Lall and Rodrigo (2001) examine *TE* variation across four industrial sectors in India during the year 1994 along with examining *TE* in relation to scale, location, extent of infrastructure investment, and some other determinants. Bhandari and Maity (2012) in a similar kind of exercise estimated *TE* score of the Indian leather industry using the approaches of both SFPF and DEA. The impact of trade policies on the technical efficiency of the Indian manufacturing sector can be found in Roy Biswas and Ghose (2012a (Chap. 6), b), who have considered the impact of policy variables like real effective exchange rate, import penetration ratio, import coverage ratio on technical efficiency.

Evolution of Nonparametric Optimization Approach: Data Envelopment Analysis

The parametric approach requires the imposition of a specific functional form for a production frontier and some assumptions like independently and identically normally distributed errors which have to be uncorrelated with the independent variables. In contrast the nonparametric approach does not require any functional form. It is based on a set of behavioral assumptions regarding production. Taking information from data on inputs and outputs the DEA method generates a discrete piecewise frontier by optimizing on *each* individual observation given the set of Pareto-efficient DMUs or peers. The technical efficiency scores are derived as the ratio of the actual output to the ideal output specified by the generated frontier.

DEA is based on mathematical programming. It is around the same time that Koopmans and Debreu independently conceive of the notion of technical efficiency in an optimization framework in two different but related contexts. Koopmans (1957) recognizes the connection between the existence of non-negative prices and quantities in a Walras-Cassel economy which leads to the problem of optimizing an objective function subject to a set of linear inequality constraints. A point in the commodity space is defined as efficient whenever an increase in the net output for one commodity requires a decrease in that of the other. Because of its similarity with the notion of Pareto optimality, this definition is known as the Pareto–Koopmans condition for *TE*. In the same year Debreu (1951), from the perspective of cost of resources, introduced the concept of *coefficient of resource utilization* as a measure of *TE* for the economy as a whole and interpreted any deviation of this measure from unity as a deadweight loss for the society on account of inefficient utilization of resources.

A few years later, Farrell (1957) observed that the efficiency of a firm which can reflect the ability to achieve the maximum level of output attainable by the state of technology or to use the inputs in optimal proportions, given their respective prices, can be analyzed in a diagrammatic framework of radial contraction of inputs/expansion of outputs from an observed point to the frontier. The main contribution of Farrell in his seminal work is to provide a simple but structured intuition to derive efficiency at the firm level using the principles of programming in a diagrammatic framework with iterative search for peer group and a piecewise discrete frontier. An assumption of constant returns to scale (CRS) technology in production is made. Hoffman (1957) suggests that the dual simplex method, an algorithm to solve a linear programming (LP) problem, can be applied to obtain Farrell's measure of efficiency. Farrell in his later work with Fieldhouse (Farrell and Fieldhouse 1962) attempted to solve the problem using the dual simplex algorithm where the case of increasing returns to scale is also considered. However, the problem is that the objective function is a fractional one and thus the tools of LP are not sufficient to provide a satisfactory solution.

Charnes, Cooper, and Rhodes (Charnes et al. 1978) contribute to convert the fractional programming into a linear programming by selecting suitable weights

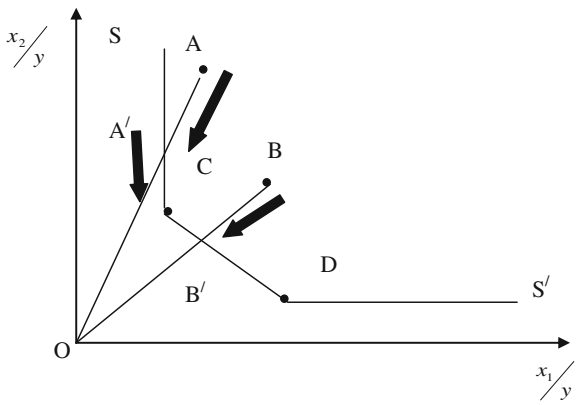


Fig. 10.1 Input-oriented radial and slack efficiency

which are nothing but the virtual prices of inputs and outputs. In their subsequent work in 1979 and 1981 a generalized DEA in a multiple-output multiple-input framework is established under a CRS assumption. The imposition of a CRS structure for the production technology implicitly assumes that producing units operate on optimal scales. As the different firms operate in different market structures, externalities, and different financial restrictions this assumption can be a handicap to analyzing real situations.

The variable returns to scale (VRS) model developed by Banker, Charnes, and Cooper (Banker et al. 1984) is able to decompose the *technical inefficiency* and *scale inefficiency* by defining and estimating the former at a given scale of operation under the assumption of a unique optimum (Maindiratta 1990). The Banker et al. (1984) model is the most general in DEA, which is used widely in the economics and management literature; it incorporates a constraint in the Charnes et al. 1978 framework to accommodate VRS and thus integrates the concept of technical and scale efficiencies together in a standard LP model.

The next challenge is to incorporate unrealized output potential and/or avoidable input waste in the framework for analyzing efficiency. In case of technical efficiency there are two distinct dimensions: first, whether the firm has selected the correct *technique* of production, i.e., choosing the correct ray (in two-dimension input-plane), and second, if the correct technique is chosen whether the *scale* has been selected optimally or not (i.e., the exact location on the ray). Since an efficient firm has to achieve both, the exact sequencing of choice does not have much bearing on the final outcome. To select an appropriate scale one seeks the maximum equi-proportionate increase in all outputs or decrease in all inputs which is known as radial efficiency. However, the radial projection of an observed input–output bundle onto the frontier does not necessarily exhaust the potential for

[4] For output-oriented technical efficiency the interpretation of radial and slack inefficiencies will just be reversed.

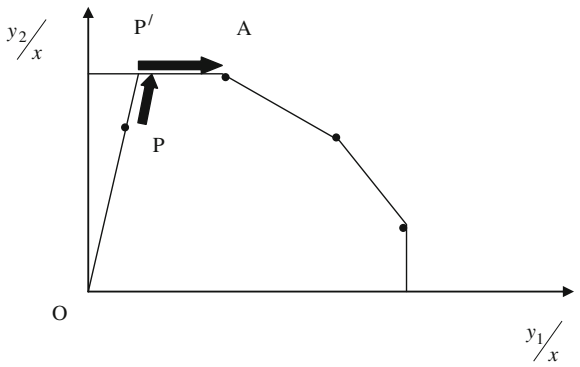


Fig. 10.2 Output-oriented radial and slack efficiency

expansion in all outputs or reduction in all inputs. In this case one needs some non-radial movement along with production frontier to reach the efficient point. This movement calls for a change in input (output)-proportions. Hence, two types of slacks are encountered (a) input slacks and (b) output slacks. Figures 10.1 and 10.2 explain the notions of input-oriented and output-oriented technical efficiencies with corresponding illustrations of radial as well as on-frontier movements to illustrate the notion of both technique-related inefficiency and scale-related inefficiency.[4]

Suppose we have four firms A, B, C, and D of which C and D are efficient firms. Now with the help of observed input–output data the piecewise linear isoquant (SS′) can be constructed (Fig.10.1)[5] Firms A and B represent two inefficient firms. So, the extent of their technical inefficiency will be $\frac{OA'}{OA}$ and $\frac{OB'}{OB}$ respectively. However, it is not the ultimate efficient point because one could reduce the input x_2 by the amount CA′and still produce the same output. Therefore, this movement along isoquant is known as the input slack. On the other hand in case of firm B only radial movement is enough to ensure efficient input–output combination. No slack movement is required here.

Similarly, consider Fig.10.2[6] where a firm operates at P, a point inside the production possibility frontier. The position is clearly inefficient as through a radial movement the firm may operate at P′and produce greater amounts of both y_1 & y_2. This distance from P to P′is called the radial inefficiency. However, it is interesting to note that even at point P′the firm has scope to increase its output by using the same level of inputs. A movement along the frontier from P′to A will help it to increase the amount of y_1 without sacrificing that of y_2, and, hence, the value of total output would certainly be higher at A compared to that at P′. This distance from P′to A is known as output slack inefficiency.

---

[4] For output-oriented technical efficiency the interpretation of radial and slack inefficiencies will just be reversed.

[5] Here x_1 & x_2 indicate two inputs and y represents one output.

[6] Here y_1 & y_2 indicate two outputs and x represents one input.

Formulation of Standard DEA Problem

This section spells out the steps in the formulation of the estimation technique in standard DEA. The nonparametric models can be computed using linear or integer programming. The computation of DEA is based on linear programming method. The computer solves as many linear programs as there are firms. In what follows we will elaborate the steps involved in the formulation of a problem in the DEA framework.

In the traditional DEA model, production technology with the following properties is hypothesized:

(i) The production possibility set is convex, i.e., if (x^0, y^0) and (x^1, y^1) are both feasible input–output bundles then $(x^{/}, y^{/})$ is also a feasible bundle where $x^{/} = \lambda\, x^0 + (1 - \lambda)\, x^1$, $y^{/} = \lambda\, y^0 + (1 - \lambda)\, y^1$, $0 \leq \lambda \leq 1$.

(ii) Inputs and outputs are freely disposable, i.e., if $(x, y) \in$ T then $(x^{/}, y) \in$T when $x^{/} \geq x$ and $(x, y^{/}) \in$ T when $y^{/} \leq y$.

(iii) When a sample of input–output bundles $(x^{\mathrm{i}}, y^{\mathrm{i}})$ is observed for N firms, we assume further that $(x^i, y^i)\ \varepsilon$ T for $i = 1,2,\ldots..,N$.

(iv) The technology satisfies variable returns to scale.

We select $T^{v} = \{(x, y) : x \geq \Sigma\lambda_j x^{j;}\ y \leq \Sigma\lambda_j y^{\mathrm{j}};\ \Sigma\lambda_j = 1;\ \lambda_j \geq 0,\ j = 1(1)n\}$, the smallest of all the sets satisfying assumptions (i)–(iv). This is the inner approximation of the underlying technology set.

Let there be N firms each producing m outputs from n inputs. Firm t uses input bundle $x^{\mathrm{t}} = (x_{1\mathrm{t}},\ldots..,x_{\mathrm{nt}})$ to produce output bundle $y^{\mathrm{t}} = (y_{1\mathrm{t}},\ldots..,y_{\mathrm{nt}})$. We use vector of virtual prices of inputs and outputs u_{it} and v_{jt} respectively, and get the average productivity of firm t as:

$$AP_t = \frac{\sum_{j=1}^{m} v_{jt} y_{jt}}{\sum_{i=1}^{n} u_{it} x_{it}}$$

The production relation has to satisfy the constraint stated as $\sum_{j=1}^{m} v_{jt} y_{jt} \leq \sum_{i=1}^{n} u_{it} x_{it} \forall t$ along with non-negativity restriction on virtual prices. So, the problem is:

$$\begin{aligned} \text{Max} : AP_t &= \frac{\sum_{j=1}^{m} v_{jt} y_{jt}}{\sum_{i=1}^{n} u_{it} x_{it}} \\ \text{Subject to} : \sum_{j=1}^{m} v_{jt} y_{jt} &\leq \sum_{i=1}^{n} u_{it} x_{it};\ \forall t = 1, 2, \ldots, N \\ u_{it}, v_{jt} &\geq 0. \end{aligned} \qquad (\text{Model} - 10.1)$$

This is a fractional programming problem. A price normalization constraint can be incorporated by virtue of which it can be written as:

$$\begin{aligned}
\text{Max}: & \sum_{j=1}^{m} p_{jt} y_{jt}, \\
\text{Subject to}: & \sum_{j=1}^{m} p_{jt} y_{jt} - \sum_{i=1}^{n} w_{it} x_{it} \leq 1 \ \forall t = 1, 2, \ldots, N \\
& \sum_{i=1}^{n} w_{it} x_{it} = 1, \ w_{it}, p_{jt} \geq 0 \ \forall t = 1, 2, \ldots, N
\end{aligned} \qquad (\text{Model} - 10.2)$$

This is a standard Linear Programming Problem and here $w_{it} = \lambda u_{it}$ and $p_{jt} = \lambda v_{jt}$with $\lambda > 0$.

Therefore, the dual of Model-10.2 can be written as:

$$\begin{aligned}
& \text{Min: } \theta \\
\text{Subject to}: & \sum_{t=1}^{N} \lambda_t y_{jt} \geq y_{jt}; \ \forall j = 1, 2, \ldots, m \\
& \sum_{t=1}^{N} \lambda_t x_{it} \leq \theta x_{it}; \ \forall i = 1, 2, \ldots, n \\
& \lambda_t \geq 0; \ \forall t = 1, 2, \ldots, N \\
& \text{Where}: \theta \text{ is free}
\end{aligned} \qquad (\text{Model} - 10.3)$$

From Model-10.3 one can estimate θ which is nothing but the input-oriented technical inefficiency score of tth firm under CRS assumption. Again, if we define $\varphi = \frac{1}{\theta}$ and $\mu_t = \frac{\lambda_t}{\theta}$ then minimization of θ is equivalent with maximization of φ. In terms of redefined variable the LP problem (Model-10.3) now becomes

$$\begin{aligned}
& \text{Max}: \phi \\
\text{Subject to}: & \sum_{t=1}^{N} \mu_t y_{jt} \geq \varphi y_{jt}; \ \forall j = 1, 2, \ldots, m \\
& \sum_{t=1}^{N} \mu_t x_{it} \leq x_{it}; \ \forall i = 1, 2, \ldots, n \\
& \mu_t \geq 0; \ \forall t = 1, 2, \ldots, N \\
& \text{Where}: \phi \text{ is free}
\end{aligned} \qquad (\text{Model} - 10.4)$$

The score generated from the expression $\frac{1}{\varphi}$ is nothing but the output-oriented technical efficiency of the tth firm under CRS. These two models consist of the first generation model of efficiency score measurement proposed by CCR.

The (in)efficiency measurement with additional constraint $\sum_{t=1}^{N} \lambda_t = 1$ in Model-10.3 and $\sum_{t=1}^{N} \mu_t = 1$ in Model-10.4 gives as the BCC model which considers the VRS assumption instead of CRS assumption.

Both CCR and BCC models calculate only radial (in) efficiency. For radial and slack calculation together one has to use the third generation models given below as extensions of Model-10.3 and Model-10.4.

Model-10.5 is the input version of the efficiency with slacks given as:

$$\text{Min: } \tilde{\theta} = \theta - \varepsilon\left(\sum_{j=1}^{m} s_j^+ + \sum_{i=1}^{n} s_i^-\right)$$

$$\text{Subject to : } \sum_{t=1}^{N} \lambda_t y_{jt} - s_j^+ = y_{jt};\ \forall j = 1, 2, \ldots, m$$

$$\sum_{t=1}^{N} \lambda_t x_{it} + s_i^- = \theta x_{it};\ \forall i = 1, 2, \ldots, n$$

$$\lambda_t, s_j^+, s_i^- \geq 0;\ \forall t = 1, 2, \ldots, N;\ \forall j = 1, 2, \ldots, m;\ \forall i = 1, 2, \ldots, n.$$

Where : θ is free.

(Model − 10.5)

Model 10.6 below gives the output version of the same as follows:

$$\text{Max: } \tilde{\phi} = \phi + \varepsilon\left(\sum_{j=1}^{m} s_j^+ + \sum_{i=1}^{n} s_i^-\right)$$

$$\text{Subject to : } \sum_{t=1}^{N} \mu_t y_{jt} - s_j^+ = \phi y_{jt};\ \forall j = 1, 2, \ldots, m$$

$$\sum_{t=1}^{N} \mu_t x_{it} + s_i^- = x_{it};\ \forall i = 1, 2, \ldots, n$$

$$\mu_t, s_j^+, s_i^- \geq 0;\ \forall t = 1, 2, \ldots, N;\ \forall j = 1, 2, \ldots, m;\ \forall i = 1, 2, \ldots, n.$$

Where : ϕ is free

(Model − 10.6)

s_j^+, s_i^-, indicates the output and input slacks and ε is any pre-assigned positive number, however small. Positive sign means output should be increased and negative sign means input should be decreased (Ray 2004).

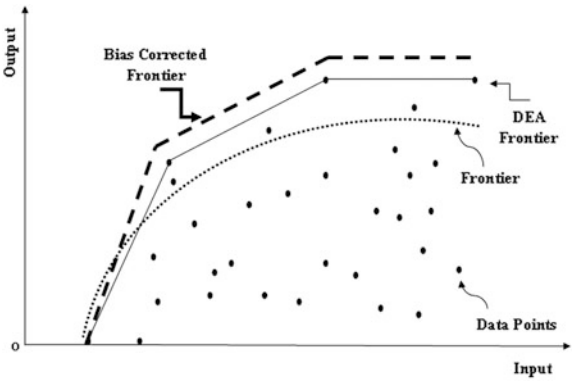


Fig. 10.3 Correction of bias in DEA

Statistical Foundations of DEA Models

One major drawback of the DEA model is that the estimates do not take into account sources of random variations and thus do not satisfy desirable statistical properties. Elements of random variations can be from four possible sources: random sampling of observations inside the production possibility set (*sampling error*), random sampling of observations outside the production possibility set (*outliers*), random outcome of production process (*stochastic technology*) or random measurement errors, omitted variables, and other disturbances (*stochastic noise*).Some common misconceptions include confusing stochastic noise with sampling variation, outliers, or stochastic technology. Also, sometimes statistical inference on sampling error is believed to improve robustness to noise. Another possible mistake is to treat robustness to outliers in the same way as robustness to noise (or at least closely related).

A review of development of statistical foundations of DEA shows that the aspect of sampling error has been explored in Banker (1993), Korostelev et al. (1995), Kneip et al. (1998), Simar and Wilson (2000a,b). We can summarize that in cases of deterministic technology, with no outliers or noise and the data being randomly sampled from the production possibility set, the DEA frontier converges to the true frontier as the sample size approaches to infinity. However in a finite sample, DEA frontier is downward biased (Fig. 10.3). Statistical inference on sampling error is possible by using asymptotic sampling distribution (Banker 1993) or Bootstrapping (Simar and Wilson 1998). However, such inferences have nothing to do with outliers, stochastic technology, or stochastic noise.

The purpose of the smooth consistent bootstrap (Simar and Wilson 1998, 2000a, b) is to mimic the original random sampling to estimate the sampling bias. Once corrected for bias, DEA frontier will lie above the original DEA frontier. In noisy data, there is a tendency that DEA will overestimate the frontier. Assuming away noise, and correcting for the small sample bias by bootstrapping, the frontier will shift upward. If noise is a problem, then bias correction will make it worse.

The problems related to outliers are addressed in the super-efficiency approach (Wilson 1995), context-dependent DEA (Seiford and Zhu 1999), robust efficiency measures or what is known as efficiency depth (Kuosmanen and Post 1999, Cherchye et al. 2000), and conditional order-m and order-α quantile frontiers (Aragon et al. 2005; Cazalse et al. 2002; Daouia and Simar 2007). The general opinion is that in case of a deterministic technology, robustness in the presence of outliers can be improved by not enveloping the most extreme observations. It is important to note that outliers are different from noise, and affect all observations, while outliers do not.

Land et al. (1993), Olesen and Petersen (1995), Cooper et al. (1996) Huang and Li (2001) address the problems of stochastic technology. Huang and Li (2001) assume inputs and outputs as multivariate normal random variables, with known expected values and covariance matrix to give a framework for chance constrained DEA models. However, knowledge about the expected values of inputs and outputs remains a question which cannot be estimated from cross-sectional data. In a panel data estimation would require that the true inputs and outputs do not change over time. Regarding variances and covariances of the error terms, also the uncertainty of the parameter estimates is not taken into account in the model. These aspects make the model very complicated to address real situations.

Some of the important models that have dealt with noise in the DEA framework are DEA + Gstach (1998), Banker and Natarajan (2008), Stochastic DEABanker et al. (1991), Stochastic FDH/DEA estimators Simar and Zelenyuk (2006), Stochastic Nonparametric Envelopment of Data (StoNED) by Kuosmanen (2006); Kuosmanen and Kortelainen (2007). To estimate a fully deterministic frontier based on data affected by noise, the shape of the frontier can be estimated without parametric assumptions, but the estimation of inefficiency (efficiency scores) becomes challenging in cross-section data if the observed output contains the noise term. Only conditional expected value can be estimated. Even the SFA efficiency estimator is not consistent in the presence of noise. In cross-section data, identifying inefficiency and noise requires some strong assumption, the strongest of all being assuming away the noise altogether. However, it has been shown that distributional assumptions do not influence the efficiency rankings (Ondrich and Ruggiero 2001).

Banker and Natarajan (2008) have established a two-stage methodology as a nonparametric counterpart of stochastic production frontiers. Their analysis claims that the estimates generated by a two-stage method comprising DEA in the first stage and an OLS or Tobit in the second stage can satisfy some desirable statistical properties not ensured by the nonparametric DEA approach. The model is developed in a framework where the set of input variables used in stage 1 and contextual variables in stage 2 are independent. It is also shown that the model works well, with a slight loss in the degree of robustness, in the presence of low negative correlation, and reasonably high positive correlations between the two

[7] Data for the pre-globalization period could not be used due to non-comparable data collection procedures in the ASI framework.

sets of variables mentioned above. It is also established that DEA followed by OLS outperforms all the one-stage as well as two-stage procedures. This is a simple formulation with known basic frameworks used in both the stages. Although not full proof this framework can address real situations and are comparatively easier to handle.

An Application

This section elaborates on applications of SFA and DEA to the garments sector in India. We have used ASI unit level data for NIC 18 classification at the five-digit level for alternate years starting from 1999–2000 till 2005–2006.[7] We use three inputs (viz., intermediate input, fixed assets, man-days worked) and one output (viz., ex-factory value of outputs) framework for the purpose. The exercise intends to track the change, if any, in technical efficiency in the sector in the post globalization era and analyzed various aspects of technical efficiency and its relationship with state-specific location and firm-level characteristics like size and age. It is to be noted that in the present context the results per se might not be very important as the main objective is to spell out the steps to be taken to use each of the techniques.

The starting point of any empirical analysis is the summary statistics of the data used for the same. It is clear from Table 10.1 that for all the years there is a considerable variation in the sample.

Application of Stochastic Frontier Analysis

We use a Translog Production Function for SFA. The first step is to check the OLS residual plots. Figure 10.4 shows the residual plots in OLS.

We find that there is a concentration of OLS residuals around zero. This implies that the effect of disturbance factor other than classical error is possibly rather weak in the present data. To get a clearer idea we do a histogram analysis of these residuals in Fig. 10.5. We find that the histogram of OLS residuals resembles a standard normal distribution. Maximum observations are within the range between −1.5 and +1.5. This is consistent with the previous observation. Table 10.2 gives the description of OLS residuals which shows a negative skewness indicating a mild presence of inefficiency in the data.

[8] As the variation of u is low, the estimation of technical efficiency under conditional mean model cannot be successfully performed for our sample of firms in the Garments Sector. There are convergence problems. .

Table 10.1 Summary statistics of output and input variables

Years	Statistics	OUT	INTRINP	FA	MDW	AGE
1999–2000	Obs	504	504	504	504	504
	Mean	104,131,652	62,134,359	21,932,206	69,525	10
	Median	49,789,313	30,542,680	6,963,219	34,541	7
	Skewness	6	7	7	6	4
	Kurtosis	53	72	82	60	30
	Max	2,288,191,437	1,599,884,694	733,782,000	1,420,206	100
	Min	34,600	19,285	7,376	285	1
	SD	190,194,464	116,401,188	52,223,368	105,972	11
	CV	2	2	2	2	1
2001–2002	Obs	821	821	821	821	821
	Mean	103,083,691	64,245,564	22,160,961	64,622	10
	Median	45,935,770	27,660,157	7,679,025	31,212	8
	Skewness	9	9	8	13	3
	Kurtosis	121	125	88	266	16
	Max	3,624,426,253	2,378,734,879	662,845,012	2,723,663	87
	Min	14,685	29,186	13,622	250	1
	SD	211,099,519	138,692,462	50,568,494	123,444	9
	CV	2	2	2	2	1
2003–2004	Obs	918	918	918	918	918
	Mean	115,295,376	69,484,538	25,673,055	71,899	11
	Median	54,286,090	32,466,694	9,357,901	32,569	8
	Skewness	8	8	8	8	2
	Kurtosis	116	119	84	129	12
	Max	3,756,780,801	2,320,723,530	807,455,452	2,367,990	84
	Min	5,850	172,294	1,429	358	1
	SD	204,976,287	125,919,819	61,662,825	125,400	9
	CV	2	2	2	2	1
2005–2006	Obs	1,028	1,028	1,028	1,028	1,028
	Mean	167,534,637	100,423,564	34,117,530	95,927	10
	Median	71,342,181	41,816,345	12,149,144	41,478	8
	Skewness	8	7	9	10	2
	Kurtosis	98	82	98	152	10
	Max	5,767,225,141	3,113,075,166	1,282,362,557	3,702,801	86
	Min	13,464	78,338	3,876	72	1
	SD	349,647,589	210,036,423	89,225,209	212,018	9
	CV	2	2	3	2	1

Source Author's calculation from unit level ASI data

Table 10.3 presents the results of SFA corresponding to a half-normal and an exponential distribution of inefficiency errors component. We would like to analyze the case for half-normal distribution in detail, as the exponential case is pretty straightforward.[8]

[9] There are some tests performed to check certain conditions regarding the applicability of the model. The rationale for these tests and the results for our application are spelt out in Appendix 2.

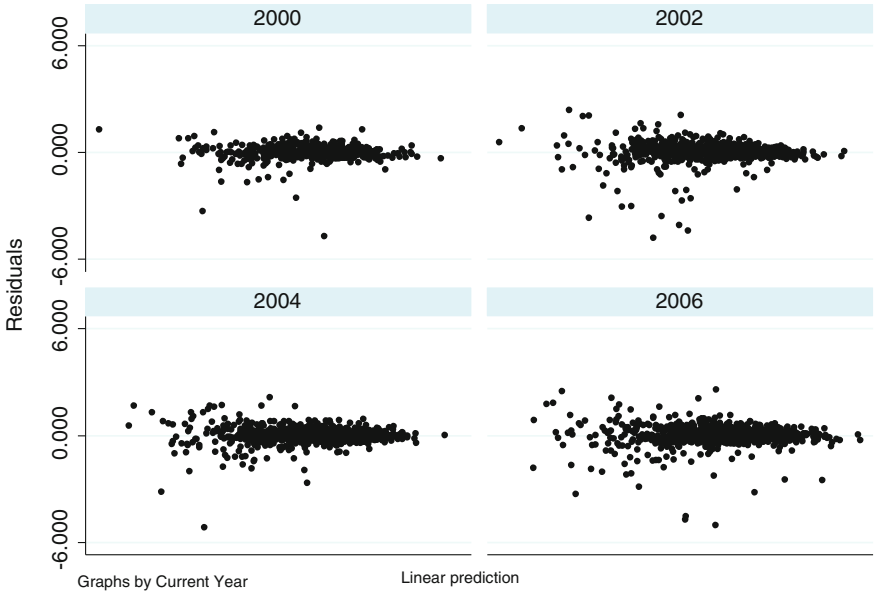


Fig. 10.4 Scatter plots of OLS residuals. *Source* Author's calculation from unit level ASI data

Table 10.4 reports the summary statistics of the inefficiency errors with a half-normal distributional assumption. We find that the mean of the inefficiency factor remains below 0.53. An even lower median (than mean) for all the years and thus a positive skewness supports the presence of low level of concentration of inefficiency error.

We also cross check with the histogram and fitted distribution curve for each year (Fig. 10.6) which approximates a half-normal distribution. This signifies that the distributional assumption on inefficiency error is not incorrect.

Table 10.5 reports the summary statistics of the firm-level technical efficiency scores for the half-normal model[9] in the four alternate years. The average efficiency is high, nearly 0.70. Median value is always greater than mean indicating concentration at high value. Negative skewness in a histogram (Fig. 10.7) makes this point clearer.

We have also done an analysis of the distribution of garments firms in India across states (Table 10.6) and the technical efficiency scores across states (Table 10.7). Tamil Nadu, Karnataka, and Delhi are identified as the highest

[10] Size: Each year individual firms are arranged in *ascending* order of size (measured by values of intermediate inputs used by them) and then the firms are classified into different quartile groups like the lowest 25 % denoted by *Very Small*, the next 25 % denoted by *Small,* the third quartile denoted by *Large*, and the highest 25 % denoted by V*ery Large*.

Age: Here classification of firms in terms of three age groups, namely, *Very Old, Old*, and *Young* is proposed utilizing information on whether the firms were established before 20 years, between 10 and 20 years and within 10 years' time, respectively.

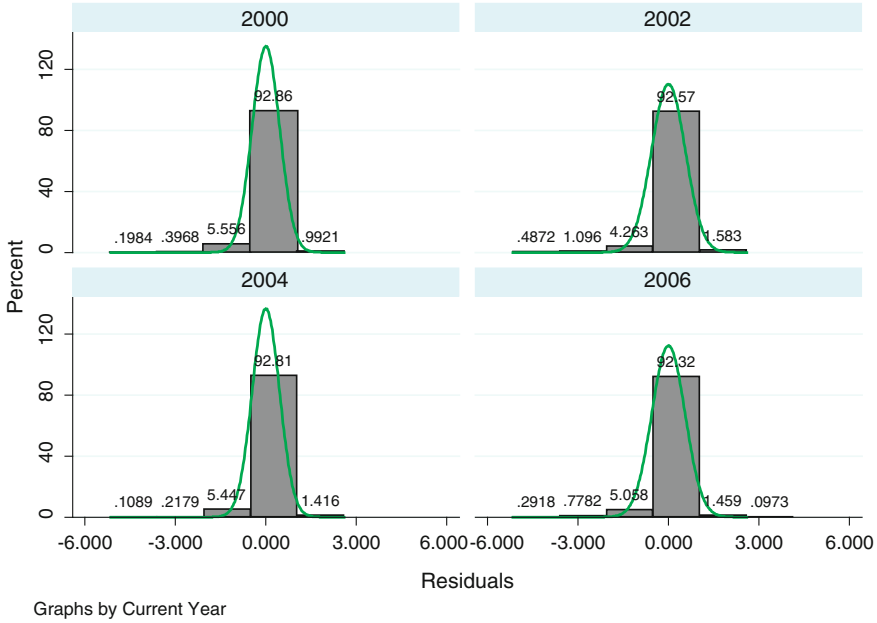


Fig. 10.5 Histogram of OLS residuals. *Source* Author's calculation from unit level ASI data

Table 10.2 Summary statistics of OLS residuals

Statistics	Years			
	1999–2000	2001–2002	2003–2004	2005–2006
OBS	504	821	918	1,028
Mean	0.000	0.000	0.000	0.000
Median	0.016	0.013	0.006	0.018
Skewness	−3.424	−3.036	−2.164	−2.593
Kurtosis	32.311	25.278	27.345	24.631
Minimum	−4.707	−4.788	−5.149	−5.036
SD	0.458	0.561	0.452	0.550
Maximum	1.391	2.371	2.129	2.590

Source Author's calculation from unit level ASI data

concentrating zone (with more than 15 % firms). Among the selected eight major states very low concentration is found in Rajasthan.

The average efficiency is high in all the states (not less than 0.65). Delhi, Rajasthan, and Uttar Pradesh (Low Concentration State) perform significantly well among other states.

Keeping in mind the probable impact of size and age on technical efficiency we have divided the firms into four categories according to size and three categories according to age.[10] We have analyzed the *TE* scores for each of these size

Table 10.3 Stochastic frontier results

Years	1999–2000						2001–2002					
No. of observation	504						821					
Variables	Half-normal			Exponential			Half-normal			Exponential		
ln_out	β-Coef.	z	P > z	β-Coef.	Z	P > z	β-Coef.	z	P > z	β-Coef.	Z	P > z
ln_intrinp	1.137	6.570	0.000	1.143	7.250	0.000	1.032	4.940	0.000	0.909	5.180	0.000
ln_fa	0.114	0.860	0.390	0.092	0.790	0.427	0.026	0.160	0.873	0.051	0.400	0.689
ln_mdw	−0.320	−1.660	0.098	−0.291	−1.730	0.084	−0.110	−0.530	0.598	0.084	0.490	0.627
ln_intrinp2	−0.007	−0.690	0.492	−0.011	−1.230	0.217	−0.024	−2.940	0.003	−0.012	−1.600	0.110
ln_fa2	0.021	3.040	0.002	0.022	3.550	0.000	−0.007	−1.080	0.279	−0.004	−0.880	0.377
ln_mdw2	−0.011	−0.760	0.448	−0.006	−0.450	0.655	−0.050	−3.510	0.000	−0.031	−2.550	0.011
ln_intfa	−0.030	−2.210	0.027	−0.023	−1.900	0.057	0.007	0.550	0.582	0.006	0.530	0.593
ln_intmdw	0.052	2.870	0.004	0.054	3.390	0.001	0.060	3.690	0.000	0.031	2.140	0.032
ln_famdw	−0.022	−1.350	0.178	−0.033	−2.150	0.031	0.008	0.610	0.542	0.002	0.200	0.838
_cons	0.352	0.250	0.804	0.243	0.200	0.843	1.395	0.820	0.413	1.105	0.820	0.415
log likelihood	−266.500			−208.918			−595.830			−470.776		
/lnsig2v	−2.920	−20.390	0.000	−2.826	−26.360	0.000	−2.520	−23.060	0.000	−2.648	−31.760	0.000
/lnsig2u	−1.027	−10.070	0.000	−2.430	−16.250	0.000	−0.633	−7.970	0.000	−1.996	−18.550	0.000
sigma_v	0.232			0.243			0.284			0.266		
sigma_u	0.598			0.297			0.729			0.369		
sigma2	0.412			0.147			0.612			0.207		
lambda	2.577			1.219			2.569			1.385		

(Continued)

Table 10.3 (Continued)

Year	2003–04						2005–06					
No. of Observation	918						1028					
Variables	Half-Normal			Exponential			Half-Normal			Exponential		
ln_out	β-Coef.	z	P > z	β-Coef.	Z	P > z	β-Coef.	z	P > z	β-Coef.	Z	P > z
ln_intrinp	1.917	10.310	0.000	1.586	9.100	0.000	1.355	7.640	0.000	1.010	6.210	0.000
ln_fa	−0.207	−1.760	0.078	−0.084	−0.820	0.413	−0.492	−3.760	0.000	−0.352	−3.030	0.002
ln_mdw	−0.471	−2.580	0.010	−0.322	−1.920	0.055	−0.381	−2.040	0.042	−0.094	−0.560	0.574
ln_intrinp2	−0.047	−5.960	0.000	−0.031	−4.030	0.000	−0.069	−7.990	0.000	−0.044	−5.180	0.000
ln_fa2	0.000	0.070	0.947	0.004	0.780	0.434	−0.003	−0.740	0.456	−0.003	−0.810	0.418
ln_mdw2	−0.030	−2.630	0.009	−0.024	−2.320	0.020	−0.030	−3.430	0.001	−0.022	−2.910	0.004
ln_intfa	0.007	0.770	0.440	−0.002	−0.200	0.844	0.063	6.060	0.000	0.049	5.170	0.000
ln_intmdw	0.054	3.930	0.000	0.044	3.240	0.001	0.098	7.010	0.000	0.065	4.860	0.000
ln_famdw	0.010	0.770	0.441	0.004	0.340	0.738	−0.045	−3.440	0.001	−0.036	−3.270	0.001
_cons	−2.859	−2.240	0.025	−1.817	−1.590	0.113	3.995	3.260	0.001	4.256	3.950	0.000
log likelihood	−525.256			−437.253			−760.561			−635.416		
/lnsig2v	−2.523	−24.390	0.000	−2.677	−31.890	0.000	−2.292	−25.110	0.000	−2.433	−33.960	0.000
/lnsig2u	−1.166	−11.810	0.000	−2.348	−20.080	0.000	−0.744	−9.380	0.000	−2.034	−20.120	0.000
sigma_v	0.283			0.262			0.318			0.296		
sigma_u	0.558			0.309			0.689			0.362		
sigma2	0.392			0.164			0.576			0.219		
lambda	1.971			1.179			2.168			1.220		

Table 10.4 Summary statistics of "u"

Statistics	Years			
	1999–2000	2001–2002	2003–2004	2005–2006
Mean	0.443	0.527	0.421	0.511
Median	0.383	0.442	0.364	0.436
SD	0.356	0.45	0.302	0.395
Kurtosis	54.727	36.878	57.817	39.824
Min	0.055	0.052	0.049	0.048
Max	4.684	4.658	4.608	4.625
Skewness	5.699	5.122	5.625	5.108

Source Author's calculation from unit level ASI data

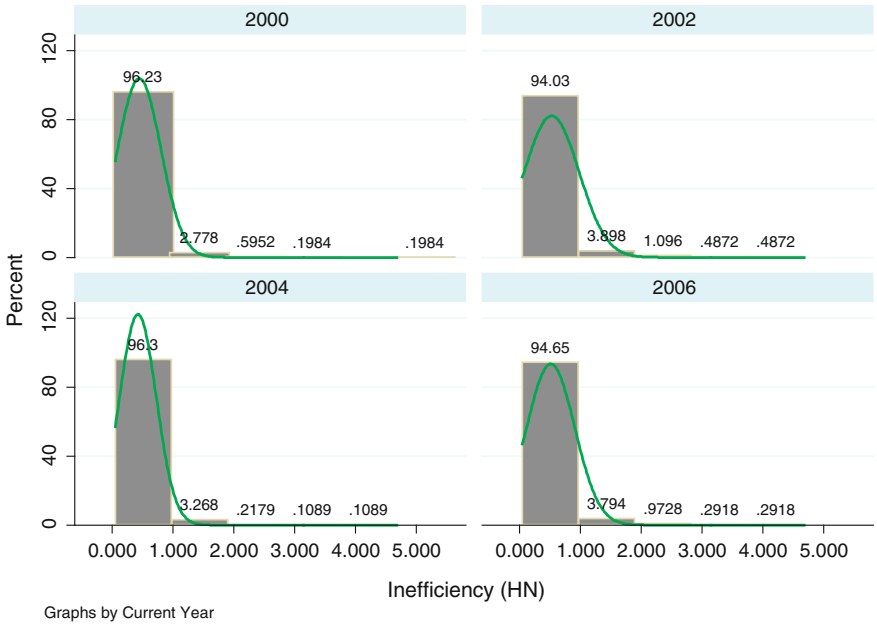


Fig. 10.6 Histogram of "u". *Source* Author's calculation from unit level ASI data

Table 10.5 Summary statistics of "technical efficiency"

Statistics	Years			
	1999–2000	2001–2002	2003–2004	2005–2006
Mean	0.679	0.641	0.690	0.646
Median	0.695	0.660	0.710	0.666
SD	0.144	0.149	0.130	0.142
Kurtosis	5.847	6.825	7.286	6.920
Min	0.009	0.010	0.010	0.010
Max	0.948	0.951	0.953	0.954
Skewness	−1.250	−1.514	−1.609	−1.525

Source Author's calculation from unit level ASI data

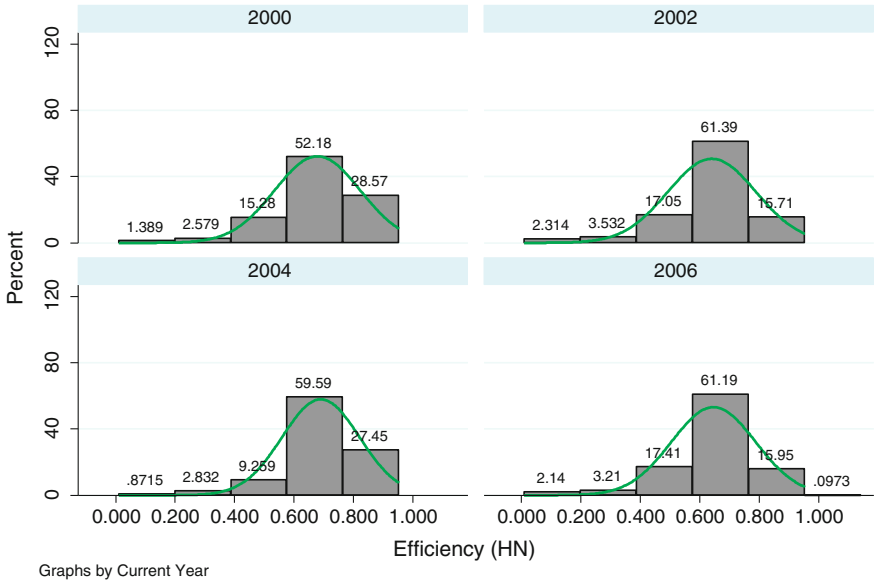


Fig. 10.7 Histogram of "*TE* Score." *Source* Author's calculation from unit level ASI data

Table 10.6 State-wise distribution of firms

States	1999–2000	2001–2002	2003–2004	2005–2006
Delhi	15.08	18.64	16.12	12.94
Gujarat	4.56	3.05	3.05	3.50
Haryana	6.55	8.89	8.50	12.16
Karnataka	21.43	21.07	17.65	14.79
Maharashtra	10.52	6.70	10.35	9.34
Rajasthan	1.98	2.80	3.49	3.11
Tamil Nadu	26.59	26.07	22.44	21.79
Uttar Pradesh	5.95	6.58	8.28	12.06
Other	7.34	6.21	10.13	10.31
Total	100.00	100.00	100.00	100.00

Source Author's calculation from unit level ASI data

(Table 10.8) and age (Table 10.9) categories for all the years and also a two-way table for these scores taking size and age together (Table 10.10).

We find that Very Large and Very Old groups always have higher efficiency scores. On the other hand Very Small and Young firms reveal low scores during the selected time period. However, there are not many differences as far as the TEs of other groups are concerned.

Table 10.7 State-wise average of "technical efficiency"

State	1999–00	2001–02	2003–04	2005–06
Delhi	0.72	0.70	0.75	0.69
Gujarat	0.62	0.63	0.70	0.62
Haryana	0.66	0.63	0.69	0.66
Karnataka	0.67	0.60	0.67	0.62
Maharashtra	0.70	0.67	0.68	0.64
Rajasthan	0.74	0.71	0.75	0.66
Tamil Nadu	0.67	0.63	0.67	0.64
Uttar Pradesh	0.70	0.62	0.69	0.67
Other	0.65	0.60	0.666	0.61

Source Author's calculation from unit level ASI data

Table 10.8 Size-wise average of "technical efficiency"

Firm Size	1999–2000	2001–2002	2003–2004	2005–2006
Very small	0.641	0.584	0.642	0.585
Small	0.704	0.644	0.706	0.667
Large	0.686	0.672	0.704	0.673
Very large	0.685	0.664	0.710	0.661

Source Author's calculation from unit level ASI data

Table 10.9 Age-wise average of "technical efficiency"

Age group	1999–2000	2001–2002	2003–2004	2005–2006
Young	0.678	0.637	0.686	0.647
Old	0.681	0.641	0.694	0.647
Very old	0.683	0.662	0.706	0.643

Source Author's calculation from unit level ASI data

Application of Data Envelopment Analysis

We have applied the Banker et al. (1984) model to derive the scores in a generalized framework with slacks. Table 10.11 records the summary statistics of DEA scores for the four alternate years. Here we find that the mean efficiency is always less than 0.50. Median is slightly lesser than mean indicating concentration at low level. Slightly positive skewness supports this observation. (See histogram in Fig. 10.8).

In DEA methodology we can also get the distribution of efficient and inefficient firms in the sector. Table 10.12 elaborates on the percentage distribution. The main observation suggests that the percentage of efficient firms is very low over the entire period (not more than 5 %) and it is falling until 2006.

One of the interesting insights from DEA is testing for the existence of input and output slacks. In our exercise the main observations suggest that very few inefficient firms have slack in the intermediate input. Also, the percentage of firms

Table 10.10 Size–age-wise average of "technical efficiency"

Years	Firm size	Young	Old	Very old
1999–2000	Very small	0.635	0.662	0.636
	Small	0.717	0.676	0.682
	Large	0.687	0.678	0.695
	Very large	0.670	0.704	0.730
2001–2002	Very small	0.579	0.601	0.586
	Small	0.639	0.643	0.695
	Large	0.669	0.667	0.696
	Very large	0.666	0.644	0.697
2003–2004	Very small	0.642	0.636	0.651
	Small	0.696	0.717	0.741
	Large	0.705	0.697	0.713
	Very large	0.706	0.707	0.733
2005–2006	Very small	0.583	0.580	0.597
	Small	0.683	0.651	0.610
	Large	0.667	0.675	0.701
	Very large	0.653	0.665	0.688

Source Author's calculation from unit level ASI data

Table 10.11 Summary statistics of "technical efficiency"

Statistics	1999–2000	2001–2002	2003–2004	2005–2006
Mean	0.48	0.47	0.41	0.40
Median	0.45	0.44	0.39	0.37
SD	0.21	0.23	0.20	0.24
Skewness	0.61	0.41	0.85	0.61
Kurtosis	3.43	2.76	4.01	2.95
CV	0.43	0.50	0.50	0.59

Source Author's calculation from unit level ASI data

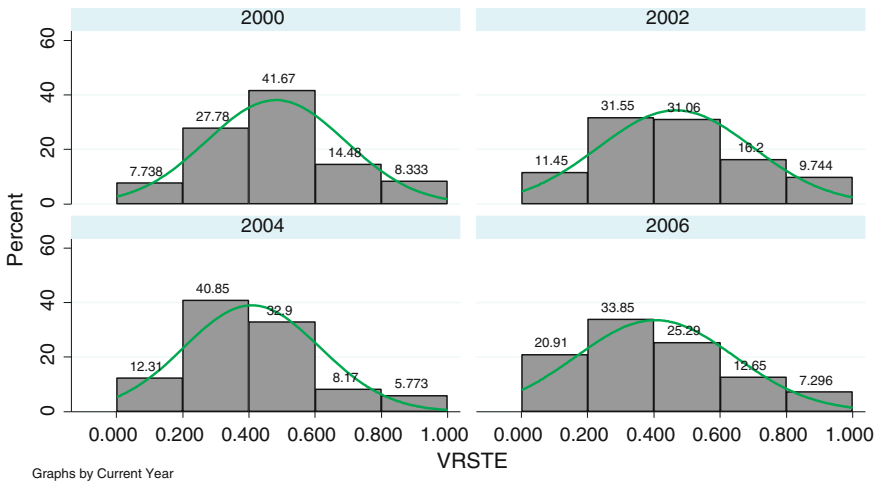


Fig. 10.8 Histogram of "TE Scores". *Source* Author's calculation from unit level ASI data

Table 10.12 Percentage distribution of inefficient and efficient firms

Years	Inefficient	Efficient	Total
1999–2000	95.04	4.96	100.00
2001–2002	96.10	3.90	100.00
2003–2004	97.39	2.61	100.00
2005–2006	96.50	3.50	100.00

Source Author's calculation from unit level ASI data

Table 10.13 Percentage of input slack of inefficient firms

Years	INTRINP	FA	MDW	Output-only
1999–2000	1.25	5.43	16.49	78.71
2001–2002	0.00	40.81	42.84	29.02
2003–2004	1.68	53.13	56.38	9.62
2005–2006	1.01	39.52	24.19	42.54

Source Author's calculation from unit level ASI data

Table 10.14 State-wise average of "technical efficiency"

State	1999–2000	2001–2002	2003–2004	2005–2006
Delhi	0.601	0.587	0.484	0.541
Gujarat	0.397	0.478	0.399	0.338
Haryana	0.460	0.483	0.426	0.431
Karnataka	0.436	0.393	0.391	0.327
Maharashtra	0.531	0.526	0.423	0.459
Rajasthan	0.479	0.570	0.406	0.380
Tamil Nadu	0.438	0.418	0.368	0.341
Uttar Pradesh	0.550	0.417	0.396	0.453
Other	0.472	0.455	0.398	0.368

Source Author's calculation from unit level ASI data

Table 10.15 State-wise percentage distribution of frontier firms

State	1999–2000	2001–2002	2003–2004	2005–2006
Delhi	36.00	37.50	12.50	22.22
Gujarat	0.00	3.13	4.17	0.00
Haryana	0.00	6.25	0.00	5.56
Karnataka	16.00	12.50	16.67	2.78
Maharashtra	12.00	9.38	20.83	19.44
Rajasthan	0.00	6.25	0.00	2.78
Tamil Nadu	12.00	12.50	16.67	16.67
Uttar Pradesh	8.00	0.00	4.17	19.44
Other	16.00	12.50	25.00	11.11

Source Author's calculation from unit level ASI data

Table 10.16 Size-wise average of "technical efficiency"

Firm Size	1999–2000	2001–2002	2003–2004	2005–2006
Very small	0.413	0.338	0.264	0.268
Small	0.446	0.371	0.381	0.372
Large	0.490	0.480	0.436	0.422
Very large	0.577	0.671	0.554	0.555

Source Author's calculation from unit level ASI data

Table 10.17 Size-wise percentage distribution of frontier firms

Firm size	1999–2000	2001–2002	2003–2004	2005–2006
Very small	28.00	37.50	33.33	30.56
Small	24.00	6.25	16.67	27.78
Large	16.00	18.75	4.17	8.33
Very large	32.00	37.50	45.83	33.33

Source Author's calculation from unit level ASI data

Table 10.18 Age-wise average of "technical efficiency"

Age Group	1999–2000	2001–2002	2003–2004	2005–2006
Young	0.465	0.447	0.394	0.403
Old	0.529	0.486	0.433	0.406
Very old	0.478	0.529	0.432	0.408

Source Author's calculation from unit level ASI data

having slack in labor is always greater than that having slack in capital excepting for 2006. Table 10.13 gives the results of the slack analysis.

We have also done an analysis of the technical efficiency scores across states (Table 10.14). We find that among the selected states, Delhi, Maharashtra, and Uttar Pradesh perform better in terms of efficiency scores but these scores are slightly falling over the period. On the other hand the other states like Gujarat, Haryana, Karnataka, etc., show a stable low average efficiency score during this period.

For DEA we can also attempt an analysis of the state-wise distribution of efficient firms which are on the frontier (Table 10.15). We find that for Delhi and Karnataka the representation of frontier firms is higher at the initial time period which falls over time. On the other hand the importance of three states (Maharashtra, Tamil Nadu, and Uttar Pradesh) in terms of frontier representation increases over time.

We have attempted an analysis of the efficiency scores and distribution of efficient firms across size classes using DEA (Tables 10.16 and 10.17). We find that Very Large firms as a group perform better than the others in terms of average efficiency score as well as the frontier representations.

We have also done similar analysis for the age groups of firms and taking age and size together. (Tables 10.18 and 10.19). Though the Young firms as a group do

Table 10.19 Age-wise percentage distribution of frontier firm

Age Group	1999–2000	2001–2002	2003–2004	2005–2006
Young	44.00	62.50	66.67	63.89
Old	52.00	12.50	12.50	25.00
Very old	4.00	25.00	20.83	11.11

Source Author's calculation from unit level ASI data

Table 10.20 Size and age-wise average of "technical efficiency"

Years	Firm size	Young	Old	Very old
1999–2000	Very small	0.388	0.492	0.399
	Small	0.451	0.463	0.371
	Large	0.481	0.497	0.534
	Very large	0.544	0.639	0.619
2001–2002	Very small	0.312	0.388	0.402
	Small	0.366	0.368	0.425
	Large	0.467	0.486	0.551
	Very large	0.671	0.650	0.719
2003–2004	Very small	0.254	0.270	0.306
	Small	0.371	0.407	0.392
	Large	0.434	0.424	0.469
	Very large	0.549	0.550	0.586
2005–2006	Very small	0.255	0.285	0.296
	Small	0.380	0.368	0.336
	Large	0.429	0.386	0.482
	Very large	0.560	0.544	0.564

Source Author's calculation from unit level ASI data

Table 10.21 Comparison of SFA and DEA scores

Years	Statistics	SFA	DEA	DIFF (%)
1999–2000	Mean	0.68	0.48	41
	Median	0.70	0.45	54
2001–2002	Mean	0.64	0.47	38
	Median	0.66	0.44	51
2003–2004	Mean	0.69	0.41	69
	Median	0.71	0.39	81
2005–2006	Mean	0.65	0.40	60
	Median	0.67	0.37	79

Source Author's calculation from unit level ASI data

not perform well in terms of average efficiency scores over the years, in terms of frontier representation they have a prominent role.

We can also attempt an analysis of the distribution of efficient firms according to size and age (Table 10.20). The average efficiency score is always high in Very

Large and Very Old group. On the other hand the average efficiency score is always low in Very Small and Young group. However, the average is falling over the period for both these groups.

We conclude this section on application by a comparison of the scores of the two methods applied (Table 10.21). We find that in terms of mean and median the SFA score is higher than DEA score throughout the selected years. The differences in percentage terms are highest in terms of mean as well as median in the year 2004. The mean (median) SFA score reported is higher by 69 % (81 %) of the DEA scores. However, from the year-wise histograms it is evident that the distributional pattern of SFA and DEA scores are similar.

Conclusions

The chapter brings together the evolution of two families of techniques used to estimate technical efficiency scores in the economics and management literature with an objective to give an overview of the two different schools of thoughts aimed towards the same goal. The nuances of the techniques are understood in the perspective of a practical problem by applying both the techniques in analyzing the impact of globalization on the Indian garments sector. We find that the impact, analyzed in the literature, cannot be traced through the technical efficiency scores. The pattern, whatever little, as seen from the results is not prominent enough to give any conclusive remarks. It deserves special mention that the elaboration of steps associated with each technique is more important than interpretation of results in this exercise as the main objective of the chapter is to give an elaborate account of the steps involved in the two estimation methods and understand the intuition behind them.

Standard packages like R, Frontier or STATA are used for the estimation of SFA models. Among the tests spelt out in Appendix 2, the LR test can be performed in STATA and the quasiconcavity test can be performed in R or Matlab. For standard DEA efficiency score estimation, Excel or DEAP can serve the purpose. For stochastic DEA the regression component can be tackled in STATA or any other standard package that can perform regressions.

When we talk about two alternative approaches for empirical estimation the immediate question that comes up is regarding the advantages and disadvantages of the two. We find that the weakness which the parametric school faces is a result of the assumption of a functional form and the distributional specifications of the inefficiency component of the error. Whereas the nonparametric school does not impose any functional form, it lacks the statistical foundation which is useful in analyzing random variations in real situations. The literature on integrating the two frameworks by providing statistical foundation to the nonparametric DEA is also

reviewed. We find that there has been serious research to design rigorous models to fulfill the gap due to non-stochasticity in basic DEA framework.

Then there is the question of choice. Is there a way to determine which approach is better than the other? Well, the answer would be that it is specific to the problem being addressed. It is not possible to judge from inspection of the data which model would fit better. There can be data-related characteristics which prove one technique to be more useful than the other. Sometimes the most general model of conditional mean in SFA with truncated distributions might face convergence problems in datasets having low variation in inefficiency errors. Later developments of DEA like chance constrained DEA incorporating stochastic technology, though general, might not be operational and solvable with all kinds of datasets.

However, one rule of thumb could be the size of the data determining the family of technique to be used. But in case the size of the data is such that two versions can be tried it would be more insightful to try both the versions and analyze the results. There are some standard models which are more widely used than others. In SFA the truncated normal model is widely used. In DEA the basic variable returns to scale model with slacks is the most widely used. Among the models which claim to have stochasticity incorporated in DEA, stochastic DEA and bootstrapping are the most widely used. However, in order to find out what is the best alternative there should be considerable effort on trial and error to try out different versions of the techniques..

Acknowledgments We thank Arijit Chakrabarty of ISI, Delhi, and Anup Bhandari of CDS, Trivandram, for their valuable inputs.

A.1 Appendix 1

A Note on Normal–Half-Normal Model

The basic distributional assumption in this case is:

(i) $v_i \sim iid - N(0, \sigma_v^2)$
(ii) $u_i \sim iid - N^+(0, \sigma_u^2)$
(iii) u_i & v_i – *Independent*

So the joint density of u & v is:

$$f(u,v) = \frac{2}{2\pi\sigma_u\sigma_v}\exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right\} \tag{10.17}$$

Now assume $\varepsilon = v - u$. Therefore the expression of joint density of u & ε is:

$$f(u,\varepsilon) = \frac{2}{2\pi\sigma_u\sigma_v}\exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{(\varepsilon+u)^2}{2\sigma_v^2}\right\} \tag{10.18}$$

The marginal density of ε is:

$$\begin{aligned} f(\varepsilon) &= \int_0^\infty f(u,\varepsilon)du \\ &= \frac{2}{\sqrt{2\pi}\sigma}\left[1-\Phi\left(\frac{\varepsilon\lambda}{\sigma}\right)\right]\exp\left\{-\frac{\varepsilon^2}{2\sigma^2}\right\} \\ &= \frac{2}{\sigma}\varphi\left(\frac{\varepsilon}{\sigma}\right)\Phi\left(-\frac{\varepsilon\lambda}{\sigma}\right) \end{aligned} \tag{10.19}$$

where; $\sigma = \left(\sigma_u^2+\sigma_v^2\right)^{1/2}$ and $\lambda = \frac{\sigma_u}{\sigma_v}$

The log likelihood function of the sample of N producer is:

$$\ln L = \text{CONS} - N\ln\sigma + \sum_{i=1}^{N}\ln\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\varepsilon_i \tag{10.20}$$

The conditional distribution of $u|\varepsilon$ is:

$$\begin{aligned} f(u|\varepsilon) &= \frac{f(u,\varepsilon)}{f(\varepsilon)} \\ &= \frac{1}{\sqrt{2\pi}\sigma_*}\exp\left\{-\frac{(u-\mu_*)^2}{2\sigma_*^2}\right\}\Bigg/\left[1-\Phi\left(-\frac{\mu_*}{\sigma_*}\right)\right] \end{aligned} \tag{10.21}$$

where; $\mu_* = -\frac{\varepsilon\sigma_u^2}{\sigma^2}$ and $\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma^2}$

Estimated Technical Efficiency of ith Firm:

$$\begin{aligned} TE_i &= E(\exp\{-u_i\}|\varepsilon_i) \\ &= \left[\frac{1-\Phi\left(\sigma_* - \mu_{*i}/\sigma_*\right)}{1-\Phi\left(-\mu_{*i}/\sigma_*\right)}\right]\exp\left\{-\mu_{*i}+\frac{1}{2}\sigma_*^2\right\} \end{aligned} \tag{10.22}$$

Table 10.22 Likelihood ratio test for half-normal frontier model

Inputs	Statistics	Years			
		1999–2000	2001–2002	2003–2004	2005–2006
Intermediate Inputs	*LR chi2(4)*	954.98	1453.32	1853.80	1758.39
	Prob > chi2	0.00	0.00	0.00	0.00
Fixed Assets	*LR chi2(4)*	23.00	3.86	9.77	37.53
	Prob > chi2	0.00	0.43	0.04	0.00
Mandays	*LR chi2(4)*	8.28	15.06	16.96	57.38
	Prob > chi2	0.08	0.00	0.00	0.00

Source Author's calculation from unit level ASI data

Table 10.23 Firms having positive input elasticity

	Years	Total no. of firms	No. of firms	% of firms
INTRINP	2000	504	504	100.00
	2002	821	821	100.00
	2004	918	918	100.00
	2006	1,028	1,028	100.00
FA	*2000*	504	420	83.33
	2002	821	714	86.97
	2004	918	845	92.05
	2006	1,028	688	66.93
MDW	2000	504	284	56.35
	2002	821	419	51.04
	2004	918	348	37.91
	2006	1,028	379	36.87

Source Author's calculation from unit level ASI data

A.2 Appendix 2

Some Conditions for SFPF Model

Likelihood Ratio Test

Suppose, we have a three input production function $Y = f(x, y, z)$. In order to estimate the SFPF model, the chapter uses fixed assets, intermediate input, and man-days worked as inputs. It may be possible that some of the inputs may be unimportant. This may be judged by performing statistical significance of the parameters associated with that particular input variable using likelihood ratio test. The higher the value of the LR statistic and the closer the value of the probability to zero, the better the fit. The results of our study are summarized in Table 10.22. It is only for the year 2001–2002 for the input Fixed Asset that we find the fit is not so good. Still, we continue taking this input for the year.

Monotonicity Requirement

Three conditions need to be satisfied:

(i) $f_x = \frac{\partial Y}{\partial x} \geq 0$; (ii) $f_y = \frac{\partial Y}{\partial y} \geq 0$; and (iii) $f_z = \frac{\partial Y}{\partial z} \geq 0$

Table 10.23 reports the number of firms having positive input elasticities for each of the inputs in our study. It is to be noted that these firms that have positive input elasticity with respect to an input would satisfy the Monotonicity Requirement for that input autmatically.

Table 10.24 Firms satisfying regularity property

Properties	Years	Total no. of firms	No. of firms	% of firms
Monotonicity	2000	504	212	42.06
	2002	821	348	42.39
	2004	918	333	36.27
	2006	1,028	372	36.19
Quasiconcavity	2000	504	496	98.41
	2002	821	807	98.29
	2004	918	902	98.26
	2006	1,028	980	95.33
Both	2000	504	210	41.67
	2002	821	340	41.41
	2004	918	331	36.06
	2006	1,028	341	33.17

Source Author's calculation from unit level ASI data

Quasiconcavity Requirement

For a three-input model, three conditions need to be satisfied [Notations carry their usual meanings]:

(a) $\begin{vmatrix} 0 & f_x \\ f_x & f_{xx} \end{vmatrix} \leq 0$, which is a trivial requirement and is satisfied always;

(b) $\begin{vmatrix} 0 & f_x & f_y \\ f_x & f_{xx} & f_{xy} \\ f_y & f_{xy} & f_{yy} \end{vmatrix} \geq 0$;

(c) $\begin{vmatrix} 0 & f_x & f_y & f_z \\ f_x & f_{xx} & f_{xy} & f_{xz} \\ f_y & f_{xy} & f_{yy} & f_{yz} \\ f_z & f_{xz} & f_{yz} & f_{zz} \end{vmatrix} \leq 0.$

We have to check the number of firms satisfying both (b) and (c) above. Table 10.24 reports the number of firms satisfying these requirements in our study.

A note of caution may apply. The estimated parameters need to be adjusted according to the form of the production function and the form in which the variables are used to derive the function for the empirical estimation. For instance, $f_x = \frac{\partial Y}{\partial x} = \frac{\partial \ln Y}{\partial \ln x} \times \frac{Y}{x} = \varepsilon_x^Y \times \frac{Y}{x}$; and similarly for $f_y = \varepsilon_y^Y \times \frac{Y}{y}$ and $f_z = \varepsilon_z^Y \times \frac{Y}{z}$.

Again, for the second order derivatives $f_{xx}, f_{xy}, \ldots, f_{zz}$ in the matrix under (b) and (c) above need to be adjusted. For instance, f_{xx} should be $\left[\beta_{11} + \left(\varepsilon_x^Y\right)^2 - \varepsilon_x^Y\right] \times \frac{Y}{x^2}$. Similarly, $f_{yy} = \left[\beta_{22} + \left(\varepsilon_y^Y\right)^2 - \varepsilon_y^Y\right] \times \frac{Y}{y^2}$ and $f_{zz} = \left[\beta_{33} + \left(\varepsilon_z^Y\right)^2 - \varepsilon_z^Y\right] \times \frac{Y}{z^2}$. Cross partials also need to be adjusted. In general, $f_{ij} = \left[\beta_{ij} + \varepsilon_i^Y \varepsilon_j^Y\right] \times \frac{Y}{x_i x_j} \; \forall \; i \neq j$; and $i, j = 1, 2, 3$.

References

Afriat SN (1972) Efficiency estimation of production functions. Int Econ Rev 13(3):568–598
Aigner DJ, Chu SF (1968) On estimating the industry production function. Am Econ Rev 58(4):826–839
Aigner DJ, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. J Econom 6(1):21–37
Ali M, Flinn JC (1989) Profit efficiency among basmati rice producers in Pakistan Punjab. Am J Agr Econ 71(2):303–310
Aragon Y, Daouia A, Thomas-Agnan C (2005) Nonparametric frontier estimation: a conditional quantile based approach. Econom Theory 21:358–389
Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manage Sci 30(9):1078–1092
Banker R, Datar S, Kemerer C (1991) A model to evaluate variables impacting the productivity of software maintenance projects. Manage Sci 37(1):1–18
Banker R, Gadh V, Gorr W (1993) A Monte Carlo comparison of production frontier estimation methods. European J Oper Res 37(3):332–343
Banker RD, Natarajan R (2008) Evaluating contextual variables affecting productivity using data envelopment analysis. Oper Res 56(1):48–58
Battese GE, Broca SS (1997) Functional forms of stochastic frontier production functions and models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan. J Prod Anal 8(4):395–414
Battese GE, Coelli TJ (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. J Econom 38(3):387–399
Belbase K, Grabowski R (1985) Technical efficiency in Nepalese agriculture. J Dev Areas 19(4):515–525
Bhandari AK, Maity P (2012) Efficiency of the Indian leather firms: some results obtained using two conventional methods. J Prod Anal 37:73–93
Bhandari AK, Maiti P (2007) Efficiency of Indian manufacturing firms: textile industry as a case study. Int J Bus Econ 6(1):71–88
Bhavani T (1991) Technical efficiency in Indian modern small scale sector: an application of frontier production function. Indian Econ Rev 26(2):149–166
Bjurek H, Hjalmarsson L, Førsund FR (1990) Deterministic parametric and non- parametric estimation of efficiency in service production. J Econom 46(1/2):213–227
Cavalluzzo L, Baldwin D (1993) Unionization and productive efficiency. In: Fried HO, Lovell CAK, Schmidt SS (eds) The measurement of productive efficiency. Oxford University Press, New York, pp 210–220
Cazals C, Florens JP, Simar L (2002) Nonparametric frontier estimation: a robustb approach. J Econom 106(2002):1–25
Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444
Cherchye L, Kuosmanen T, Post P (2000) Why convexify ? An assessment of convexity axioms In Dea, Helsingin Kauppakorkeakoulu Helsinki School Of Economics And Business Administration, Working Papers W-270
Coelli TJ, Battese GE (1996) Identification of factors which influence the technical inefficiency of Indian farmers. Australian J Agr Econ 40(2):103–128
Coelli TJ, Perelman S, Romano E (1999) Accounting for environmental influences in stochastic frontier models: with application to international airlines. J Prod Anal 11(3):251–273
Cooper WW, Huang ZM, Li SX (1996) Satisficing DEA models under chance constraints. Ann Oper Res 66:279–295
Cummins JD, Zi H (1998) Comparison of frontier efficiency methods: an application to the U. S. Life Insurance Industry. J Prod Anal 10(2):131–152

Daouia A, Simar L (2007) Nonparametric efficiency analysis: a multivariate conditional quantile approach. Open Access Publications from University of Toulouse 1 Capitole http://neeo.univ-tlse1.fr, University of Toulouse 1, Capitole

Debreu G (1951) The coefficient of resource utilization. Econometrica 19(3):273–292

Farrell MJ (1957) The measurement of productive efficiency. J R Stat Soc Series A (General) 120(3):253–290

Farrell MJ, Fieldhouse M (1962) Estimating efficient production frontiers under IRS. J R Stat Soc Series A 125(2):252–267

Fuwa N, Edmonds C, Banik P (2007) Are small-scale farmers in Eastern India really inefficient? Examining the effects of microtopography on technical efficiency estimates. Agr Econ 36(3):335–346

Gabrielsen A (1975) On estimating efficient production functions. Working Paper No. A-85, Christian Michelsen Institute, Department of Humanities and Social Sciences, Bergen, Norway

Goldar B (1985) Unit size and economic efficiency in small scale washing soap industry in India. Artha Vijnana 27(1):21–40

Goldar B (1988) Relative efficiency of modern small scale industries in India. In: Suri KB (ed) Small scale enterprises in industrial development. Sage Publication, New Delhi

Goldar B, Renganathan VS, Banga R (2004) Ownership and efficiency in engineering firms: 1990–91 to 1999–2000. Econ Polit Weekly 39(5):441–447

Greene WH (1990) A gamma-distributed stochastic frontier model. J Econom 46(1/2):141–163

Gstach D (1998) Small sample performance of two approaches to technical efficiency estimation with multiple outputs. Department of Economics, Working Papers from Vienna University of Economics, Department of Economics

Hattori T (2002) Relative performance of U. S. and Japanese electricity distribution: an application of stochastic frontier analysis. J Prod Anal 18(3):269–284

Herzog HW, Hofler RA, Schlottmann AM (1985) Life on the frontier: migrant information, earnings and past mobility. Rev Econ Stat 67(3):373–382

Hoffman AJ (1957) Discussion on Mr. Farrell's paper. J R Stat Soc Series A 120(3):284

Hofler RA, Polachek S (1985) A new approach for measuring wage ignorance in the labor market. J Econ Bus 37(3):267–276

Hjalmarsson L, Kumbhakar SC, Heshmati A (1996) DEA, DFA and SFA: a comparison. J Prod Anal 7(2–3):303–327

Huang ZM, Li SX (2001) Stochastic DEA models with different types of input–output disturbances. J Prod Anal 15:95–113

Huang CJ, Liu JT (1994) Estimation of a non-neutral stochastic frontier production function. J Prod Anal 5(2):171–180

Hunt-McCool JC, Warren RS Jr (1993) Earnings frontiers and labor market efficiency. In: Fried HO, Lovell CAK, Schmidt SS (eds) The measurement of productive efficiency. Oxford University Press, New York, pp 197–209

Jondrow J, Lovell CAK, Materov IS, Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. J Econom 19(2/3):233–238

Kalirajan KP (1981) An econometric analysis of yield variability in paddy production. Can J Agr Econ 29(3):283–294

Kneip A, Park BU, Simar L (1998) A note on the convergence of nonparametric DEAestimators for production efficiency scores. Econom Theory 14(6):783–793

Koopmans TC (1957) Three essays on the state of economic science. McGraw-Hill Book Company, New York

Korostelev AP, Simar L, Tsybakov AB (1995) Efficient estimation of monotoneboundaries. Anal Stat 23(2):476–489

Kumbhakar SC, Ghosh S, McGuckin JT (1991) A generalized production frontier approach for estimating determinants of inefficiency in U. S. dairy farms. J Bus Econ Stat 9(3):279–286

Kumbhakar SC, Lovell CAK (2000) Stochastic frontier analysis. Cambridge University Press, Cambridge

Kuosmanen T (2006) Stochastic nonparametric envelopment of data: combining virtues of SFA and DEA in a unified framework. MTT Discussion Paper No. 3/2006
Kuosmanen T, Kortelainen M (2007) Stochastic DEA myths and misconceptions, PPT url http://www.docstoc.com/docs/50906992/Stochastic-DEA-Myths-and-misconceptions-TimoKuosmanen-HSE-MTT-Andrew-Johnson-Texas-A-M-University-Mika-Kortelainen-University-of-Manchester-XI-EWEPA-2009-Pisa
Kuosmanen T, Post GT (1999) Robust efficiency measurement. Rotterdam Institute for Business Economic Studies (RIBES), 1999/3/25, Rotterdam. Report, vol 9911
Lall SV, Rodrigo GC (2001) Perspectives on the sources of heterogeneity in Indian industry. World Dev 29(12):2127–2143
Land KC, Lovell AK, Thore S (1993) Chance constrained data envelopment analysis. Manag Decis Econ 14(6):541–554
Levin H (1974) Measuring efficiency in educational production. Public Financ Quart 2(1):3–24
Little IMD, Mazumdar D, Page JM Jr (1987) Small manufacturing enterprises: a comparative study of India and other economies, Oxford University Press, Washington
Lundvall K, Battese GE (2000) Firm size, age and efficiency: evidence from Kenyan manufacturing firms. J Dev Stud 36(3):146–163
Maindiratta A (1990) Largest size efficient scale and size efficiencies of decision- making units in data envelopment analysis. J Econom 46(1/2):57–72
Meeusen W, Van den Broeck J (1977) Efficiency estimation from Cobb-Douglas production function with composed error. Int Econ Rev 18(2):435–444
Neogi C, Ghosh B (1994) Intertemporal efficiency variations in Indian manufacturing industries. J Prod Anal 5(3):301–324
Nikaido Y (2004) Technical efficiency of small-scale industry: application of stochastic production frontier model. Econ Polit Weekly 39(6):592–597
Nourzad F (2002) Real money balances and production efficiency: a panel data stochastic production frontier study. J Macroecon 24(1):125–134
Olesen OB, Petersen NC (1995) Incorporating quality into data envelopment analysis: a stochastic dominance approach. Int J Prod Econ 39:117–135
Olson J, Schmidt P, Waldman D (1980) A Monte Carlo study of estimators of the z stochastic frontier production functions. J Econom 13(1):67–82
Ondrich J, Ruggiero J (2001) Efficiency measurement in the stochastic frontier model. Eur J Oper Res 129:434–442
Page JM (1984) Firm size and technical efficiency: application of production frontiers of Indian survey data. J Dev Econ
Pitt MM, Lee L-F (1981) Measurement and sources of technical inefficiency in the Indonesian weaving industry. J Dev Econ 9(1):43–64
Ramaswamy VK (1994) Technical efficiency in modern small-scale firms in Indian industry: applications of stochastic production, frontier. J Quant Econ 10(2):309–324
Ray SC (2004) Data envelopment analysis: theory and technique for economics and operation research, Cambridge University Press, Cambridge
Reifschneider D, Stevenson R (1991) Systematic departures from the frontier: a framework for the analysis of firm inefficiency. Int Econ Rev 32(3):715–723
Richmond J (1974) Estimating the efficiency of production. Int Econ Rev 15(2):515–521
Robinson MD, Wunnava PV (1989) Measuring direct discrimination in labor markets using a frontier approach: evidence from CPS female earnings data. South Econ J 56(1):212–218
Rossi MA, Canay IA (2001) Measuring inefficiency in public utilities: does the distribution matter? http://www.aaep.org.ar/espa/anales/pdf_01/rossi_canay.pdf
Roy Biswas P, Ghose A (2012a) Growth efficiency and productivity of Indian manufacturing industry: an econometric analysis, LAP Lambert Academic Publishing, Saarbrucken ISBN 978-3-8484-8354-9
Roy Biswas P, Ghose A (2012b) Role of trade related factors in determining technical efficiency of West Bengal's manufacturing industries : evidence from stochastic frontier approach. Trade Dev Rev 5(1):1–36

Schmidt P (1976) On the statistical estimation of parametric frontier production functions. Rev Econ Stat 58(2):238–239
Seaver BL, Triantis KP (1989) The implications of using messy data to estimate production-frontier-based technical efficiency measures. J Bus Econ Stat 7(1):49–59
Seiford LM, Zhu J (1999) An investigation of returns to scale in data envelopment analysis, Omega, Elsevier, vol 27(1), pp 1–11
Shapiro KH, Müller J (1977) Sources of technical efficiency: the roles of modernization and information. Econ Dev Cult Change 25(2):293–310
Sharma KR, Leung PS (1998) Technical efficiency of carp production in Nepal: an application of the stochastic frontier production function approach. Aquacult Econ Manage 2:129–140
Simar L, Wilson PW (1998) Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. Manage Sci 44:49–61
Simar L, Wilson PW (2000a) Statistical inference in nonparametric frontier models: the state of the art. J Prod Anal 13:49–78
Simar L, Wilson PW (2000b) A general methodology for bootstrapping in nonparametric frontier models. J Appl Stat 27:779–802
Simar L, Zelenyuk V (2006) On testing equality of two distribution functions of efficiency score estimated via DEA. Econom Rev 25(4):497–522
Stevenson RE (1980) Likelihood functions for generalized stochastic frontier estimation. J Econom 13(1):57–66
Timmer CP (1971) Using a probabilistic frontier production function to measure technical efficiency. J Polit Econ 79(4):776–794
Van den Broeck J, Førsund FR, Hjalmarsson L, Meeusen W (1980) On the estimation of deterministic and stochastic frontier production functions. J Econom 13(1):117–138
Waldman D (1982) A stationary point for the stochastic frontier likelihood. J Econom 18(2):275–279
Wang H-J, Schmidt P (2002) One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. J Prod Anal 18(2):129–144
Winsten CB (1957) Discussion on Mr Farrell's paper. J R Stat Soc Series A 120(3):282–284
Wilson SW (1995) Classifier fitness based on accuracy. Evol Comput 3:149–175
Wilson P, Hadley D, Asby C (2001) The influence of management characteristics on the technical efficiency of wheat farmers in Eastern England. Agr Econ 24(3):329–338

Chapter 11
Telecommunications Industry in the Era of Globalization with Special Reference to India

Debabrata Datta, Soumyen Sikdar and Susmita Chatterjee

Introduction

Globalization opens up possibilities for gains in efficiency through international exchange based on the principle of comparative advantage. These gains are very significantly augmented with the development of communications system that reduces cost of negotiations, monitoring, and coordination. The advent of telegraph as a communication device in 1839 in Britain marked a signal change in this scenario of cost of communication. This facility for telecommunication, subsequently aided by information and communication technology (ICT) has facilitated resource reallocation process for specialization and business diversification at the global level and a symbiotic relation between globalization and communication has been established. ICT improves connectivity between countries and promotes globalization. Given this role of telecommunication industry in promoting economic efficiency of resource mobilization at the global level, the efficiency of telecom sector itself emerges as a matter of prime concern. For several reasons that will be discussed later on, the role of the nation state and the optimum regulation of telecom industry remain decisive factors in achieving the efficiency.

In the initial stage of telecommunications it was a wired-line service. The wired-line service is subject to increasing returns to scale (IRS) on account of huge cost involved in roll out of wire network and its maintenance. So it was economically wasteful to set up network for each company, and hence in almost all countries the early market structure of telephone industry was framed on the

D. Datta
Institute of Management and Technology, Gaziabad, India

S. Sikdar
Indian Institute of Management Calcutta, Kolkata, India
e-mail: soumyen1956@gmail.com

S. Chatterjee (✉)
Globsyn Business School, Kolkata, India
e-mail: susmita.cb@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_11, © Springer India 2013

principle of natural monopoly to be operated by either government owned or regulated firm with a responsibility of maximizing social welfare by suitable pricing.[1] Although government owned firm is a usual choice under natural monopoly, Demsetz (1968) shows that with suitable bidding mechanism at the entry stage, government can serve the purpose of social welfare maximization by choosing a private firm as well. This natural monopoly characteristic of telecommunication industry ceased to exist with the advent of wireless technology in the 1980s. Once the shackle of wired network is removed, IRS barrier to entry is over. The wired-line telephony has faced the twin disadvantages of declining demand for fixed point service and high cost of maintenance of wired line.[2] Thus, the incumbent firm has been led per force to compete on the wireless telephony only. As a result, private players become a part of telecommunication industry and the market becomes oligopolistic in nature, subject to government regulation in most of the countries. Execution of this policy, however, needed thorough reforms of the government laws and regulations. Indian economy also has been witnessing similar process of reforms since 1994.

In most of the countries in the world, the erstwhile state-owned telephone system has been allowed to continue. In most of the cases, however, the departmental undertaking is corporatized and this new public-owned corporate body is to compete with the entrant private firms. In many developing economies, the private firms are fully or partly under foreign ownership. The evolving market structure thus becomes a mixed oligopoly with coexistence of private and public firms.

The Plan of the Paper

In what follows, we address in terms of manageable theoretical models several issues that affect the telecommunication industry. We start with an overview of technology that enables the general readers to understand the complicated nature of this industry and the role of spectrum in it. Thereafter, we explain the concept of mixed oligopoly and show some results under Cournot competition in this setup. (The results under Bertrand competition are discussed in the appendix). In the following part, we consider the issue of spectrum allocation and relative merits and demerits of different mechanism by which spectrum is allocated among different

[1] In the USA, a private company American Telephone and telegraph Co (AT&T) worked as a national monopoly under a deal with government. In 1982, AT&T was broken up into regional companies with the hope that local monopoly would pave the way for competition in the market with the development of satellite technology. In Israel, incumbent monopoly firm was instructed to not to reduce the connection price to ensure the entry of new firms. The entire scenario, however, changed with advent of wireless technology (Shy 2001).

[2] In an ITU estimate, the cost of wired-line service is 80 % higher than that of wireless telephony. In effect, in many countries, there has been steady decline in fixed line—mobile connection ratio since the introduction of mobile telephony. See Oestmann (2003).

firms. Given that volatility of technology is a major characteristic of telecom industry, the next part of the paper considers the issue of introduction of 3G technology and competition between first and second mover firm in the context of Stackelberg model. An open area in this issue but not addressed in this paper is a possible entry prevention strategy to be adopted by the incumbent firm. The last theoretical part considers switching cost and presents a general discussion on pricing. What this paper again has not touched but can be an area of enquiry is the possibility of tacit collusion among firms. In the final part of the paper, we discuss the Indian telecommunication industry in the backdrop of the foregoing theoretical analysis and try to get an idea about the relevance of the theoretical results of the models in a real economic scenario.

This of course still leaves several areas of research unaddressed. The optimum auction design for spectrum allocation is still an open area. There is an important area of tacit collusion in a situation of infrastructure sharing. Strategic entry prevention by litigation is another possible area of research. These areas may be future research agenda.

Let us first focus on the market structure of the telecommunications industry. We begin by giving a nontechnical overview of the telecom technology and an idea about the possible cost structure. Thereafter, we present different variants of a simple theoretical model in order to explain this mixed oligopoly characteristics of the telecom industry. As argued earlier, the incumbent public monopoly does not have the option of entry preventing limit price using the old technology. Instead, our focus is on mixed oligopoly where public and private firms coexist with the public firm continuing to bear some burden of inefficiency, a legacy of lack of competition in the past.

An Overview of Telecom Technology

Telecommunication is the transmission of information over significant distances to communicate. The fixed-line telecommunication was invented in 1876 and, up until the recent past, was defined as a telephone line that traveled through metal wire or optical fiber as part of a nation-wide telephone network. Obviously, this technology involves cost in terms of laying of telephone lines and erection of towers. The cost per user is particularly high in difficult geographic terrain and in sparsely populated areas.

Since the late 1980s wireless telephony has come into the market. Although initial entry cost in this business may be high due to high cost of spectrum acquisition, the operating cost is substantially low for the following reasons.

1. Mobile investment per user is a function of economies of scale. The base station is a single bulk investment and costs per user decreases continually as more mobile users share the service. In contrast, fixed-line investment can only capture the revenue of a single user unless it is a public payphone.

2. Economies of scale have reduced the cost of base stations and other cellular infrastructure considerably. In some cases cellular infrastructure is re-used, lowering costs even further.
3. A major enhancement of GSM technology for rural areas is the option to implement "range extension". The reach of a GSM-900 base station can be stretched from its theoretical limit of 35–70 km, or even as far as 120 km. This is achieved at the expense of cell sector coverage by focusing the geographical spread of the signal into one direction.

The major issue of the mobile telephony is spectrum availability. The word 'Spectrum' basically refers to a collection of various types of electromagnetic radiations of different wavelengths. In India, the radio frequencies are arbitrarily confined between 9 kHz and 3000 GHz and are being used for 40 different types of services like fixed communication, mobile communication, broadcasting, radio navigation, radiolocation, fixed and mobile satellite service, aeronautical satellite service, radio navigational satellite service, etc., in conformity with the provisions of national and international laws.

Use of radio frequency spectrum is susceptible to overlapping interference and requires the application of complex engineering tools to ensure interference free operation of various wireless networks. Unlike other natural resources, radio frequency spectrum is not consumed upon its usage. Here wastage means suboptimal and inefficient use of spectrum.

The International Telecommunication Union (ITU) allocates spectrum frequencies for the use of various countries. Allocations are made on a regional basis and for different types of services. It is mandatory for all administrations to adhere to these allocations. For the purpose of spectrum allocation, each member country submits its proposals to ITU, based on their requirements and priorities for opening of the bands.

Low-frequency radio wave at kilohertz level is earmarked for radio transmission. Lower the frequency wave less is the loss of data over distance. Hence, radio set can catch signal from distant radio stations like BBC or VOA. But low-frequency radiowave is inefficient in data transmission on account of high interference among different transmission under low bandwidth.[3] (Bandwidth is the difference between two band waves). Thus, mobile needs high-frequency wave and in India it is in the 890–960 MHz and 1710–1930 MHz range. Major share of this bandwidth is in the possession of defence (65 MHz in India out of 100 MHz). The rest is distributed among the telecom companies.

Spectrum thus appears to be a scarce natural resource and government takes up the responsibility of its distribution. The economic issues involved here are discussed subsequently but here we address one issue linked to market competition. If government auctions spectrum, the entry cost of the new firms is very high and only companies with deep pocket will be able to enter the market. One argument

[3] FM transmission is clearer than AM transmission because of high data wave frequency but its range is limited.

(Maskin 2000) suggests that the ability to pay for the asset out of future earnings might get around the inefficiency associated with capital constraints. Another view is that financing of auction is not a problem in an economy, where there is a functioning capital market along with a limited liability clause for corporate borrowers. In such market aggressive bidding for acquisition of spectrum is common. If the bidder fails to earn sufficient revenue, burden falls on the lending bank or the government In fact early path of entry of private telecom firms is littered with bankrupt and sold out firms in many countries including India and the USA. See Zheng (2001), Board (2007) in this regard.

Mixed Oligopoly

Since the 1970s, the world has witnessed full or partial privatization of many erstwhile public sector enterprises. China introduced its privatization policy in 1978, the Thatcher government in the UK started a privatization drive from 1979, and similar policies were subsequently followed in East Europe, South America, and other Asian countries, including India. In many industries, these measures led to the coexistence of public and private firms creating a market structure, termed as mixed oligopoly in the literature. In the literature, the discussion on mixed oligopoly dates back to Merrill and Schneider (1966). Thereafter, a huge literature has come up with contributions by De Fraja and Delbono (1989, 1990). We present below a simple model to derive some results of mixed oligopoly.

There are two firms in this model—private (X) and public (Y) with production levels x and y. The basic behavioral assumption is that the private firm maximizes profit but the public firm maximizes social welfare which is the sum of consumers' surplus and the total profit of the firms. Let us introduce two more distinctions between the public and private firms, both of which are empirically validated by the survey among customers of BSNL, the state-owned firm in India (Datta and Chatterjee 2012). First, the public firm is perceived to be less efficient and less consumer friendly by the consumers and therefore has a lower demand. This is captured by assuming that maximum reservation price for the private firm is 1 and for the public firm $\lambda < 1$

Thus, the demand functions for the private firm and the public firm are $p = 1 - x - y$ *and* $p = \lambda - x - y$ respectively.

We have the restriction that $\lambda < x + y$

Second, the public firm cannot continue with solely the old technology and it has to switch over to the new technology (mobile telephony). But it has a higher cost on account of inefficiency.[4] For simplicity, we assume that the private firm has

[4] Inefficiency should not always be taken in a pejorative sense. A part of the inefficiency may be a legacy of the inefficiencies of its past monopoly. But higher cost may be due to better adherence to regulatory norms compared to private firms. BSNL scores higher than the private firms in terms of transparency but that imposes a burden on BSNL. See Datta and Chatterjee (2012).

zero average and marginal cost and for the public firm this average and marginal cost is $c > 0$. We assume further that the firms are engaged in Cournot competition with their respective objectives, demand and cost structures, and with no capacity limit.

Whether firms play quantity competition or price competition game is not unequivocally settled in the literature. A general consensus has emerged that the Cournot and Bertrand models are best viewed as reduced forms of some dynamic process where decisions are taken about both types of variables, price and quantities. The variable that is more difficult to adjust would be the dominant strategic variable. There are results suggesting that even when firms choose prices if they make irreversible production decision first, a quantity setting reduced form may be appropriate. Payoffs would depend on decided quantities, already incorporating the effect of subsequent price competition.

Only a few very recent papers look at the issue of Cournot and Bertrand in the context of mixed oligopoly. Ghosh and Mitra (2010) compare Cournot with Bertrand completion in a mixed oligopoly with differentiated good and finds reversal of the results that Bertrand competition yields lower prices as well as profit and higher consumer surplus (Singh and Vives 1984; Vives 1999). According to this paper, for substitute goods, the public firm's output is higher and the private firm's output lower under Cournot competition, compared to the Bertrand competition. In addition, profits of both firms are lower under Cournot competition than they are under Bertrand competition. Matsumura and Ogawa (2012), however, suggest that choosing Bertrand competition is a dominant strategy for both public and private firms when there is no efficiency gap between the public and private firms. Choi (2012) considers the case where public firm is less efficient than the private firm, an assumption we hold in our model. Choi's paper develops a two stage game model and shows that for substitute goods adoption of price competition or quantity price competition depends on degree of efficiency gap. All these models consider differentiated goods market. Given this possibility of both quantity and price competition even in the case of differentiated goods, we shall mainly consider Cournot competition here for its analytical simplicity and tractability in a market of homogenous good. But for the interested reader we shall present a simple Bertrand model in a differentiated good framework in the appendix.

The profit function of the private firm is

$$\pi_X = x - x^2 - xy \tag{11.1}$$

The private firm maximizes this profit taking y as given.

The public firm wants to maximize social welfare taking x as given

$$W_Y = \lambda y - xy - y^2 - cy + x - x^2 - xy + \frac{(x+y)^2}{2} \tag{11.2}$$

From the necessary maximizing conditions we get the following equilibrium values.

$$\left.\begin{aligned} x^* &= 1 - \lambda + c \\ y^* &= 2\lambda - 2c - 1 \\ p^* &= 1 - \lambda + c \end{aligned}\right\} \tag{11.3}$$

The above result shows that if the public firm has no disadvantage with regard to demand and cost, i.e., if $c = 0$ and $\lambda = 1$ there is no need for private firm from the point of view of economic efficiency. The public firm can ensure Pareto efficiency by following social welfare maximization principle (Lange 1936). But as the post World War II experiences have shown in many countries, higher cost resulting from legacy of the past is a reality and taking this into account the governments of various countries have taken recourse to privatization.

Let us now consider n number of firms out of which $n - 1$ firms are private and the nth firm is the public firm. If the public firm is also profit maximizing like the private firms but face higher cost and lower demand then the solutions are

$$x_i^* = \frac{2 - \lambda}{n + 1}, \quad y_n^* = \frac{1 - n(1 - \lambda)}{n + 1} \tag{11.4}$$

If, on the other hand, the public firm maximizes social welfare, the solutions are

$$x_i^* = 1 + c - \lambda, \quad y_n^* = 1 - n(1 - \lambda) - nc \tag{11.5}$$

It is observed that as n rises, space for public firm goes down.

The above result shows that with privatization and entry of new private firms in the fray, the public sector firm may find it more and more difficult to survive in the market, unless adequate cost management is exercised and adequate support is provided for its social obligation.

We can consider a different scenario where an industry x is allowed with public private partnership and run on pure profit motive. We thus introduce the case of partial privatization by assuming that α is the extent of public holding in x. It is also assumed that the public firm puts weight α on social welfare in its objective function.

$$\prod_y = (1 - \alpha)\pi_y + \alpha W \tag{11.6}$$

where W is the social welfare (sum of profit in x and y industry and consumers' surplus)

$$\prod_y = \pi_y + \alpha\pi_x + \alpha(CS)$$

where CS denotes consumer surplus.

Then under the assumption $\lambda = 1$, we get

$$x^* = \frac{1 - \alpha + c}{3 - 2\alpha}, \quad y^* = \frac{1 - 2C}{3 - 2\alpha} \tag{11.7}$$

The interesting result in this case is that the social welfare W curve will be 'U' shaped with regard to public firm's cost c (Fig. 11.1). If $c = 0$, then as we have shown earlier the public monopoly with objective of social welfare maximization is best with welfare level of W_{PC}. If now cost is higher for public firm welfare initially decreases but after a certain level as less costly private output replaces inefficient public production gradually welfare starts rising and ultimately reaches W_{Mon}, the welfare corresponding to private monopoly. So if inefficiency of the public firm crosses a certain level, it is better to rely on private firms only. In the following figure, it is thus shown that if the public firm is fully efficient, it can maximize welfare on its own (Lange 1936) i.e. at $c = 0$, maximum welfare is attainable with monopoly public firm producing the competitive output. But if it is marginally inefficient, there should be entry of private firms and reforms to improve the efficiency of the public firm. But if it is grossly inefficient, full scale privatization is better.

In this setup, let us now introduce foreign investment and ownership of domestic firm. Let there be three firms. X is private firm owned domestically, Y is public firm and Z is private firm partly owned by foreigners. Then

$$p = 1 - x - y - z \tag{11.8}$$

$$\left.\begin{aligned} &\text{x maximizes } \pi_x = x - x^2 - xy - xz \\ &\text{z maximizes } \pi_z = z - z^2 - xz - yz \\ &Y \text{ maximizes } \Pi_y = R_y - cy + \alpha\pi_x + \alpha\theta\pi_z + \alpha(x + y + z)^2/2 \end{aligned}\right\} \tag{11.9}$$

where $\alpha =$ extent of public holding and $\theta =$ domestic holding in z, $\prod_y, R_y =$ social welfare pay-off and revenue of public firm respectively.

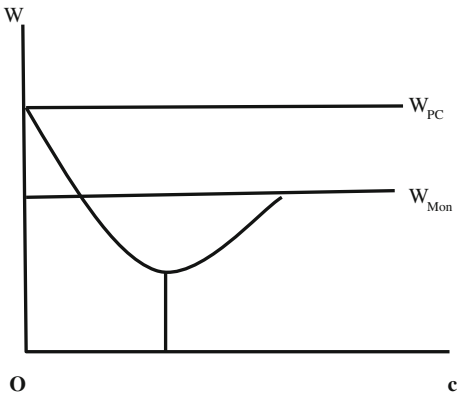


Fig. 11.1 Social welfare and public sector inefficiency

We get the following equilibrium

$$\left.\begin{aligned} x^* &= z^* = \frac{1-\alpha+c}{4-2\alpha-\theta\alpha} \\ y^* &= \frac{1+\alpha(1-\theta)-3c}{4-2\alpha-\theta\alpha} \end{aligned}\right\} \tag{11.10}$$

If $c = 0$, social welfare is maximized with $\alpha = \theta = 1$

i.e., if the public firm is fully efficient no privatization or foreign investment is necessary.

But suppose $c = 1/3$. *So* $y = 0$ *and* $x = z = 1/3$. In this case it is better to privatize or invite foreign ownership. For example, if $\alpha = 1$it is best to go for full foreign ownership i.e. $\theta = 0$. But suppose for some reason, policy-makers cannot introduce full foreign ownership. Then, if only 50 % foreign investment is allowed i.e. $\theta = 1/2$ the best degree of privatization is 11/15. If, on the other hand, foreign ownership is more at $\theta = 1/3$, the privatization is less at 13/21. So there is a tradeoff between privatization and foreign ownership in this model.

Mixed Oligopoly and Regulation

Let us now consider the case of regulation in mixed oligopoly. As it has been shown earlier that if the public firm has no efficiency problem, public production with the objective of social welfare maximization can ensure Pareto optimal outcome but given the cost inefficiency of the public firm there is scope for regulation. We present below this aspect of regulation in a situation where quality of production is also an issue. In our two-firm scenario, one public (Y) and another private (X), we now assume that both firms have a constant unit cost of production, which can however be reduced with deficiency in quality (b). Increase in the value of b implies fall in quality of services. We further assume that given the quality level of, say X's service, reduction in the quality of Y's service reduces demand for y but increase the demand for x. The firms play the Cournot game i.e., they optimize having assumed that the quality and quantity of the other firm are unchanged.

A relevant question here is whether the firms face capacity constraint on account of limited availability of spectrum. This is a valid question but as Oestmann (2003) shows and our model assumes that firms compromise with quality to increase the supply of quantity of services.[5] There is, however, discussion of mixed oligopoly under capacity constraint in the literature. Main result obtained in Papers of Nishimori and Ogawa (2004) and Lu and Poddar (2005) is that, in the

[5] In fact, those days are gone even in India when one had to wait a long period to get a telephone connection but as TRAI has documented quality of services in Indian telephony is pathetic.

mixed duopoly market, the public firm strategically chooses under capacity while the private firm chooses excess capacity.

Let us first consider ordinary Cournot game. The firms have the following demand and profit functions.

$$\left.\begin{aligned} p_x &= (1 - ab_x^2 + ab_y^2) - x - y \\ p_y &= (1 + ab_x^2 - ab_y^2) - x - y \end{aligned}\right\} \tag{11.11}$$

$$\left.\begin{aligned} \pi_x &= (1 - ab_x^2 + ab_y^2)x - x^2 - xy - (c_x - b_x)x \\ \pi_y &= (1 + ab_x^2 - ab_y^2)y - y^2 - xy - (c_y - b_y)y \end{aligned}\right\} \tag{11.12}$$

where represents where b_i represents quality reduction of ith firm $i = x, y$, a is a coefficient to capture the effect of quality reduction. Here we assume without any loss of generality $a = 1$.

If we assume both firms have symmetric cost, the model will imply a symmetric quality and quantity as follows.

So,

$$b_x^* = b_y^* = 1/2a, \;\; x^* = y^* = 1 - c + b^*/3 \tag{11.13}$$

$$p^* = 1 + 2c - 2b^*/3, \;\; \pi_x^* = 1 + c - 2c^2 + 4b^*c - 2b^{2*} - b^*/9 \tag{11.14}$$

where $b^* = \frac{1}{2a}$

Whether the firms will maintain a positive quality deficiency depends on the profit configuration. In the following Fig. 11.2 we draw a 45 ° line OK and another line IJ where $\pi = \pi_{b=0}$. To the right of IJ profit is more with quality deficiency. Given $c > b$, thus the shaded area IJKL gives the configurations of b and c for which profit is increased by quality cutting. So quality deficiency is profitable only at a cost not very close to zero. This is understandable because if cost of production is very low, there is no point in cutting quality.

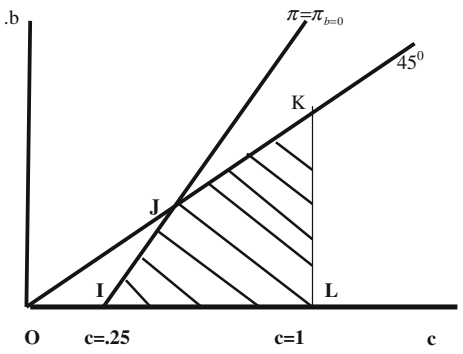


Fig. 11.2 Quality deficiency and profitability

Let us now introduce the public firm in this scenario with the objective of social welfare maximization. For the sake of simplicity, we assume $a = 1$ without any loss of generality. We further assume that consumer welfare is reduced by quality cutting at the rate m. We get the following profit functions now.

$$\left.\begin{aligned} \pi_x &= (1 - b_x^2 + b_y^2)x - x^2 - xy - (c_x - b_x)x \\ \pi_y &= (1 - b_x^2 + b_y^2)y - y^2 - 2xy - (c_y - b_y)y \\ &+ (1 - b_x^2 + b_y^2)x - x^2 - (c_x - b_x)x + \frac{(x+y)^2}{2} - m(b_x + b_y) \end{aligned}\right\} \tag{11.15}$$

$$\left.\begin{aligned} \frac{\partial \pi_x}{\partial x} &= (1 - b_x^2 + b_y^2) - 2x - y - (c_x - b_x) = 0 \\ \frac{\partial \pi_x}{\partial b_x} &= -2b_x x + x = 0 \rightarrow b_x* = 1/2 \\ \frac{\partial \pi_y}{\partial y} &= (1 + b_y^2 - b_x^2) - y - x - (c_y - b_y) = 0 \\ \frac{\partial \pi_y}{\partial b_y} &= -2b_y y + y + 2b_y x - m = 0 \end{aligned}\right\} \tag{11.16}$$

From the above equations given $b_x^* = 1/2$,

We can solve for. x^* y^* and b_y^*. We get three roots of b_y^*, which represents the possibility of multiple equilibria.

Assuming $c_x = 0.75$, $c_y = 0.9$, $m = 0.1$, we get either $b_y^* = 0.53$, or $b_y^* = 0.26$ compared to $b_x^* = 0.5$, For $b_y^* = 0.53, y^* = 0.47$ *and* $x^* = 0.13$, For $b_y^* = 0.26$, $y^* = 0.27$ *and* $x^* = 0.39$.

So the quality level of the services of the public firm may be better or worse than that of the private firm. The quality value, however, will depend on m that shows the quality consciousness of the public firm. If m is higher, quality of service of the public firm is likely to improve. If initial b_y is greater than one-half, quality of service of the public firm definitely rises in our model as m rises (For example, with initial $b_y^* = 0.53$ *for* $m = 0.1$, when $m = 0.2$ we get $b_y^{**} = 0.46$, $y^{**} = 0.33$ and $x^{**} = 0.19$ and this implies an improvement of quality). If public product is of better quality the output by the public firm will be less and private output will be more. It will be the reverse for low-quality public good.

Before concluding this section, we should mention that the issue of increasing returns to scale often arises in the context of telecommunications industry. Instead of pursuing this area of research here we mention two interesting recent papers in this context. Sengupta and Tauman (2011) arrive at incentive mechanisms by which a benevolent regulator is able to implement, without running a deficit, a welfare improving outcome by making every Cournot oligopolistic firm but one exits the market. Greve and Keiding (2010) consider the mechanism of truthful revelation of costs by firms in a market with regulated competition under increasing returns to scale. In both these cases of IRS subsidization becomes necessary.

Spectrum Allocation, Multiple Technologies, and Switching Cost

The above model shows that once the market is privatized, need for regulation comes only from the oligopolistic nature of the market. For the telecom industry; however, this need definitely remains as entry cost to this industry is very high; the need for regulation also arises on account of the requirement of a scarce resource namely spectrum. If the distribution and regulation of this natural resource is not regulated, the firms can get into economic warfare to corner these resources and this will kill the industry. So let us now focus on the issue of spectrum allocation.

Once cellular telephony came into being, spectrum allocation among users appeared as a formidable task for the government in all countries of the world. Distribution of scarce resources like land has been historically settled more or less (although issues of land disputes erupt often) but how to decide in the modern day allocation of a scarce resource like spectrum? Given that society can never allow any scarce resource to be usurped by power, four methods can be contemplated: (a) First-come first served, (b) lottery, (c) discretionary allocation, known as beauty contest in the literature, and (d) auction. First-come first-served principle has an incumbency bias, and thus can foster monopoly by acting as an entry barrier. So this principle is not favored in spectrum allocation. In the old days of radio service, lottery system was followed in the USA but it was found that this system was grossly inefficient as it brought into the bidding a large number of casual and insincere parties seeking to try luck to get some scarce resources. This leaves the last two as viable modes of spectrum allocation and the merits and demerits of the two have become a major debatable issue in the literature.

There is a huge literature on spectrum allocation and auction in the telecom industry. Wilson (1992) is a discussion on the theory of auctions. Milgrom and Weber (1982) have elaborated on many ideas of spectrum auctions. McMillan (1994) gives an overview of the issues involving sale of spectrum rights. Economides (2004) gives an introduction to the telecommunication regulation. Malik (2004) discusses the Indian telecommunication policy and regulation while Virmani (2000, 2004) argues for complete de-licensing of entry into (and exit from) telecom services coupled with competitive auction of *tradable* spectrum permits.

The issue of pros and cons of spectrum auction have received wide attention of the literature. Zheng (2001) shows how high bid for spectrum auction can lead to bankrupt winners. This may be a problem of auction designs and several papers (Bulow et al. 2009; Board 2007) have tried to address this issue. Cave and Valetti (2000) have shown that the spectrum auction through ascending auction need not necessarily be bad as the license fee is only a sunk cost. Haan and Toolsema (2003) show that consumers' price may be lower under spectrum auction than under beauty contests. On the other hand, Burguet (2005) shows how beauty contest method is a better approach to spectrum distribution than the auction of licensing. It is thus observed that the relative merits and demerits of spectrum

auction and beauty contests have remained unresolved. When possibility of allegation of corruption is high and when credibility of political leadership is low, beauty contest may better be avoided. This has the risk of derailing the progress of telecom industry under series of litigation. But if the corruption is not an issue and government wants to ensure the spread of telecom services to the low income people, beauty contest may be adopted. Otherwise auction has no alternative. But since the history also shows that auctioning was not smooth in several countries, designing of auction should be carefully contemplated.

Choice of Technology

The problem of initial allocation of spectrum was no doubt difficult but this allocation is perhaps more difficult in an existing industry characterized by more than one technology. Presently, the telecom industry is characterized by the presence of GSM and CDMA technology and it is stated that their costs are not identical. This scenario makes the application of beauty contest principle, i.e., discretionary allocation, difficult as the regulatory body does not know the exact cost nor other parameters. There is no empirical formula which can establish the spectrum requirement. Additional spectrum requirement depends on various parameters like the number of subscribers, the density of subscribers morphology of the service area—urban, semi-urban or rural terrain, type of applications (voice, data, multimedia application such as video/audio streaming, video on demand, mobile TV, etc.) and pattern of traffic, (voice, data, and multimedia services), and the technology itself, etc. [2G (CDMA IS95A, GSM), 3G (WCDMA/CDMA2000), WiFi, WiMax]. In a multi-operator scenario where market forces are in a position to decide the business strategies, it is difficult to assess the value of various parameters mentioned above. While a few parameters such as current subscriber base, current growth rate, etc., are known, there are several indeterminate factors, which can substantially alter the spectrum requirement estimates; the prominent among them is the nature of traffic, i.e., extent of voice and data traffic in the network.

When beauty contest faces difficulties, one has to rely on auction but as we shall show in terms of our model that auction may not necessarily be efficient here. Here we confront a market with asymmetric cost structure with consumers' switching cost. Costs are asymmetric because CDMA technology users are said to be less costly service providers and a better users of spectrum, compared to GSM. Had it been the case one could argue that all spectrums could be given to CDMA technology users for the sake of efficiency. But it is not that easy where a large number of subscribers have GSM connection. As it is well-known that in the absence of number portability, telecom consumers are subjected to switching cost. My phone number is used by others and therefore if I change this number, I lose contact with many who know my old number but not the new one. This is known as switching cost. We show in terms of a simple Cournot type model how in the presence of

switching cost higher spectrum allocation may reduce welfare, rather than increase it. The mobile number portability (MNP) assumes significance in this case.

Singapore was the world's first country to implement MNP in 1997. The United Kingdom and the Netherlands first implemented MNP in Europe in 1999. Countries such as Spain (in 2000), Sweden and Denmark (all in 2001), Belgium, Italy, Germany, and Portugal (all in 2002) followed suit (Buehler et al. 2005).

Consumers of mobile telecommunications services typically face switching costs which derive from the real and psychological costs that consumers confront when changing suppliers (see, e.g., Klemperer 1987a, b, 1988, 1995). These switching costs are endogenous if they emanate from customer loyalty programs (such as Deutsche Telekom's so-called *Happy Digits* programme) or contractual clauses that make the change of suppliers more costly (such as contract termination penalties). There are also *exogenous* switching costs resulting from the transaction costs associated with switching providers (e.g., for changing the network assignment of a given number). Introducing MNP eliminates at least part of these switching costs.

On an overall basis, the competitive effects of introducing MNP are fairly complex and ambiguous. MNP is likely to affect retail prices, termination charges, price elasticities, market shares, as well as entry and investment decisions. So far, it is fair to say that most analyses on MNP have supported the notion that, on the whole, MNP intensifies competition in mobile telecommunication. There may of course be tacit collusion but the antitrust commission should look at it.

Available empirical evidence on the portability of premium rate numbers appears to support this conclusion (Viard 2004). Yet, it is unlikely that introducing MNP reduces prices always. Whereas prices for mobile services are likely to decrease as competition intensifies (Buehler and Haucap 2004), the pro-competitive effects of MNP are likely to vary across countries, depending on the degree of competition prevailing before introducing MNP.

NERA/Smith study of the costs and benefits of introducing MNP in Hong Kong estimated the additional benefit from increased competition amounts to 1€ per customer over a period of 10 years. (See NERA/Smith 1998). In contrast, Oftel (1997) study estimated the additional benefits of increased competition resulting from MNP to lie around 69 € per customer over a 10-year period—quite a marked difference which is due to Oftel's assumption that competition would be significantly more intense with MNP than without MNP (see Oftel 1997).

Since license fees, network investments and customer acquisition costs are all specific investments, operators are vulnerable to expropriation through renegotiation of the regulatory contract (Goldberg 1976; Sidak and Spulber 1999). In dynamic and innovative industries such as mobile communications, dynamic efficiency aspects are highly relevant. Therefore, regulators should be cautious in introducing regulations, as it may adversely affect investment and innovation incentives.

Let us now present a simple model that shows additional spectrum may not lead to Pareto improvement in the absence of mobile number portability. There are two firms, selling homogeneous goods or services in the market and they are involved

in quantity competition. Since it is no longer the mixed oligopoly situation, we denote the firms by 1 and 2. For convenience, we assume that the firms face a linear demand function of the simplest form

$$\left.\begin{aligned} p &= 1 - q_1 - q_2 \\ Q &= q_1 + q_2 \end{aligned}\right\} \tag{11.17}$$

where p = price, q_i = output if ith firm, Q = industry output.

The cost structure is asymmetric and without loss of any generality we assume $c_1 < c_2$.

We further assume that although the product is homogeneous a consumer suffers welfare loss if she/he is to change the firm in the second period because of the switching cost. If, however, there is mobile number portability, this switching cost is absent.

We consider the following two-stage game. In the first stage, the firms play Cournot game and make profit. With that profit they buy spectrum through auctions. We assume that at the highest bid both firms get spectrum as per their revenue payments. The additional spectrum reduces their cost but the firm buying more spectrum enjoys greater cost reduction. In the second stage, the firms again play Cournot game with new spectrum support.

Proposition: *In the second stage output of the high cost firm may be less than its output at the first stage for which some consumers may be forced to change firm and suffer welfare loss.*

Proof The profit functions of the firms are

$$\pi_i = q_i - q_i^2 - q_i q_j - c_i q_i \tag{11.18}$$

Then the equilibrium quantity, price, and profit configurations in the first period are

$$q_i^* = \frac{1 - 2c_i + c_j}{3}, \quad p^* = \frac{1 + c_i + c_j}{3}, \quad \pi_i^* = \frac{1}{9}(1 - 2c_i + c_j)^2 \tag{11.19}$$

$$\pi_1^* > \pi_2^* \text{ given } c_1 < c_2$$

So low cost firm 1 will be able to buy more spectrum from the government in the next auction.

Then in the second stage their costs of production will be

$$c_1\prime = c_1 - A \text{ and } c_2\prime = c_2 - \beta A \text{ where } \beta < 1$$

The cost of firm 1 will fall by a greater amount.

Then the equilibrium quantity, price, and profit configurations in the second period are

$$\left.\begin{aligned} q_1\prime &= \frac{1-2(c_1-A)+(c_2-\beta A)}{3} > q_1* \\ q_2\prime &= \frac{1-2(c_2-\beta A)+(c_1-A)}{3} \geq \text{or} \leq q_2* \\ Q\prime &= \frac{2-(c_1-A)-(c_2-\beta A)}{3} > Q* \\ p\prime &= \frac{1+(c_1-A)+(c_2-\beta A)}{3} < p* \end{aligned}\right\} \tag{11.20}$$

Since some consumers of firm 2 may have to switch to firm 1 there may not be welfare improvement in the Pareto sense in the absence of mobile number portability.

The above result highlights the role of number portability in efficient spectrum auction. Apparently, these two issues of MNP and spectrum auction may not appear to be closely related but our model shows that this measure is important for an optimal solution.

Our analysis has assumed that the firms can buy any quantum of spectrum, provided they can pay for it. But spectrum availability itself is a problem in every country. The overall spectrum of frequencies range from 300 Hz (cycles per second) to 300 GHz. But with today's technology, much of this range is not practically available for deployment in the telecom sector. With phenomenal growth of subscribers' number base, there is steady increase in demand for more spectrums. In many countries certain band of spectrum is reserved for defense. Thus apart from arrangement of additional spectrum, the need remains for efficient use of spectrum. The upgradation from 1G to 2G and then to 3G and 4G is technological response to the problem of scarcity of spectrum. The latest technologies can transmit more data compared to the old one. However, with improvement of technology the demand for data transmission also rises. Initially, mobile is a voice transmission instrument but now mobile handset is an instrument that transmits voice, data, and image. Thus, as in many other areas, supply and demand increases together keeping the scarcity of resources undiminished.

The coexistence of several technologies side by side makes the issue of efficient use of spectrum quite critical. The GSM technology is more efficient in long distance telephony while the CDMA is less costly in the short distance telephony. The service quality of 3G will not be up to the mark if service is not transmitted at appropriate spectrum level. (For example, in India the service is provided at 5 MHz fo each firm, a third of global average). These factors are to be taken into consideration for efficient spectrum management.

Advent of New Technology

The advent of 3G technology has brought about a paradigm shift in the concept of mobile communications. NTT DoCoMo introduced the 3G services in October 2001 in Japan. Early wireless networks always operated at low bandwidths,

suitable only for voice and data communication. In late 1990s, the wireless industry developed three types of high-speed wireless systems: the third-generation (3G) digital networks, WiFi, and Bluetooth (Elliott 2002). Among these systems, 3G combines the high-speed mobile access with the internet and enables new ways to communicate, access information, and conduct business. M-commerce, which refers to conducting business transactions while on the move, requires the application infrastructure to maintain business communications and transactions via a mobile device (Kalakota and Robinson 2002), and 3G system is just compatible with this requirements.

Countries have to maintain the ITU standardization in case of 3G. 3G is a new type of technical enabler which not only increases the transmission capacity several fold, but also transforms the manner by which data are transmitted going from packet-based connection to the Internet related connection (Jonason and Eliasson 2001). The role of 3G product is to provide customers with high speed and high capacity data services (Harmer and Friel 2001). Under 3G, for the mobile telecom operator, the main upstream includes content, security, and infrastructure equipment providers.

In India, 3G services were first introduced in February 2009 with two state run firms BSNL and MTNL receiving the required spectrum (2.1 MHz) and beginning the service. For some unknown reason, however, these state—run firms were provided spectrum, much less than the global standard (15 MHz). Introduction of 3G in India has stimulated keen interest in the world as India has already emerged as the world's second largest mobile market after China with nearly 850 million subscribers in 2011. The auction of spectrum for 3G brought huge revenue for the government.

Following the analysis of Porter (1985) , competition forces in telecom industry are determined by bargaining power of suppliers like infrastructure equipment provider and application provider, Complementors like content providers, bargaining power of buyers, rivalry among the existing competitors, threats of new entrants, threats of substitutes, and opportunity of legacy operator. In the subsequent analysis we focus on the issue of legacy operator. In the context of telecom industry, the public firm is usually the legacy operator. When private firms are allowed entry in this industry, the question of level-playing field between the state-owned operators and private service providers comes to the fore. The legacy operator advantage is identical as the first mover advantage, which is enjoyed by the public firm. These advantages can be as following: (a) defining the standards for technology and for other activities, (b) early profits if the technology is valued highly by buyers. But there are some disadvantages as well like (a) pioneering costs, (b) cost of educating buyers, (c) developing infrastructure, and (d) demand uncertainty.

On this account, we can say the incumbent firm can only make profit if costs of disadvantages are overcome by benefits of advantages. First mover advantage will be limited if MNP is introduced. As we have shown with the help of an economic model in the next part of the paper that it is not unlikely that second mover makes higher profit than the first mover. This can happen on account of second mover

having better information about the market, arising out of the opportunity to observe the experience of the first mover incumbent firm.

Our next model takes into account the above scenario. For the emerging 3G market in India, we feel that a Stackelberg model with demand uncertainty and asymmetric cost is appropriate a la Liu (2005). We make the following assumptions in this model.

1. The leader, being the first mover faces demand uncertainty,
2. The follower does not face any uncertainty as it can observe the demand in the first period when only the leader is in the market.
3. Firms have asymmetric cost patterns.
4. There is switching cost for consumers for which the follower cannot start a price game.

General demand function is

$$\left.\begin{aligned} p &= x - Q \\ Q &= q_1 + q_2 \end{aligned}\right\} \tag{11.21}$$

The new parameter x is introduced to capture uncertainty. We assume that the leader firm only considers the expected value of x Let a be the maximum value of x. For simplicity, we assume that x follows a uniform distribution.

Consider h as a particular value of x such that if firm 2 observes that $x \leq h$, it does not enter. Then profit function of firm 1 has two components as given below.

$$\pi_1 = \int_{q_1}^{a} \left(x - q_1 - \frac{x - q_1}{2}\right) q_1 f(x)dx + \int_{h}^{q_1} (x - q_1) q_1 f(x)dx - c_1 q_1 \tag{11.22}$$

where $f(x) = \frac{1}{a-h}$ is the density function under uniform distribution.

The equilibrium values are

$$\left.\begin{aligned} q_1^* &= \frac{8ah \pm \sqrt{64a^2h^2 - 12(4a - a^2 + 2h^2 - c_1)}}{6} \\ q_1^* &= \max\left(\frac{a - c_2 - q_1^*}{2}, 0\right) \end{aligned}\right\} \tag{11.23}$$

Since q_1^* involves both a and h it is possible that $q_1^* < q_2^*$

For example if $a = 4.1$, $h = 0$, $c_1 = c_2 = 0.4$ then $q_1^* = 0.2$, $q_2^* = 1.75$

Similarly, profit of second mover can be greater than that of the first mover. Cost is another factor that can make the second firm's output and profit bigger than that of firm 1. The result can be shown diagrammatically as follows.

Firm 1 determines the leadership output assuming G_SH_S to be the expected reaction function of second firm (S stands for small). But firm 2 observes larger

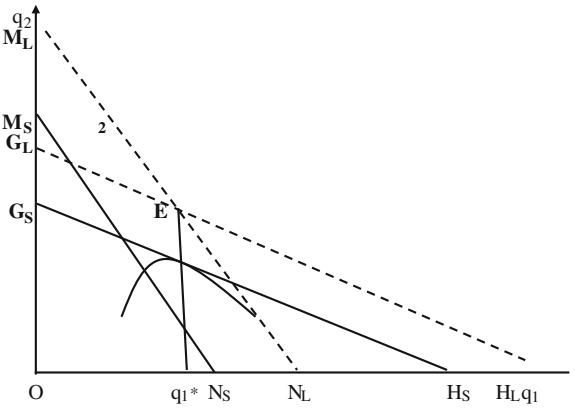


Fig. 11.3 Output determination under demand uncertainty

demand and produces $E\,q_1^*$, which is greater than output of firm 1. Given this miscalculation, Firm 1 finds that given $q_2 = E\,q_1^*$, it is best response to produce $O\,q_1^*$, corresponding to firm 1's reaction function as $M_L N_L$ (Fig. 11.3).

Simulation Exercise of the Model

We work out a genetic algorithm (GA) simulation for the above model. GAs are used as a model to represent learning. GA emulates the process of natural evolution and application of this algorithm is based on the assumption that firms in a situation of uncertainty will gradually optimize by following the natural selection path. The following steps are involved in optimizing the profit functions for Player1 and Player 2, leading to discovery of optimal values for Market Demand and Output.

Firms make a production quantity decision in a given period. Initially, their quantity production decisions are random, however; each period they learn a little more. Each firm is endowed with a single string which is used as its quantity production decision. It then compares this string against other firms' strings. After all agents have made a quantity production decision, the quantities are aggregated and plugged into a demand function to get a price. Each firm's profit is then calculated. Fitness values are then calculated as a function of profits. After the offspring pool is generated, hypothetical fitness values are calculated. These hypothetical values are based on some sort of estimation of the price level, often just by taking the previous price level.

In our case at the outset a random population of elements (called chromosomes in GA), encoding the probable values of Market Demand xi and Output qi, is generated, the range being 1.5–287 million for Player 1, while 0–287 million for Player 2. These figures are taken on the basis of present ownership level of 3G compatible handsets (1.5 million), which are assumed to be the lowest boundary of

the demand and the total number of subscribers at present (287 million), who constitute the upper boundary of the demand. It is assumed that second firm will not get any captive demand, which is already captured by the pioneering firm. So second firm has to start with a zero captive demand. The population has been subjected to the whole process of selection and crossover, until it converges to some optimum value.

After the first round of convergence, some *mutant* elements are inserted to capture the uncertainty of the atmosphere. The above two steps are repeated until all the elements converge to the same value and the ultimately converged value is recorded as optimum value.

It is found out that first firm gets an optimum demand of 20.4 million and makes a revenue of 382 million while second firm faces a demand of 23.5 million and makes a higher profit of 48.8 million. The interesting finding is that starting with a lower captive demand, second mover can end up with bigger revenue. The first mover firm makes a lower profit on account of its information disadvantage at the beginning and the consequent baggage of history.

Pricing in Telecom Industry: A General Discussion

Above model considers only quantity competition. But pricing competition will also play a major role in shaping up the future of the industry. Price level and structure is one of the most important factors in the adoption and diffusion of 3G services. Pricing models tested and used in different countries consists of combinations of two basic rates: a monthly fat rate based on connection speed as well as quality of service and an added component for premium services. Task of the 3G operators is to balance between the basic services and the value added services.

With regard to pricing of telecom services, regulator usually wants the free market operations. To let the market operate 'freely' some requirements are to be fulfilled. These are as follows:

1. Producers must be able to advertise and sell their product to any consumer.
2. Consumers should have no restriction in case of choice of the producer.
3. Products must be well described and information must be transparent.
4. Trust between consumer and producer should be ensured.
5. Given that the radio interface cannot be perfect, perfect product quality cannot be absolutely guaranteed. There must be a fair method of specifying how much confidence a consumer can place on the product.

The 3G market is oligopolistic in nature. In order to ensure benefits of competition to the consumers, it is necessary that there should be more than one service provider in each area of services and services from different providers are interchangeable. So the basic operation of the market needs MNP also as a prerequisite. In the United Kingdom, MNP was introduced in 1999 but the progress was slow

due to cost consideration and networking problem among operators. In Japan also it was introduced for both 2G and 3G.

Despite close monitoring by the regulator, the telecom industry may have the problem of tacit collusion. Tacit collusion is more likely in view of the practice of infrastructure sharing in this industry. One measure that can prevent tacit collusion is to have around six players in the market (Selten 1973). Whatever be the allegation of corruption of distribution of 2G licensing on first-come first-served basis in India during 2007–2008, one major benefit that emerged for the Indian consumers was price reduction following price competition among a large number of telecom players. Scrapping of licenses by Supreme Court order is likely to reduce competition in the market.

In the context of emergence of 3G technology, pricing of services of two coexisting technologies assumes significance. If we discuss different tariff plans of different companies abroad, we observe that usually same price is charged for 2G and 3G for standard use. Vodafone offers same discount services for 2G and 3G services. NTT DoCoMo charges similar rate for voice communications to facilitate smooth migration from 2G to 3G. Hutchison 3G in Sweden offers Blue Plan and Green plan with different rates for standard use and a specialized use like video calls. In the UK, these companies offer no monthly subscription but there is a monthly rate and charges different prices for 'add-ons'(Add Text, football, entertainment, Add Unlimited, etc.). In South Korea, SK Telecom uses the varying tariff option with use of 3G services. Singtel (Singapore operators) charges 3G data services at a rate 30 % lower than that of its GPRS service. In Singapore, they are offering different content services with different rates.

So we can figure out two strategies practiced abroad, (a) low monthly fixed charge with incremental charge with value-added uses and (b) high monthly fixed charge with 3G packages. Revenue sharing with content providers will reduce the price level that will induce the customers to adopt the 3G-services. Similarly, more advertisement revenue from mobile Internet can also reduce the charge for 3G service. Multi-part tariff system is applied in different countries and it is seen that average revenue per user (ARPU) provides a benchmark measure for mobile operators. ARPU includes revenues from traffic and does not include revenues received from handsets or advertisements. According to industry, analysts 3G revenue generation from advertisement is a long run phenomenon. It can be concluded that 3G businesses have to depend on volume rather than high prices for revenue. Discount plan at the inception period (first to subscribe/registration) may prove to be helpful to increase the subscriber base. Identification of consumers with different needs and pricing according to their demand will also be effective. To facilitate a smooth migration from 2G—charges for voice communications under 3G should be similar to 2G (as followed by the NTT DoCoMo, Japan).

The pricing policy is very much dependent on the demand elasticity and propensity to spend on a particular service. So the identifying preference and reservation price of the buyers is a big problem in the inception stage. Here, we can only assume that if the price tag is same for voice and data communication under

2G and 3G, then people with 3G-enabled handsets are likely to switchover from 2G to 3G but the demand for value-added services is uncertain in a developing economy like India. It may take time to become viable in a country like India.

The Indian Scenario

The saga of blossoming of Indian telecommunication industry into a huge network of the present has begun since 1985 when Division Indian was separated from the Department of Posts and Telegraph and a new Department of Telecommunications (DoT), operating under the Ministry of Communications was formed in order to regulate, maintain, and serve as primary policy-maker for the telecom industry. Subsequently, India, like many other countries of the world, have adopted a gradual approach to telecom sector reform through selective privatization and managed competition in different segments of the telecom market. The process of privatization started in 1991 when the government separated fixed and value-added services like cellular telephony, paging, internet etc., and permitted private operators to enter the market in value-added market. The National Telecom Policy, released in May of 1994 formally opened up the basic telephony for private sector investment. The Telecom Regulatory Authority of India (TRAI) was constituted in 1997 as an independent regulator in this sector. Competition was also introduced in national long distance (NLD) and international long distance (ILD) telephony. Presence of the government in this sector as a direct service provider was further reduced with the incorporation of Bharat Sanchar Nigam Ltd. (BSNL) as a public limited company in 1 October 2001.

In December 1991, the government invited competitive sealed bids for two non-exclusive digital mobile licenses, for a 10-year period, extendible by 5 years, for the four metropolitan cities of Mumbai, Delhi, Calcutta, and Madras. The license specified the use of GSM standards for offering cellular services. The annual license fee for the first 3 years was a given parameter, while the license fee from the fourth year onwards was fixed at Rs. 5,000 per subscriber (based on unit call rate of Rs. 1.10) subject to a minimum total amount. Along with the license fee, call charges were also a given parameter and the bidding was for the lowest rental to be charged from customers. The evaluation was on the basis of financial strength, experience of the partners, committed rollout, and lowest rentals. At the end of the tender process eight licenses were issued in October 1994.

In the first bidding round for auction the government invited bids for each circle for basic wireless services but when the bids were opened in August 1995, it was found out that the winning bids were extremely high and unsustainable. There was a complaint that the government was more concerned with raising revenue rather than speedy roll out of net work. The second bidding round also faced major problems. As more lucrative circles were already awarded in the first round, there was lack of enthusiasm and only six bids were received. Naturally initial service

rollout was slow, as a result of narrow licensing conditions and the high cost of license fees.

The initial auction process was alleged to lack transparency and legal disputes ensued following the auction. Ultimately, eight licenses were issued in 1995 but bidders could realize subsequently that they overestimated demand and revenue generation. This is called the winners' curse problem in the literature—which is the chief bane of auction. The government then relieved them of the obligation to pay further installments of their committed license fees and allowed them to move to a revenue sharing regime.

In the year 1999, the government released the National Telecom Policy in 1999 (NTP 99) in order to help the transition and update policy and auction stipulations to encourage faster telecom development, essentially increasing the number of players in each service area. In the same year, the government decided to go technology neutral, in consideration of rollout of Code Division Multiple Access (CDMA) technology. But this led to a protracted legal battle since 2000. GSM-based operators who had paid a significant license fee (2.5 billion dollars) contested this decision of allowing CDMA players to provide mobile telephony. This legal battle continued for 3 years and was finally resolved when the government introduced unified licensing making cellular services technology neutral and allowing CDMA players to provide full mobility after payment of an entry fee equal to what the GSM operators had paid.

The NTP 99 for the first time explicitly recognized that the number of cellular operators in a given geographical area would necessarily be limited by the availability of spectrum. According to NTP 99, based on the immediately available frequency spectrum band, apart from the two private operators already licensed, DOT/MTNL would be licensed to be the third operator in each service area. In order to ensure a 'level playing field' between different service providers in similar situations, license fee was payable by DoT also. 'However, as DoT is the national service provider having immense rural and social obligations, the Government reimbursed full license fee to the DoT.'

The above history shows that auctioning was never smooth in the country and in regards to spectrum allocation the Indian Spectrum Management Committee (created in May 1999) failed to specifically improve the effectiveness of the auction itself. It was decided in 2003 to go for first-come first-serves basis to allocate spectrum. In 2005, under the new guidelines, restriction on the number of entrant in a circle was withdrawn. Allocation of spectrum was subject to availability as per this guideline. In 2008, sharing of infrastructure was allowed and the access deficit charges for BSNL for rural services were abolished. To broaden the service network, nine telecom companies were issued spectrum and licenses for second generation (2G) mobile phone services at Rs. 1,658 crores for a pan-India operation. As many as 122 circle-wise licenses were issued. But subsequently complaints of corruption and grafts in course of giving the airwaves cheap, that too in the controversial manner of first-come first-served basis occupied the center stage and it was alleged that the exchequer had lost billions of dollars in this 2G allocation. Later, based on the auction of airwaves for third generation (3G)

services, which got nearly $15 billion to the exchequer, and that for broadband access, which fetched over $8.5 billion, the notional loss was estimated at $38 billion to the exchequer by the CAG. In 2010, the 3G auction took place that brought huge revenue to the government. At the end of 2010, MNP was introduced.

One problem that has gripped the Indian telecom industry is the continuous decline of the incumbent public firm BSNL and MTNL in terms of profit and market share. From a mighty profit making firm BSNL is now registering huge losses. The market share of BSNL and MTNL, the two PSUs, was more than 85 % at the time of formation of BSNL, but it has since declined to <20 %. The share of BSNL alone has dropped from about 20 % in March 2006 to less than 10 % in 2011. Not only has the market share been declining but BSNL's subscribers have lower average revenue per user (ARPU) than the average subscriber of private sector operators. Our primary survey has shown that BSNL has poor image in terms of office service, complaint handling but enjoys better image than price operators in case of transparency. Our analysis of the BSNL performance shows that it is constrained by poor image and also by higher cost of operations owing to a bloated work force. Much more damage was caused by huge increase in wage during 2008 that played a signal role in converting the profitable company into loss making. This provides an empirical validation of our theoretical assumptions that public firm has poor brand image and higher cost. But BSNL was also troubled by many anti incumbency decisions of the government. The access deficit charge that used to be given for loss of BSNL on rural telephony was abolished. Many times tenders for procuring modernization equipments were delayed. In consistence with the result of our model, the earliest allocation of low level of spectrum (2.1 MHz in place of present level of 5 MHz) for rolling out 3G ahead of others did not help BSNL.

Finally, we briefly mention the 2G scam. India is divided into 22 telecoms zones and there are a total 281 zonal licenses in the market. According to the telecom policy of India, when a license is allotted to an operator some start-up spectrum is bundled along with it. The policy had not a provision for auctioning the spectrum. In year 2008, 122 new second generation (2G) Unified Access Service (UAS) licenses were given to telecom companies at the price of 2001 and on first-come first-served basis. As per the charge sheet filled by Central Bureau of Investigation (CBI), several rules were violated and bribes were paid to favor certain firms while awarding 2G spectrum licenses. On 2 February 2012, the Supreme Court of India declared allotment of spectrum as "unconstitutional and arbitrary" and quashed all the 122 licenses issued in 2008.

This judgment will have serious implication for the Indian telecommunication industry. First of all, this will affect the market structure if the telecom firms who received licenses in 2008 failed to remain in the market. This may give advantage to the firms who received licenses before 2001. Second, hefty auction charge may increase the price of telecom services by cost escalation and reduction in the number of firms. Selten (1973) has shown that in case of cartel in imperfect competition, four are small and six are many, implying that it will be easy to form

a cartel if the number of firms in the market is less than four. So the Indian telecommunication industry may witness tacit collusion in future. This episode opens up an important area of research where legal judgment of regulatory bodies may impact the market structure.

Conclusion

Telecommunications is a difficult industry to understand and analyze on account of the involvement of critical natural resources like spectrum, fast changing technology, and the history of initial monopoly of public firms and subsequent entry of private firms. These two issues of efficient spectrum use and technology related uncertainty have left telecom industries in almost all countries in the world under a regulatory regime. So it is not a simple mixed oligopoly but mixed oligopoly with regulation. Our theoretical results show that in order to make this industry maximize social welfare, several issues like optimum allocation of spectrum, efficiency improvement of the public firm, smooth transition to new technology, and speedy roll out of infrastructure are to be sorted out. The rules framed in this regard may lead to legal disputes, and therefore dispute management is also a serious task of the regulator.

In the process of reforming and invigorating the public firm, four points are to be kept in focus. First of all, it is to be ensured that in the over enthusiasm of liberalization and privatization, anti incumbent measures are not imposed on the public firm. As our theoretical setup shows, the public firm is the most important instrument in the way to maximize social welfare. So the cost borne by social obligation placed on the public firm should not be glossed over in the making of the policies. Second, as The Economist (2012) has shown state run corporatized firms of the emerging markets of China, Russia, and Latin America are occupying important slots in the global corporate world and these state-directed firms are no longer seen as a way station on the road to liberal capitalism but as sustainable instruments of development. This potentiality of state-run firms should also be kept in consideration while framing policy decisions. Third, public firm is by nature risk minimizing and slow in decision making on account of bureaucratic system of management. So they are not expected to do very well in uncertain atmosphere, as our model has shown. While designing policy, this factor should be kept in mind. Otherwise, so-called advantage granted to public firm may turn into a disadvantage. Fourth, this firm has a baggage of product and cost disadvantage from their initial incarnation. This history is to be reversed with careful planning.

When we look at the Indian telecom industry we find that despite all the uncertainties and delay in policy formulations, litigations over spectrum allotment and the barrage of allegations of corruptions, the industry has achieved phenomenal growth over a span of 15 years. The early trouble with the auction system has made the government more interested in discretionary allocation, but it is good that they have switched over to auction in case of 3G. It was apprehended in an early

paper (Datta and Chaudhury 2009) that, discretionary allotment in a situation of different technologies might open the floodgates of corruption and legal wrangling which would stifle the growth of this vital sector. Unfortunately, that apprehension of allegation of corruption and legal tangles proved correct. So spectrum allocation has perhaps no alternative other than auction. Auction mechanism has successfully been tried out in many countries. Further, the argument that auction raises the cost of services in the industry can be taken care of by using part of the auction revenue to subsidize infrastructure investment in telecommunication (Datta 2012). We have also shown (Datta and Chatterjee 2011) that auction need not be per se efficient but with number portability, auction can prove to be a better measure for spectrum allocation. It is good that number portability has been introduced now and this has enhanced competition in the process. Now TRAI should be vigilant that the firms do not resort to tacit collusion to reduce the degree of competition in the industry. Revenue sharing is also a good option and it has helped the growth of telecom industry in India at a critical period. However, solution based on linking revenue sharing with additional spectrum will have its own limit.

India's domestic telecom market has been attracting huge investments. The 3G auctions have generated huge revenue for the government. The telecom industry's total revenue is estimated to be around 3.6 % of total GDP of the country. So it is the responsibility of not just the policy-makers but the entire society to make sure that the growth of this crucial sector does not get bogged down in the quagmire of endless political controversy.

Appendix

Telecommunications Bertrand model

Let us show the implication of the public and the private firm playing a game of price competition. In this mixed oligopoly framework, we maintain the assumption that the public firm (Y) is relatively inefficient compared to the private firm (X). We also maintain the assumption that the public firm maximizes social welfare.

Let the utility function of the representative consumer is

$$U = x + y - \frac{1}{2}(x^2 + 2bxy + y^2)$$

b represents the degree of substitutability. $b = \in (0, 1)$

From the above utility function we get the following inverse demand functions

$$p_x = 1 - x - by$$
$$p_y = 1 - bx - y$$

From the above we get direct demand function as under.

$$x = \frac{1 - b + bp_y - p_x}{1 - b^2}, \quad y = \frac{1 - b + bp_x - p_y}{1 - b^2}$$

The private firm (X) has zero marginal cost and the public firm (Y) has a marginal cost of c. $1 > c > 0$.

c measures the degree of inefficiency of the public firm.

The public firm maximizes social welfare, which is a sum of consumer's surplus and firms' profit.

$$SW = U - p_x x - p_y y + \pi_x + \pi_y$$

$$SW = x + y - \frac{x^2 + y^2}{2} - bxy - cx$$

Replacing x and y from equation and finding out first-order condition of social welfare maximization by the public firm and profit maximization by the private fir we get the following optimum values.

$$p_y^* = \frac{b(1 - b) + 2c}{2 - b^2}$$

$$p_x^* = \frac{1 - b + bc}{2 - b^2}$$

$$y^* = \frac{1 - b - c}{1 - b^2}$$

$$x^* = \frac{1 - b + bc}{(1 - b^2)(2 - b^2)}$$

The public firm will produce when inefficiency is not very large under the condition

$$1 - b > c$$

If $1 - b < c$, Bertrand competition leads to limit pricing equilibrium.

It is observed that if the inefficiency is zero, the public firm definitely charges a lower price. But if inefficiency rises the survival of the public firm will be at stake. In that case the private firm will prefer to get into price competition.

For a fuller treatment of comparison of Cournot and Bertrand models in mixed oligopoly see Ghosh and Mitra (2010) and Choi (2012).

References

Board S (2007) Bidding into the red: a model of post-auction bankruptcy. J Finance LXII(6):2695–2723

Buehler S, Dewenter R, Haucap J (2005) Mobile number portability in Europe, discussion paper no. 41. Department of Economics, University of the Federal Armed Forces Hamburg

Buehler S, Haucap J (2004) Mobile number portability. J Ind Competition Trade 4:223–228

Bulow J, Levin J, Milgrom P (2009) Winning play in spectrum auctions. NBER working paper no. 14765
Burguet R (2005) License, allocation and performance in telecommunication market. Working paper, Institute for Economic Analysis, Barcelona, Spain
Cave M, Valetti T (2000) Are spectrum auctions ruining our grandchildren's future? Info: J policy regul strategy Telecommun inf 2(4):347–350
Choi K (2012) Cournot and Bertrand competition with asymmetric costs in a mixed duopoly revisited. MPRA paper no. 37704
Datta D (2012) Spectrum auction and investment in telecom industry—a suggested policy. Vikalpa 37(1):19–30, IIM, Ahmadabad
Datta D, Chatterjee S (2011) Optimal spectrum allocation market structure and subsidization. Telecommun 60(1) BSNL, Sanchar Vikas Bhawan
Datta D, Chatterjee S (2012) BSNL-consumer perception and necessary remedial measure, forthcoming in Telecommunications 60(6) BSNL, Sanchar Vikas Bhawan
Datta D, Chowdhury SM (2009) Indian telecom: regulation, spectrum allocation and dispute Management. IIMB Mgmt Rev 21(4):287–296
De Fraja G, Delbono F (1989) Alternative strategies of a public enterprise in oligopoly. Oxf Econ Pap 41:302–311
De Fraja G, Delbono F (1990) Game theoretic models of mixed oligopoly. J Econ Surv 4(1):1–17
Demsetz H (1968) Why regulate utilities? J Law Econ 11:55–66
Economides N (2004) Telecommunications regulation: an introduction. Working paper 04–20, NET Institute
Elliott R (2002) Wireless information management. Inf Mgmt J 36(5):62–66
Ghosh A, Mitra M (2010) Comparing Bertrand and Cournot outcomes in mixed market. Econ Lett 109:72–74
Goldberg VP (1976) Regulation and administered contracts. Bell J Econ 7:426–448
Greve T, Keiding H (2010): Regulated competition under increasing returns to scale. University of Copenhagen, Deptarment of Economics, discussion paper no. 11–01
Haan MA, Toolsema LA (2003) License auctions when winning bids are financed through debt. Working paper no. 200310, University of Groningen, CCSO Centre for Economic Research
Harmer JA, Friel CD (2001) 3G products—what will they enable? BT Technol J 19:24–31
Jonason A, Eliasson G (2001) Mobile internet revenues: an empirical study of the mode I- portal. J Internet Res 4:3–27
Kalakota R, Robinson M (2002) M-business vs. M-commerce. http://www.ebstrategy.com/mobile/articles/mbus_vs_mcomm.htm
Klemperer P (1987a) Markets with consumer switching cost. Quart J Econ 102(2):375–394
Klemperer P (1987b) The competitiveness of markets with switching costs. RAND J Econ 18:139–150
Klemperer P (1988) Welfare effects of entry into markets with switching costs. J Ind Econ 37:159–165
Klemperer P (1995) Competition when consumers have switching costs: an overview with applications to industrial organization, macroeconomics, and international trade. Rev Econ Stud 62:515–539
Lange O (1936) On the economic theory of socialism. Rev Econ Stud 4(1):53–71
Liu Z (2005) Stackelberg leadership with demand uncertainty. Manag Decis Econ 26:345–350
Lu Y, Poddar S (2005) Mixed oligopoly and the choice of capacity. Res Econ 59:365–374
Malik P (2004) Indian telecommunications policy and regulation: impact on investment and market structure. Discussion paper WDR 0304
Maskin E (2000) Auctions, development and privatizations: efficient auctions with liquidity constrained buyers (Alfred Marshall lecture). Eur Econ Rev 44:667–681
Matsumura T, Ogawa A (2012) Price vs. quantity in a mixed duopoly. Econ Lett 116(2):174–177
McMillan J (1994) Selling spectrum rights. J Econ Perspect 8(3):145–162
Merrill W, Schneider N (1966) Government firms in oligopoly industries: a short run analysis. Quart J Econ 80:400–412

Milgrom PR, Weber RJ (1982) A theory of auctions and competitive bidding. Econometrica 50:1089–1122
NERA/Smith (1998) Feasibility study and cost benefit analysis of number portability for mobile services in Hong Kong, Final Report OFTA. NERA/Smith, London
Nishimori A, Ogawa H (2004) Do firms always choose excess capacity? Econ Bull 12:1–7
Oestmann S (2003) Mobile operators: their contribution to universal service and public access, Intelecon Research & Consultancy Ltd. available at http://www.inteleconresearch.com
Oftel (1997) Economic evaluation of number portability in the UK mobile telephony market, a report by Office of Telecommunications, London
Porter ME (1985) Competitive strategy: techniques for analyzing industries and competitor. Free Press, New York
Selten R (1973) A simple model of imperfect competition where four are few and six are many. Int J Game Theory 2(3):141–201
Sengupta A, Tauman Y (2011) Inducing efficiency in oligopolistic markets with increasing returns to scale, mathematical social sciences. Elsevier 62(2):95–100
Shy O (2001) The economics of network industries. Cambridge University Press, Cambridge
Sidak JG, Spulber DF (1999) Deregulatory takings and the regulatory contract: the competitive transformation of network industries in the United States. J Comp Econ 27(1):198–201
Singh N, Vives X (1984) Price and quantity competition in a differentiated duopoly. Rand J Econ 15:546–554
The Economist (23 Jan 2012) The rise of state capitalism
Viard VB (2004) Do switching costs make markets more or less competitive? The case of 800-number portability. Stanford Graduate School of Business research paper series, Paper no. 1773(R1) http://ssrn.com/abstract=371921
Virmani A (2000) A communication policy for the 21st Century. Econ Polit Wkly XXXV(23):1907–1910
Virmani A (2004) Competitive access to telecom: Spectrum policy and M & A. Econ Polit Wkly XXXIX(7):14–20
Vives X (1999) Oligopoly pricing: old idea and new tools. MIT Press, Cambridge and London
Voice and Data. Cyber Media Publications, Various Issues
Wilson RB (1992) Strategic analysis of auctions. In: Aumann RJ, Hart S (eds) Handbook of Game Theory with economic applications, vol 1. North Holland, Amsterdam
Zheng CB (2001) High bid, broke winners. J Econ Theory 100:129–171

Web sites

(2003) NTT DoCoMo Inc., Website http://www.nttdocomo.com
http://www.trai.org
http://www.dot.india.in
http://www.itu.int/home/index.html
http://www.bsnl.co.in

Chapter 12
Market Structure of Crude Steel Industry and India's Position in the Era of Globalization

Tiyas Mallick

Introduction

Crucial to the development of any modern economy, steel is a product of large and technologically complex industry having strong forward and backward linkages in terms of material flow and income generation. All major industrial economies are characterized by the existence of strong steel industry and the growth of many of these economies has been largely shaped by the strength of the steel industries in their initial stages of development. So being at the center of attraction, in a situation of high growth of world economy, the steel industry is likely to have a bright prospect. Expanding automobile industry, transportation, and infrastructural projects, linkages of steel industry with different sectors, and crude steel being the first tier of steel making process it becomes extremely important to focus on this industry for analyzing the status as well as trend.

In the era of globalization, the steel industry and its market structure have and are undergoing seismic change. With competition considered as one of the founding principles of globalization, it is conjectured that globalization would lead to greater competition that would evidently be marked by efficiency gain. We interrogate this claim with respect to crude steel across various countries and try to find out whether the opening up of crude steel to global competition across countries leads to a growing competition or concentration. Our results are somewhat counter-intuitive since we find a divergence from the mentioned conjecture, namely that global crude steel market shows evidence of growing concentration instead of greater competition. In this regard, Herfindahl–Hirschman Index (H) is used to identify the changing market pattern of global crude steel from moderate concentration to high concentration. Results indicate that while China has emerged as the major global power in this market, India too is becoming an important

T. Mallick (✉)
Tata Consultancy Services, Bengaluru, India
e-mail: tiyas_mallick@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_12, © Springer India 2013

player albeit still far behind China. I shall end my discussion by briefly examining India's new position in the global crude steel industry and the challenges it faces in the post-globalization phase.

Changing Global Crude Steel Market Structure

The industry went through one of its most difficult phases between 1997 and 2001, as it faced severe recession in the global economy leading to imbalance among capacity, demand, and production. Break up of USSR and Asian Financial Crisis created a significant impact on crude steel industry worldwide. Before that USSR was huge net importers of crude steel. Due to disintegration of USSR, many new countries came into existence and some had turned to net exporter of crude steel.[1] On the other hand, due to Asian financial turmoil, demand of steel within the effected countries fell. These events coupled with similar developments across the world led to a situation where producers had much higher capacity than they could sell. Prices of steel during this period touched a 20-year low (with Hot Rolling Steel going below $200 mark in 2001[2]) and most producers made heavy losses. Many companies were forced to shutdown leading to massive redundancies. New capacities became uneconomical and excess. Fresh expansion plans had to be abandoned as financial sector withdrew support from the steel sector. The period also witnessed major steel producing nations resorting to tariff and non-tariff barriers to safeguard their domestic industry. Pushed to the wall, by 2001, advanced countries were forced to call for a global agreement organized by the OECD to limit the world output of steel and stem the slide in steel prices.

Nevertheless, there were a few "winners" in this difficult environment, and this led the World Steel Dynamics (WSD) to identify for the first time the "world-class" steel makers in the post-globalization phase. By that time, the pricing environment had improved dramatically and WSD could confidently identify companies that deserve to be called "world-class steel makers". It was spelled out that management in each case has implemented changes in the structure of the company that seem to position it well for the future. The importance of this announcement was that these companies in effect were, from then on, to be considered as setting the benchmark for other steel makers in their home countries and internationally. This information also became the benchmark through which it was possible to identify the relative competitive strength of countries in the steel market and which ones were and are emerging stronger. The companies that WSD has selected in 2005 are:

[1] *Source* http://jpcindiansteel.nic.in/profile.asp.

[2] *Source* http://www.steelonthenet.com/kb/world-hot-rolled-coil-prices-2001-2010.html.

- One in Australia—Blue Scope Steel.
- Two in Brazil—CSN and Gerdau *(Note*: *CST, which had also been listed as a "world class" steel maker, was recently merged into Arcelor).*
- Three in the European Union—Arcelor, Corus, and Thyssen/Krupp.
- One in Canada—Dofasco.
- Five in China—Anshan, Baosteel, Maanshan, Shagang, and Wuhan.
- One in India—Tata Steel.
- Two in Japan—Nippon Steel and JFE.
- One in Russia—Severstal.
- One in South Korea—POSCO.
- One in Taiwan—China Steel.
- Three in the United States—Nucor, SDI, and United States Steel. International Steel Group (ISG) was removed from the list after being acquired by Mittal Steel.
- One global company—Mittal Steel, with operations in Algeria, Canada, Czech Republic, France, Germany, Hungary, Indonesia, Kazakhstan, Macedonia, Mexico, Poland, Romania, South Africa, Trinidad, and the United States.

14 out of 22 major crude steel producers are from India China, Japan, US, Russia, Germany, and South Korea. Therefore, we concentrate our analysis in these seven major players.

Objective of the Study

In order to analyze the changing market structure of global crude steel industry, the period from 1995 to 2011 is considered and an attempt made to explore whether global crude steel industry is exhibiting increased competition or is it producing consolidation and concentration. Further this paper is also analyzing position of Indian crude steel industry within this global market.

Literature Review

Existing literatures give the idea of different concentration ratios (Sleuwaegen and Dehandschutter 1986; Schmalensee 1977; Kelly 1981). Of particular importance is the paper by Sleuwaegen and Dehandschutter (1986). Sleuwaegen and Dehandschutter made an analysis on k firm concentration ratio (C_k) and Herfindahl–Hirschman Index (H). These measures can be helpful to determine the market structure of crude steel, that is, in finding out whether the market exhibits concentration or competition. In their study k denoted the largest firm. They made the analysis on US (1956) and Belgian (1981) industrial data and concluded that there was a horn-shaped relationship between the two ratios. Implication of the

conclusion is that if number k increases then that implies lower concentration within the industry on which both indices are calculated. Here between C_k and H any one can be chosen. But for lower k, industry will be highly concentrated and under this situation H is a superior measure of concentration. Consequently, H is acceptable under any circumstances as a measure of market concentration. We can consider this H Index to reveal the changing crude steel market structure in global context.

Several literatures use H index to measure the concentration of different industry like North American steel plate industry, Vermont's retail food industry, US life insurance industry, respectively (David Cummins et al. 1972; Cotterill 1986; Maasoumi et al. 2002). Maasoumi et al. (2002) used H to measure the market power in the context of antitrust litigation in the Steel Plate Industry of North America in late 1990s. It concluded that steel plate market was not subjected to anti competitive forces.

From the review of literature it is clear that none of the previous studies use H in global context in their respective field of analysis. For example, in one study, H index was used in case of US Life Insurance Industry to measure the market pattern but analysis was not made by combining US Life Insurance Industry and Global Insurance Industry (David Cummins et al. 1972). Furthermore, the literature that analyzed steel industry and tried to find out the market structure of steel mainly focused on specific finished steel product.

As such, no analysis is available on crude steel market structure or on the global nature of the market structure of crude steel which additionally marks the distinctiveness of my intervention. Given the distinctness of my focus and methodology, I intend to explore now whether global crude steel market structure is competitive or concentrated and, in that context, throw light on India's position in the crude steel industry.

Methodology and Data Source

As mentioned earlier, Herfindahl–Hirschman Index (H) can be used to identify the changing market pattern of global crude steel. Calculation of H will help to resolve whether Global Crude Steel Industry follows the general conjecture of globalization or not. Herfindahl Index (Herfindahl—Hirschman Index or H) is a measure of the size of firms in relationship to the industry and an indicator of competition among them. It is defined as the sum of the squares of the market shares of each individual firm.

$$H = \sum_{i=1}^{n} (s_i^2)$$

where, s_i is the market share of firm 'i' in the market and 'n' is the number of firms.

Three possibilities arise:

$H < 0.1$ implies unconcentrated index.
$0.1 < H < 0.18$ implies moderate concentration.
$H > 0.18$ implies high concentration.
Source (Maasoumi et al. 2002)

This is normalized Herfindahl Index.

Another form of this Index is:

$$H = (1/n) + n.V$$

where, V is Statistical Variance of the firms' share.

$$V = \left[\sum_{i=1}^{n} (s_i - 1/n)^2\right]/\mathrm{n}$$

In a special case if all firms have equal (identical) shares (that is, if the market structure is completely symmetric) in which case $s_i = (1/n)$ for all i, then $V = 0$ and $H = (1/n)$. But in a more generalized version, firms contribute unequally i.e., market is asymmetric. So there will be a significant gap between s_i and $(1/n)$, in which case V and therefore H will be high. In this paper, H index has been used to examine the market pattern of global crude steel industry.

If the number of firms in the market is held constant then a higher variance due to higher level of asymmetry among firms' shares will result in high Index value. In the global context, major crude steel producing countries can be considered as firms and world crude steel production as industry contribution. Here, India is one of the major crude steel producers.

H Index can be constructed from crude steel production data. Information on global players can be obtained from World Steel Association Data Base presented in "Metal Bulletin Research, Ferro Alloy Monthly Report" (Issue 195, November 2008). This can be further supported by Annual Reports of SAIL and TATA.

Empirical Evidences

14 out of 22 major crude steel producers are from India China, Japan, US, Russia, Germany, and South Korea. Therefore, we concentrate our analysis in these seven major players (Table 12.1).

Now each country's market share (s_i) is calculated in Appendix I, Table 12.2 by dividing each country's production by the world's crude steel production. As seven major players are considered, $(1/n) = (1/7)$. Calculation of $(s_i - (1/n))$, its square, sum of its squares is shown in Appendix I Table 12.3. V and H are calculated in Appendix II, Tables 12.4 and 12.5, respectively.

Over the period, global crude steel industry is shown to be growing with a fast changing scenario. From Appendix 1 Table 12.2 it can be seen that Japan, US, and Russia, the key players of 1995, were unable to retain their position. Share in global crude steel production falls from 13.5053 % in 1995 to 7.2210 % in 2011 in case of Japan, from 12.6545 % in 1995 to 5.7848 % in 2011 for US and for Russia it falls from 6.8590 % in 1995 to 4.6104 % in 2011. On the other hand when most of the key players' share slides, China and India has shown an upswing in their share in global crude steel output. Thus, within 1995 and 2011, share in world crude steel production of every nation falls except India and China.

Considering theoretical background of H index, here H has been used to reestablish the changing global crude steel market structure. V is calculated to get the values of H index.

Values of Herfindahl Index (H) are plotted against years in Appendix II, Graph 12.1. If the trajectory of the graph is followed then it is clear that in the year 2001, index value exceeds 0.18 i.e., in this particular year crude steel market starts to become highly concentrated. 2009 is the year of maximum concentration. There is a change in the world crude steel market structure from moderate concentration to high concentration, where 2001 seems to suggest the edge of this change. Therefore, years of globalization seem to have flattened competition somewhat leading to a more concentrated industry structure in the post-globalization phase. Another important point that comes from this analysis is that China has evolved into a crude steel giant through this transformation. Even in the phase of global meltdown in 2008 China retained remarkable share of total world crude steel output which many suggests was due to the Beijing Olympics. By no means comparable to China yet, one also cannot ignore the growing importance of India in that industry.

Position of India in Global Crude Steel Industry

The above study makes it clear that China and India has emerged as major competitors. In 1950, both the Indian and Chinese steel industries were producing finished steel at a rate of about 1 million tons per year . From that point until about the mid-1990s, the Indian steel industry grew at a much slower rate, expanding at an annual rate of 6.5 % versus approximately 10.3 % for the Chinese steel industry.

From the mid-1990s to 2007, Chinese crude steel production registered a growth of approximately 28.48 % whereas India registered that of 10.38 % (*Source* World Steel Association, World Steel in Figures 2010) Even over the period 2008 and 2009, India's crude steel production grew at the rate of 2.7 % vis-à-vis 13.5 % that of China (*Source* World Steel Association, World Steel in Figures 2010). This only factualizes the rapid emergence of China as a dominant crude steel player within the world. It is presumed that the growth of the Indian

steel industry and its share of global crude steel production could be even higher if they were not being held back by major deficiencies in fundamental areas. Investment in infrastructure is rising appreciably but remains well below the target levels set by the government due to financing problems.

Deficiencies in Indian Crude Steel Industry

Energy Supply

Power shortages hamper production at many locations. Since 2001, the Indian government has been endeavoring to ensure that power is available nationwide by 2012, albeit not very successfully. The deficiencies have prompted many firms with heavier energy demands to opt for producing electricity with their own industrial generators.

Problems Procuring Raw Material Inputs

Since domestic raw material sources are insufficient to supply the Indian steel industry, a considerable amount of raw materials has to be imported. For example, iron ore deposits are not adequate and there are problems in mining sufficient amounts of it. India's hard coal deposits are of low quality which has led to increase in hard coal imports. The rising output of electric steel is also leading to a sharp increase in demand for steel scrap. In the coming years imports are likely to continue to increase, thanks to capacity increases.

Inefficient Transport System

In India, insufficient freight capacity and a transport infrastructure that have long been inadequate are becoming increasingly serious impediments to steel development. Although the country has one of the world's biggest transport networks—the rail network is twice as extensive as China's—its poor quality hinders the efficient supply of goods. The story of transport bottleneck stretches to port facilities and airports.

Growth Opportunities in Indian Crude Steel Industry

Although China is the world's largest crude steel producer, its export potential may not be as great as overall production might suggest, due to its relatively small number of efficient producers (Baosha, Wuhan, Anshan, Shougang to name a few). China is not a significant exporter of steel as they have focused on domestic

market. In a study (Meta Bulletin Research Battelle) it was mentioned that, given average finished steel production, China consumes 96 % of finished steel. Per capita Steel consumption is low by international standard, but it is growing, showing an increase from 60 kg in 1990 to 100 kg in 1998. In 1998 the total consumption of steel exceeds 125 million MT per year. With such a large and growing steel market China not only absorbs most of its own production but is a net importer. Most finished steel production that China exports go to regional markets of Asia implying China does not yet consistently meet the international product standard.

This indicates that India has opportunity to become a substantial exporter. Indeed, over the past few years, the gap between production and domestic consumption in India suggests that India is becoming a net exporter. One cannot deny joint support extended by Central and State government to exploit country's natural resources and that would help India to strengthen its position in global crude steel market. This also may facilitate strong sustainable growth in India's crude steel industry. Moreover, the high level of competition to get ownership of iron ore reserves probably favors domestic companies since they better understand the bureaucratic and legal complexities that need to be negotiated. Interestingly, not only are existing and fledgling steel mills vying for the ore; state-owned mining and trading companies also want access to it, and then to develop new mines, perhaps, on a joint venture basis. Better cost position as compared to offshore firms also helps India. Going forward, India's lower wages and favorable energy prices will continue to promise substantial cost advantages compared to production facilities in (Western) Europe or the US. The growth prospects of the client industries are also very good. The deployment of modern production systems is increasingly enabling India to improve the quality of its steel products and thus to enhance its export prospects. However, as we mentioned earlier, the challenges too are formidable.

Conclusion

This paper explored the changing global crude steel market structure through Herfindahl–Hirschman Index. Our result points to market concentration and concludes that China has emerged as a dominant player in terms of crude steel production. While, in comparison to the pre-liberalization and pre-globalization period, India is becoming prominent in World Crude Steel market, it still lags far behind China. Infrastructural bottlenecks fettering investment seem to be holding back its true growth potential.

Appendix I

Table 12.1 Crude steel output (Million Tones)

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
India	22	23.8	24.4	23.5	24.3	26.9	27.3	28.8	31.8	32.6	38.8
China	95.4	101.2	108.9	114.6	124	127.2	150.9	182.2	222.4	280.5	348.06
Japan	101.6	98.8	104.6	93.6	94.2	106.4	102.9	107.8	110.5	112.7	112.47
South Korea	36.8	38.9	42.6	39.9	41	43.1	43.9	45.4	46.3	47.5	47.7
Germany	42.1	39.8	45	44.1	42.1	46.4	44.8	45	44.8	46.4	44.51
US	95.2	95.5	98.5	98.7	97.4	101.8	90.1	91.6	93.7	99.7	93.09
Russia	51.6	49.3	48.5	43.8	51.5	59.1	59	59.8	61.5	65.6	64.34
World	752.3	750	798.9	777.3	789	847.7	850.3	903.8	969.1	1067	1102.9

	2006	2007	2008	2009	2010	2011
India	42.62	49.39	57.8	62.8	68.3	72.2
China	421.46	487.63	500.3	573.6	626.7	683.3
Japan	116.22	120.2	118.7	87.5	109.6	107.6
South Korea	48.53	51.18	53.6	48.6	58.5	68
Germany	47.22	48.54	45.8	32.7	43.8	44.3
US	98.54	98.21	91.4	58.2	80.6	86.2
Russia	70.47	72.49	68.5	60	66.9	68.7
World	1215.04	1309.31	1326.5	1219.7	1413.6	1490.1

Source Metal Bulletin Research, Ferro Alloys Monthly, Issue 195, November 2008
http://en.wikipedia.org/wiki/List_of_countries_by_steel_production (based on data provided by World Steel association)

Table 12.2 Share (%) of each nation in world crude steel output i.e. s_i

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
India	2.9244	3.1733	3.0542	3.0233	3.0798	3.1733	3.2106	3.1865	3.2814	3.0553	3.5179
China	12.6811	13.4933	13.6312	14.7433	15.7161	15.0053	17.7467	20.1593	22.9491	26.2887	31.5578
Japan	13.5053	13.1733	13.0930	12.0417	11.9392	12.5516	12.1016	11.9274	11.4023	10.5623	10.1974
South Korea	4.8917	5.1867	5.3323	5.1332	5.1965	5.0843	5.1629	5.0232	4.7776	4.4517	4.3248
Germany	5.5962	5.3067	5.6327	5.6735	5.3359	5.4736	5.2687	4.9790	4.6228	4.3486	4.0356
US	12.6545	12.7333	12.3295	12.6978	12.3447	12.0090	10.5963	10.1350	9.6688	9.3440	8.4402
Russia	6.8590	6.5733	6.0708	5.6349	6.5272	6.9718	6.9387	6.6165	6.3461	6.1481	5.8336

	2006	2007	2008	2009	2010	2011
India	3.5077	3.7722	4.3573	5.1488	4.8316	4.8453
China	34.6869	37.2433	37.7158	47.0280	44.3336	45.8560
Japan	9.5651	9.1804	8.9484	7.1739	7.7533	7.2210
South Korea	3.9941	3.9089	4.0407	3.9846	4.1384	4.5635
Germany	3.8863	3.7073	3.4527	2.6810	3.0985	2.9730
US	8.1100	7.5009	6.8903	4.7717	5.7018	5.7848
Russia	5.7998	5.5365	5.1640	4.9192	4.7326	4.6104

Table 12.3 Calculation of $((S_i - (1/n))^2)$

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
India	0.01291	0.01235	0.01261	0.01268	0.01256	0.01235	0.01227	0.01232	0.01211	0.01261	0.01159
China	0.00026	0.00006	0.00004	0.00002	0.00020	0.00005	0.00120	0.00345	0.00751	0.01441	0.02983
Japan	0.00006	0.00012	0.00014	0.00050	0.00055	0.00030	0.00048	0.00056	0.00083	0.00139	0.00167
South Korea	0.00882	0.00828	0.00802	0.00838	0.00826	0.00847	0.00832	0.00858	0.00904	0.00967	0.00992
Germany	0.00755	0.00806	0.00749	0.00742	0.00801	0.00777	0.00813	0.00866	0.00934	0.00987	0.01051
US	0.00027	0.00024	0.00038	0.00025	0.00038	0.00052	0.00136	0.00172	0.00213	0.00244	0.00342
Russia	0.00552	0.00595	0.00675	0.00748	0.00602	0.00535	0.00540	0.00588	0.00630	0.00662	0.00714
Total	0.03538	0.03507	0.03543	0.03674	0.03598	0.03480	0.03715	0.04117	0.04726	0.05702	0.07409

	2006	2007	2008	2009	2010	2011
India	0.01162	0.01105	0.00986	0.00835	0.00894	0.00891
China	0.04162	0.05270	0.05490	0.10721	0.09029	0.09967
Japan	0.00223	0.00261	0.00285	0.00506	0.00427	0.00499
South Korea	0.01059	0.01077	0.01050	0.01061	0.01030	0.00945
Germany	0.01081	0.01119	0.01174	0.01347	0.01252	0.01280
US	0.00381	0.00460	0.00547	0.00905	0.00737	0.00723
Russia	0.00720	0.00765	0.00832	0.00877	0.00913	0.00936
Total	0.08789	0.10058	0.10362	0.16251	0.14280	0.15241

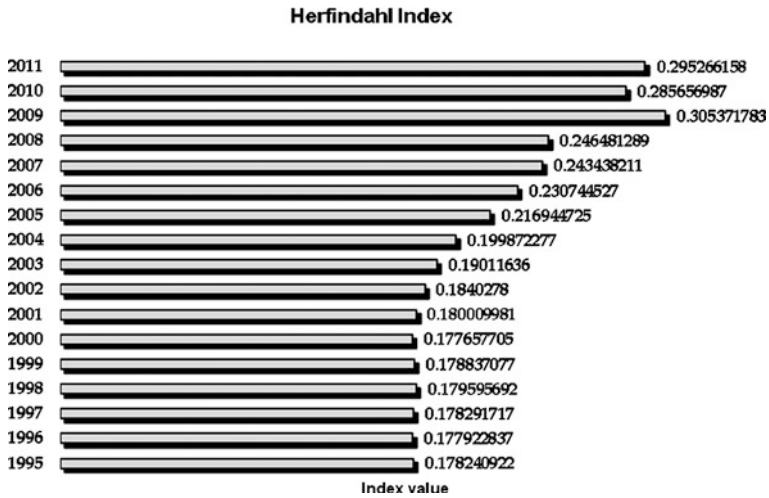
Appendix II

Table 12.4 $V = [\text{Sum}\{((S_i\text{-}(1/n))^2)\}/n]$

1995	0.005054826
1996	0.005009385
1997	0.005062082
1998	0.005248364
1999	0.005139991
2000	0.004971509
2001	0.005307548
2002	0.005881522
2003	0.006751317
2004	0.008145019
2005	0.01058394
2006	0.012555341
2007	0.014368724
2008	0.014803449
2009	0.023216377
2010	0.020399978
2011	0.021772716

Table 12.5 $H = (1/n) + n*V$

1995	0.178240922
1996	0.177922837
1997	0.178291717
1998	0.179595692
1999	0.178837077
2000	0.177657705
2001	0.180009981
2002	0.1840278
2003	0.19011636
2004	0.199872277
2005	0.216944725
2006	0.230744527
2007	0.243438211
2008	0.246481289
2009	0.305371783
2010	0.285656987
2011	0.295266158



Graph 12.1 Crude Steel Market & Herfindahl Index

References

Cotterill RW (1986) Market power in the retail food industry: evidence from vermont. Rev Econ Stat 68(3):379–386

David Cummins J, Denenberg HS, Scheel WC (1972) Concentration in the US life insurance industry. J Risk Insur 39(2):177–199

Kelly AW (1981) A generalized interpretation of the Herfindahl index. South Econ J 48(1):50–57

Maasoumi E, Prowsw SD, Slottje DJ (2002) Measuring market power in the steel industry. http://www.steelonthenet.com/kb/world-hot-rolled-coil-prices-2001-2010.html. http://en.wikipedia.org/wiki/List_of_countries_by_steel_production. (based on data provided by World Steel association) http://jpcindiansteel.nic.in/profile.asp

Metal Bulletin Research (2008) Ferro Alloys Monthly. Issue 195, Nov 2008

Schmalensee R (1977) Using the H index of concentration with published data. Rev Econ Stat 59(2):186–193

Sleuwaegen L, Dehandschutter W (1986) The critical choice between the concentration ratio and the H—index in assessing industry performance. J Ind Econ, XXXV, pp 193–208

Chapter 13
Nabadiganta: A New Horizon? Patriarchy, Globalization, and Women's Agency in the IT Sector

Mousumi Dutta and Zakir Husain

Introduction

The integration of the Indian economy with the global economic system from the 1990s created substantial employment opportunities at the high ends of the formal sector. The Information Technology (IT) sector was one of the fastest growing of these sectors. Its share in India's GDP increased from 1.2 % in 1997–1998 to 5.2 % in 2006–2007. The IT and ITES sector also constitutes the largest organized private sector employer in the country, employing about 12 % of the workforce in the organized private sector (NASSCOM and Deloitte 2008).

Given the need to fill the newly created posts, gender discrimination became a luxury that few IT and ITES companies could afford. The entry of women has increased rapidly in recent years—a MERCER-NASSCOM study found that in 2007–2008 alone employment of women increased by 60 %, so that by the end of 2008 there were 6.7 million women working in the sector.[1]

It was initially believed that the rapid growth of the IT sector would generate less exploitative avenues of employment for women. Further, economic empowerment would strengthen the bargaining power of women within the household and improve their status. Clarke and Sekher (2007) argue that growth of the IT sector has led to a partial reversal of the tendency to devalue a daughter within Indian families. Shanker (2008) notes a tendency for social relations to become more gender neutral within and outside the workplace.

[1] Reported in Economic Times (13 May 2009).

M. Dutta
Department of Economics, Presidency University, Kolkata 700073, India
e-mail: dmousumi1970@gmail.com

Z. Husain (✉)
Population Research Centre, Institute of Economic Growth, Delhi 110007, India
e-mail: dzhusain@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_13, © Springer India 2013

However, the majority of studies argue that the IT sector—for all the branding of IT companies as a "global workplace"—has been unable to isolate itself from the social backdrop. As a result, relationships at work in this sector continue to be shaped by the conflicting and asymmetrical gender relationships that prevail in Indian society, so that women remain downtrodden and lack freedom of choice (Mitter and Rowbotham 1995).

This paper takes a more positive stance. We argue that in the post-globalized world, decisions are increasingly becoming individualized, though influenced by the situational and historical context. In particular, patriarchal foundations of the Indian family are weakening, enabling women workers to break out of a passive mold and trying to carve out their own destiny.

The paper is structured as follows: We start with a review of the literature on status of women workers in India's IT sector. We then present findings from a survey of women workers in Kolkata's IT sector in support of this framework.

Gender, Organization, and the Family

Although there is a considerable body of literature arguing that employment and economic independence empowers women, this proposition is also refuted in many empirical studies. Researchers studying the status of women workers in India's IT sector (viz. Mitter and Rowbotham 1995; COD 2004; Upadhya 2005) may be said to belong to the second category.

Women Workers Within the Family

To understand the basis of their arguments it is necessary to look at the reasons for entry of women workers into the labor market. The deterioration in standard of living after the Second World War, caused by hyper-inflation, had led to women from middle and low income families being allowed to join the labor force by the men within the family. Since the objective was to simply use women to augment family income, women were not freed from household responsibilities, or from adhering to the social norm of providing care services. They worked in fixed hours jobs, sacrificing their leisure hours to undertake domestic responsibilities.

This situation may be described in terms of the Game of Chicken (Folbre and Weisskopf 1998), illustrated in Fig. 13.1a.

A husband and wife return home from work, and find that they have to cook and wash the dishes. They can share the tasks, or one spouse can do the work alone. In the latter case, the person who has to do the work becomes even more tired, while the other person relaxes and gets a higher pay-off. There is another possibility—neither does the tasks, so that they both go to bed hungry. It can be seen that the

Fig. 13.1 **a** Household chores as a Game of Chicken **b** Household chores as a Game of Chicken under patriarchy

(a)

Wife \ Husband	Do household tasks	Take rest
Do household tasks	(5, 5)	(0, 10)
Take rest	(10, 0)	(2,2)

(b)

Wife \ Husband	Do household tasks	Take rest
Do household tasks	(5+ θ, 5)	(0+ θ, 10)
Take rest	(10, 0)	(2,2)

game is characterized by multiple Nash equilibria—represented by the north-east and south-west cells. However, the equilibria are unstable as the person who has to do the household chores has the incentive to defect.

The crux of the situation is that gender stereotyping of the wife as care provider in a patriarchal society encourages the husband to relax, and imposes a social obligation on the wife to provide the services. The pay-offs to the wife gets modified—under patriarchy, she gains an additional pay-off of (say) θ for adherence to social norms by doing the household chores herself. It can be seen that if θ is greater than 2, in this case, the equilibrium in the north-east cell represents a stable Nash equilibrium. The presence of patriarchal relations, therefore, modifies the pay-offs in the Game of Chicken to Fig. 13.1b, and results in the wife doing the household chores and the husband taking rest.

Studies show that sharing of household tasks is rarely practiced even in nuclear families (COD 2004; Upadhya 2005). By and large, however, gender stereotyping of household roles persists. Women accepted the practice of husbands being responsible for "outside the house" tasks (generally related to financial matters), while they looked after household chores despite their late nights.

Gender Relations at Work

The picture is more complex within the organization. Kelkar and Nathan (2002) and Upadhya (2007) argue that firms are reluctant to recruit married women with children. Apart from this there does not seem to be any other discrimination against women in the hiring process. Rather, companies like IBM, Microsoft, and NIIT have been known to set targets for recruiting women. These companies also offer incentives to recruitment consultancies fulfilling such targets (NASSCOM and Deloitte 2008). It is with regard to organizational process and career advancement that a gender difference emerges.

Studies (Kelkar and Nathan 2002) have found a glass ceiling restricting the career growth of women employees. In an extreme form it may even result in clustering of women in low level areas with lower levels of pay and with limited opportunities for growth (Rothboeck et al. 2001; COD 2004; Upadhya 2007). There are several reasons for this:

(a) The core work hours of IT companies correspond to the office time of their (off-shore) clients. Coupled with the individualization of work[2] and the need to meet over-ambitious project deadlines (Upadhya 2007), this typically results in long hours of about 14 h per day. In many cases, work continues over night. This is a major problem for women workers because of social disapproval, objection from families, domestic responsibilities, and security-related issues involved in returning late at night (Upadhya 2005; Mitter and Rowbotham 1995). As a result, women workers typically work shorter hours than men (Rothboeck et al. 2001).
(b) Women workers are often reluctant to accept on-site (particularly off-shore) assignments and travel abroad (Ramsey and McCorduck 2005). An all-India survey found that 63 % of women workers did not accept off-shore assignments (COD 2004).
(c) Women have limited access to the informal knowledge networks important for upward mobility (Upadhya 2007). This is not only because women workers have to return home early, but also owing to men workers feeling uneasy socially interacting with their women colleagues (COD 2004).
(d) Pregnancy and the responsibility of child caring impose major limitations on the career growth of women workers, particularly as they are rated even during this leave period. This leads to many women workers postponing marriage and having children. A significant proportion of women workers also leave, being unable to bear the pressure (COD 2004; Upadhya 2007).

Satisficing Approach

An alternative approach to analyzing gender issues in the IT sector is using the satisficing approach. Satisficing is an alternative to optimization for cases where there are multiple and competitive objectives and in which one gives up the idea of obtaining the "best" solution (Simon 1955). In this approach, the actor sets lower bounds for the various objectives that, if attained, will be satisfactory, or "good enough", and then seeks a solution that will exceed these bounds. The satisfier's philosophy is that in the real world there are far too many uncertainties and conflicts in values thereby negating the hope of obtaining a true optimization.

[2] The work is disaggregated into specific components which are allocated to individual members of the project team.

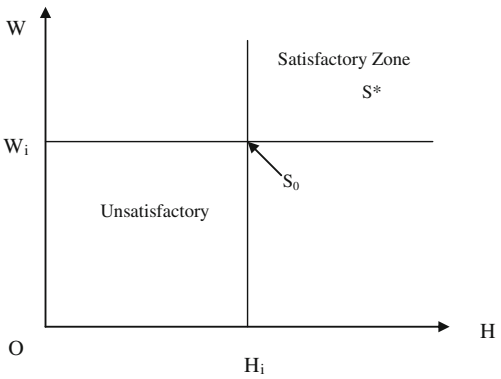


Fig. 13.2 A satisficing framework

Hence it is more sensible to set out to improve on one's current situation and do "well enough".

In the present context, for instance, women have two conflicting goals—fulfilling family commitments and attending to official duties. To what extent they are able to attain each of these targets depends upon their work, family, and social environment—the attitude of their matrimonial family, extent to which their husbands are supportive, attitude of colleagues and superiors, presence of external support, and other factors specific to the situation. In this situation, women are not able to maximize their welfare by choosing the best outcome, but have to embark on a negotiation process with their family and office. This negotiation enables them to test the structural limits on their agency; once such limits are reached, they adapt their aspirations and targets to structural realities to identify possible actions (like choice of departments, acceptance of off-shore assignments, family planning, etc.) that will enable them to balance the two conflicting objectives. This process may also be discussed in terms of a diagram (Fig. 13.2).

The two competing goals, fulfilling family commitment and attending to official duties, are represented by H_i and W_i, respectively. Initially, agents are situated at S_0, representing socially determined minimum levels of both these two goals. It can be seen that movement to the north-east quadrant (Satisfactory Zone) increases their welfare levels. However, as Simon points out, agents may not be able to identify the choices/actions that will lead to outcomes like S^*. In this situation, agents are not able to maximize their welfare by choosing the best outcome in the Satisfactory Zone. Rather they embark on a negotiation process to identify possible actions that will enable them to move to an equilibrium point in the satisfactory zone.

It should be recognized that this equilibrium point does not represent an equilibrium common to all respondents. Multiple equilibria are consistent with satisficing, and each agent chooses her equilibrium point on the basis of individual circumstances. In the remaining part of this paper—based on a primary survey of 97 women workers in Kolkata's IT sector—we shall discuss how the behavior of women workers can be analyzed in terms of the satisficing approach.

Table 13.1 Composition of sample by segment

Segment	Number	Percentage
Software development	37	38.1
Knowledge processing	25	25.8
Business processing	28	28.9
Other ITES segments*	7	7.2
Total	97	100

* Comprising of medical transcription, travel, tourism, etc

Findings

Since most companies refused access to women workers we took recourse to convenience and snowball sampling. Respondents were interviewed using a structured, mostly close-ended questionnaire. In addition, about 10 % of the respondents were interviewed in detail by the authors. The study period was June to October 2008. The composition of the sample working in different segments of the IT-industry is given in Table 13.1.

Profile of Respondents

In line with earlier studies of the IT sector it is found that the workforce is quite young. The mean and median age is 27 years. Given that the boom in Kolkata's IT sector is a recent phenomenon, this is expected. However, it imposes the limitation, that long-term effects cannot be studied in our survey.

About 46 % of the respondents are married; this proportion is higher in the software sector (65 %). The average family size is about four, for both married and unmarried respondents. About 36 % of every married respondents (that is, currently married or divorced) had one child, while 6 % had two children.

Although variations in the age of respondents are not marked, differences in work experience are observed across segments (Table 13.2). This implies that the age of entry varies sharply.

There are sharp variations in educational profile across the five segments studied (Table 13.3). In software, an essentially technical line, 84 % of respondents are technical graduates or post-graduates. In the KPO sector, 76 % are graduates or above. Interestingly, the proportion of post-graduates is low in both these segments—22 % in software companies and 24 % in KPOs. Quite a few (44 %) KPO employees also completed Management degrees. About three out of every four BPO employees interviewed is a graduate.

We also found the economic status of respondents to be high. About 70 % live in their own houses, while one out of every three respondents has a car. While the respondents from the software segment were most well-off (73 % owned a house

Table 13.2 Age and experience of respondents (in years)

Segment	Average age	Average experience in IT sector	Average age at entry
Software development	27.4	3.9	23.5
KPO*	27.2	4.6	22.6
BPO†	27.1	2.8	24.3
Miscellaneous	28.7	3.1	25.6
Total	27.4	3.7	23.9

* Knowledge process outsourcing is a form of outsourcing in which knowledge- or information-related work is carried out by workers in a different company in the same country or on an off-shore location to save cost. This typically involves high-value work carried out by highly skilled staff

† Business process outsourcing refers to the transmission of processes along with the associated operational activities and responsibilities to a third party with at least a guaranteed equal service level

Table 13.3 Educational level of respondents

Educational level	Software development	KPO	BPO	Miscellaneous	Total
HS	2.7	4.0	3.6	14.3	4.1
Graduate—general	2.7	24.0	75.0	42.9	32.0
Graduate—technical	62.2	8.0			25.8
Graduate—management	2.7	8.0			3.1
PG—general	8.1	52.0	17.9	42.9	24.7
PG—technical	21.6				8.2
Tech dip/cert		4.0	3.6		2.1

and 51 % owned a car), respondents from the Miscellaneous category appeared least well-off (only 43 % owned a house, while 14 % owned a car).

Findings indicate that women workers in the IT sector are not representative of the Indian society, but belong to an exclusive socio-economic class (Krishna and Brihmadesam 2006; Upadhya 2007).[3] The majority of respondents were from high caste, educated, and affluent urban families. As a respondent admitted,

> "I have very rarely seen women coming from a rural background … at most they come from suburbs. Women don't come from a typical rural background … father teaching in primary school, mother hasn't studied. Women in the IT sector come from the urban middle class."

As we shall see, this characteristic of the women workers is very important in determining the nature of their aspirations, and how they interacts with their environment to create a structured individuation.

Respondents identified possible reasons responsible for this bias. They opined that investment in higher education and coaching for girls is felt to be a luxury in low income and rural households. Further, such families are often averse to the

[3] In contrast, male workers come from more diverse socio-economic backgrounds.

idea of their daughters availing of hostel-based education. Reliance on campus interviews, importance placed on communication skill and ability to communicate in English[4] during the recruitment process also limits the socio-economic base from which women workers are recruited.

Agency of IT Women Workers

In-depth interviews reveal that education and economic independence provided by employment in a technical, high paying sector are crucial in defining the role of respondents within the family. Respondents felt that employment allowed them to raise questions and issues that would otherwise not have been raised; it also enabled them to avoid questions that would otherwise have been raised:

> "No matter how much lip service … people pay to equality, background, equal partnerships, at crucial moments it (remains) lip service… about independence, freedom to choose what you want to do, equal partnership, unless there is some extent of economic independence."

Another respondent said:

> "Every one in my home … gives me a lot of space … a lot of freedom …. I have my own identity as a working woman. It feels so very good that you are productive… You are having your own social status. You are known not only through your husband's name but as an associate of TCS. It gives a good feeling."

A respondent, who had given up her work after marriage to devote time to her family, said that she felt "empty headed … useless … I felt that I was not contributing anything to the family".

The combination of high levels of education and employment in the hi-tech competitive IT sector thus increases the perceived interest response[5] and perceived contribution response[6] of women workers (Agarwal 1994). The economic independence, perceived contribution response, and greater awareness of women in the IT sector give them the confidence and bargaining strength needed—what Sen (1993) refers to as an improved "fallback position"—to renegotiate their role in the family:

> "Economic freedom gives you a say. It is not deliberate, it gives you confidence … you are more confident in sticking to what you think is correct. … If you are economically independent then some questions will not rise at all - you have avoided those questions. If you are not, may be the issues are resolved in your favor, but the questions do come up."

---

[4] A high proportion of the respondents (about 66 %) had studied in English-medium schools. Only in the BPO sector is this proportion low—about half of the respondents (mostly working in smaller firms) were not from English-medium schools.

[5] The perception that one's well-being is important in the family welfare function.

[6] The perception that one is making a significant contribution to family welfare.

Another respondent mentioned that the confidence gained from working enabled her to go against her husband's wishes and insistence on having a child.

It was not that the women are always participating in decision-making. In many cases respondents are so bogged down with their official responsibilities that they deliberately did not exercise this power—"If I am so busy in work, how can I take household decisions?" However, they value the feeling that they have the power to make their voice heard.

About 80 % of respondents were satisfied with their level of control over all types of family decisions. With regard to control in household financial decisions, we found that a negligible 6–7 % did not have any say in such decisions. The socio-economic background of IT employees and the fact that women from their families have worked in earlier generations usually lowers resistance to women "neglecting the family", by reducing expectations from working women. This is an example of how aspirations of earlier generations lead to situational change, thereby giving women in subsequent generations the scope to renegotiate their role and functions within the household.

The case of Ankita (name changed), a software engineer working in a domestic company, is a typical example of this trend. She works hard to establish her reputation in her company; this entails, on several occasions, late nights at office. She tries to compensate for this by spending time with her family during weekends.[7] Frequent vacationing with her husband is also part of her "balancing strategy". Ankita is also averse to accepting on-site assignments as this will affect her work-home balance. She accepts the fact that this may affect her ratings marginally, but feels that this is a necessary cost of being able to spend time with husband, family, and friends. However, her priorities are not static but adaptable over time; after 10–12 years of marriage, she is willing to prioritize her career over family and accept even long-term on-site assignments.

Ankita's in-laws have accepted her careerism and are generally supportive. They realize that the specification of job timings has changed. While they expect that Ankita will return home by 6.30 p.m., they accept the fact that this is not always feasible in the new working environment. They retain reservations about her late nights, but Ankita's economic empowerment reduces the probability of outspoken resentment by her in-laws. They also appreciate the fact that she sets aside her weekends for them.

However, given the multiple and often conflicting objectives of women workers, uncertainties involved in identifying actions associated with a particular outcome and differences in attitudes and goals of women workers, an optimizing framework does not appear appropriate for analyzing decisions of respondents. We suggest that Simon's satisficing model offers a more relevant alternative framework to examine their decision-making. This is illustrated with reference to three

[7] Commitment toward parents is quite high, with about 90 % respondents reporting that they spent time with their parents (in-laws) regularly or on their off-days.

instances often cited by researchers on IT sector—occupational segregation, undertaking household chores, and ensuring child care.

Occupational Segregation

Researchers have pointed out that the existence of gender stereotyping in the IT and ITES sectors has led to occupational segregation within the industry. The belief that women find it difficult to acquire the "hard skills" required in the hardware sector or display the aggressiveness and drive required from project consultants has led them to concentrate on software design. It has also been argued that difficulties faced by women employees in balancing work and home responsibilities often force them to opt for soft careers like quality assurance, testing, and human relations. These sectors have fixed hours; another advantage is that on-site assignments and traveling can be avoided. This is observed not only in other developing countries (Wajcman and Lobb 2007), but also in New Zealand (Crump et al. 2007) and Germany (Ben 2007).

As pointed out previously, multiple equilibria are consistent with satisficing, and each agent makes her own career decision on the basis of individual circumstances. Respondents echo Hakim (1991, 1995, 1996a, b) when they stress that occupational segregation is the outcome of deliberate occupational choices by women, based on their orientations:

> "Women have to balance—and for balance they have to compromise (on their career). And the only thing that they can compromise is their sector… Different women are molded differently. Some want to give more time to their work than to their marriage. Some women know they can balance their work and married life. It depends upon the women what they want to do."

Viewed from this perspective, occupational segregation may be interpreted as the outcome of a deliberate attempt to balance work-home commitments through choice of career.

Women who attach greater weight to family welfare opt for "soft" careers (in departments like Quality Assurance and Human Relations). As these sectors have fixed hours and do not require traveling and on-site assignments, they enable respondents to look after family needs. At the other extreme, we have women with their focus set firmly on their career. Such women generally choose not to marry, and are prepared to accept off-site assignments, travel frequently, work long hours, and so on. This group prefers demanding assignments like project consultancy. Most women, however, try to balance both work and home. They go only for short-term on-site and traveling assignments, delay having children to establish their reputation within the company to facilitate their comeback after a long layoff, and rely on external (often hired) support system for household responsibilities. Such women are also dynamically adaptive. Over time, as their children grow older and require less care, their priorities change and are revised in favors of the

office. Such women are willing to work in areas like software development, hardware research, and other less "soft" departments.

To some extent, this may also explain the glass ceiling restricting entry of women into higher echelons of the IT industry. Both the first and third group of women workers are hampered by their focus on family and children, so that it is only the second group, forming a minority, which has the ability to break the glass ceiling.

Provisioning of Care Services

We had earlier referred to the pressures faced by working women who had to assume the dual role of a home-maker and income-earner. This is a historical problem. The first women workers from middle income families had entered the labor market solely to supplement family income. They were mainly clustered in lowly paid clerical jobs or in schools—occupations with fixed working hours. Further, the image of women as wife and mother, created by a patriarchal society, was deeply ingrained in their psyche. Bounded by "socially imposed altruism",[8] these women workers therefore took upon the dual role of bread-earner, and provider of care services.

In the past two decades, as educated women became common, and the Indian economy became more integrated with the global economic system, women—particularly educated women from high middle income urban families—began to shift away from pink collar jobs (like teaching, secretarial assistance, receptionist, etc.) and competed with men for technical jobs. Employment is no longer linked to economic survival in such families; rather, it allows women to seek psychological satisfaction and sense of fulfillment outside the family. For instance, a software engineer said, "It (my job) is so much a part of my identity … I define myself to a very large extent on the basis of my work". Similarly, another respondent said, "I never had it in my mind that I will quit my job … Whatever I have learned, if I can apply it in my practical life, that's the best thing." Other respondents referred to the empowering effect of working.

The influences of globalization and western culture also left their imprint on the educated women workers in the Indian society. Not only did these give rise to gendered consumerism, but these also began to chip away at traditional patriarchal values. The earlier values based on the concept of personal care and service by household women were replaced by a more pragmatic, market-oriented mentality willing to substitute personal care by purchased services. This modified the traditional social altruism by still retaining women as persons responsible for the care and welfare of family members, but no longer making it obligatory on their part to fulfill this responsibility personally. The availability of cheap domestic labor

[8] "Social altruism" refers to the norm that assigns women greater responsibility for the care of dependents (Badgett et al. 2001).

permitted women[9] to shift away from supplying labor to a supervisory and planning role.

Simultaneously, there was a gradual change in the family structure. The joint structure have been distintegrating after the Second World War; where it did, the family became even smaller—consisting of just parents and children, with the grand-parents being gradually discarded and treated as part of an external support system.[10] The new structure and values, coupled with her new found confidence and sense of identity, allowed the wife to renegotiate her role within the family (Lee 2004) and modify practices still embedded within the family (Kelkar et al. 2002; Kelkar and Nathan 2002).

In some nuclear families, sharing of household tasks has become accepted; some husbands help in case of household emergencies.[11] This is enough for respondents, who uniformly say "I don't expect him to do (household chores) also", citing his lack of training as the reason—"you suddenly cannot change the demarcation inculcated from childhood." Similarly, a software engineer remarked, "If I don't feel like cooking, it's not that he will cook; he will ask me to eat outside." Gender stereotyping of household roles still persists. An accepted practice is that husbands are responsible for "outside the house" tasks related to financial matters—"He is otherwise extremely useful in the house—looking after banks and other financial matters".

Although women are still the home makers, the way of doing it has changed so that the modified Game of Chicken (Fig. 13.1b) no longer holds. Our survey revealed that very few women undertook household chores themselves (Table 13.4). Only one out of five respondents cook, wash, or dust on a regular basis. A significant proportion of such respondents are unmarried women, living as paying guests, or in rented accommodation. If only married women are considered, then also the proportion of women undertaking these tasks is about the same (20 %). This is contrary to what the Game of Chicken predicts in the presence of patriarchy, viz. that working women will perform household tasks.

Respondents reported trying to balance their commitments to office and home by hiring women from the newly developed "centers" which offer household and other services. Thus, the availability of domestic labor in Kolkata breaks down the

[9] The emergence of "Centers" has also facilitated this process. Such centers hire out domestic workers on a per hour basis. This permits working women to hire such help for 10–12 h to perform the needed tasks. Affiliation of the helpers to centers guarantees their reliability and honesty. Further, as a helper becomes 'fixed' to a household, she becomes acquainted with its patterns and requirements.

[10] For instance, this study found that in about 44 % households' cases there were no elderly persons within the family; only 14 % of the households had two elderly members. The family size is quite small (as shown in Table 13.2), with 74 % of the families having less than five members.

[11] Such emergencies arise when sudden visitors come, domestic helps' absentee themselves, and so on. In such situations, the introduction of electrical appliances facilitating household work becomes important in involving the husband in household chores. For instance, the husband can wash clothes when the maidservant has absented, or heat food simply by turning some knobs.

Table 13.4 Regularity in performing household tasks

Marital status	Household chore	Never	Occasionally	Always
Unmarried	Cooking	22.4	55.1	22.4
	Washing	20.0	54.0	26.0
	Dusting	18.4	59.2	22.4
Married	Cooking	9.1	55.1	22.7
	Washing	20.0	57.8	22.2
	Dusting	17.8	64.4	17.8

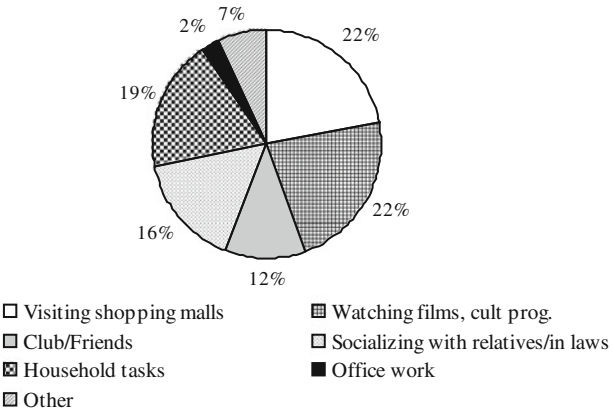


Fig. 13.3 Allocation of leisure hours by working women

analogy of the Chicken Game by providing wives with a third option. Instead of doing household tasks themselves, they can *purchase* care services and switch to a *supervisory* role.

Another example of satisficing is the way in which our sample spends their hours at home (Fig. 13.3). Only 20 % of respondents accorded first priority and 8 % accorded second priority to household duties in leisure allocation. On the other hand, activities in which the entire family takes part and enables the respondent to simultaneously undertake household tasks (shop for grocery items and other household requirements, have lunch/dinner)—like visiting shopping malls and watching films—are the preferred ways of spending leisure time.

Overall, respondents were satisfied with the balance between work and home. About 80 % of respondents (82 % of married workers) reported that they were content with the current state of affairs. Although 54 % did feel that long work hours interfered with their family life, they mostly cited the inability to devote time to their children, husband, and social interactions (with friends) as unsatisfactory.

Child Care and Pregnancy

One problem with such purchased services is that the care providers cannot fully meet the emotional needs of children (Ferguson 1989). A respondent remarked that whenever the time came for her domestic help to leave, she felt tense in case any unexpected development prevented her from leaving at her scheduled time. Respondents reported that men colleagues are generally insensitive to child-related problems:

> "There are people … when you call up to say that you will be working from home because your maid hasn't turned up and you have to look after your daughter … people don't say it but you feel it that men don't think … realize that this is an issue - but it matters a lot to you. … No one will say it why you are doing this … because that is going against the code of conduct and also that would take away the liberated image from them."

An ITES employee said that when she left at 6 p.m., her colleagues perceived her to be slacking. The need to look after children is another reason for the reluctance to travel or accept long-term on-site assignments observed earlier, though companies generally accommodate such cases.[12]

This implies that pregnancy has major implications for the working women. It has been observed by researchers (Mitter and Rowbotham 1995; COD 2004; Upadhya 2005) that pregnancy leads to a break in the career, and women are often unable to return to work thereafter. Respondents admitted that having children affected their careers as they are negatively evaluated during their leave period.

To minimize the effect on their career, women preferred to establish themselves in their companies before having a child, so that they could return after about 2 years of leave (including unpaid leave). The 25-year-old software engineer, Ankita, was aware that pregnancy and rearing up the child is demanding and will mark a difficult phase in her life. Although she wanted a child very much and wanted to give the child "the best of everything", she was equally determined not to sacrifice her career for her child. Other respondents spoke of the fulfillment in having and bringing up a child despite the immense pressure involved.

IT women therefore plan their pregnancy carefully. For instance, Ankita had made up her mind to delay her pregnancy for at least 2–3 years, and then take a leave of about 2 years. This interval would enable her to accumulate enough leave, save enough to compensate for the temporary loss in income during her leave period and facilitate rejoining her company after her maternal leave by creating a good reputation.

Once leave ends the situation becomes more complex. While some teams had a culture of working late hours—"When you rejoin, people expect you to swing back to work like they used to before … they expect these people to stay back till

[12] Male workers, on the other hand, are not handicapped in this way. Their responsibility in child care activities is limited to looking after the child when there is an emergency.

late at night"—others emphasized the output from the employee, allowing working mothers to leave at regular hours.

Working mothers generally rely on paid help, parents, or in-laws in such cases. In Kolkata, neighbors are quite supportive. If meeting project deadlines prevent Ranjana, an employee in a company developing e-learning modules, from returning home before her domestic help leaves, her daughter is taken care by her neighbors. She also keeps in touch with their child during office hours through the telephone. Respondents also sacrifice leisure hours to look after the child. Three-fourths of the women having children said that they always looked after their children themselves after returning home, while 11 % of such women kept aside their off-days for their children. We had also seen earlier that about 44 % of respondents spent their leisure hours in activities that may involve their children (visiting shopping malls, watching films, etc.). Respondents insist that this is deliberate—even those respondents whose households are run by their mother-in-laws find time to supervise their children's education and respond to their emotional needs. This is in keeping with the observation that women often incorporate the welfare of their children into their utility function (Sen 1993).

At the same time, respondents do not sacrifice their careers for their child. Ranjana had to leave her daughter for 5 months to participate in an induction program in Mumbai before joining her present company. She also mentioned that sometimes when her daughter rings her up she is busy. Although she realizes that the information that her daughter wants to give may be of significance to the child, in some cases she does not answer the phone. Emotional deprivation of the child is accepted as a cost of careerism.

In all, respondents avail of better reproductive techniques to phase conception at a stage of their career when the perceived cost of withdrawal from the labor force is least. Initially, they look after the child taking leave. Subsequently, they shift the responsibility of child care to relatives and hired helps, but sacrifice their leisure hours to spend time with their children. This again constitutes an example of how respondents adapt their aspirations to attain a satisfactory outcome.

Conclusion

To sum up, our findings reveal that women workers in Kolkata's IT sector belong to an exclusive socio-economic category—urban-based, middle income, English-speaking, and educated families. As a result their motivations for seeking employment, attitudes toward life and family, and actions are different from the low income women workers (Noronha and D'Cruz 2008). Predictably, this leads to differences in outcomes. The primary motivation underlying the decision of respondents to work is to create a personal space and to support their consumerist life styles. As a result, they are shedding their traditional caring and family oriented behavior, and relying more upon purchasing care services. Instead of personally undertaking household tasks, women are shifting to a planning and

supervisory role. Pregnancy, too, is a carefully planned decision. The lives of these women, therefore, represent a careful balancing of multiple, competing objectives. Starting from their choice of occupation, their job decisions, allocation of leisure time, pattern of spending after-office hours, the different choices of these women workers represent a dynamic adjustment of aspirations to structural realities, which may be conceptualized in terms of a satisficing framework.

But have these empowered women? Have their agencies really been enhanced, or is it a mirage sugar-coated by the gendered consumerism that they can afford?

Historically, there has been a division of labor between the affluent men and women, with the former performing the role of bread-earner and the latter the role of house maker and caregiver. Over time, however, social changes have led to a disintegration of this bi-polar outcome. The confines binding women to the home have broken down, offering them a choice between work and home. The burden emerging from the contradictory nature of work and home, however, rests solely on women. Whether this really represents a choice is debatable. Proctor and Padfield (1998) argue that for choice to be meaningful, work and family must be socially organized to permit either one or both to be experienced, as is the every day experience of men in society.

When we raised this issue before respondents, they admitted the validity of this proposition. They referred to the social limits on their aspirations and choices, the asymmetrical gender relations still persisting at home and work. Respondents complained about the long hours of monotonous work. They also suffer from fractured identities, as is evident from their guilt in not giving enough time to their family and particularly to their child. However, they also pointed out that advantage of the employment opportunities created by globalization the outcomes now are not hoisted on the respondents by a patriarchal family taking, but are deliberate choices made by the women: “Earlier it was not the girl’s choice. They had to do what their family allowed.” Although society may still circumscribe their choice set, the decision to choose their individual outcomes is that of the actors. This, we argue, is the most important effect of the IT revolution on women’s empowerment.

Acknowledgments The survey was part of the Rosa Luxemburg Foundation-Institute of Development Studies Kolkata project, ‘Women, Work and Education’. We are grateful to Achin Chakraborty (IDSK) and Uttam Bhattacharya (IDSK) for their suggestions.

References

Agarwal B (1994) A field of one’s own: gender and land rights in South Asia. Cambridge, New Delhi

Badgett M, Lee V, Folbre N (2001) Gender norms and economic outcomes. In: Loutfi MF (ed) Women, gender and work: what is equality, and how do we get there?. ILO, Geneva, pp 327–341

Ben ER (2007) Defining expertise in software development while doing gender. Gend Work Organ 14(4):312–332
Centre for Organization Development (2004) Final report on women in information technology. Report submitted to Department of Women & Child, Ministry of Human Resource Development, Government of India, Hyderabad
Clarke A, Sekher TV (2007) Can career-minded young women reverse gender discrimination? A view from Bangalore's high tech sector. Gend Technol Dev 11(3):285–319
Crump BJ, Logan KA, McIlroy A (2007) Does Gender Still Matter? A Study of the Views of Women in the ICT Industry in New Zealand. Gend Work Organ 14(4):349–370
Ferguson A (1989) Blood at the Root: Motherhood, Sexuality and Male Dominance. Pandora Press, London
Folbre N, Weisskopf TE (1998) Did father know best? Families, markets and the supply of caring labor. In: Ben-Ner A, Putterman L (eds) Economics, values and organization. New York, Cambridge, pp 171–205
Hakim C (1991) Grateful slaves and self made women: Fact and fantasy in women's work orientation. Eur Sociol Rev 7(2):101–121
Hakim C (1995) Five feminist myths about women's employment. Brit J Sociol 46(3):429–455
Hakim C (1996a) The sexual division of labor and women's heterogeneity. Brit J Sociol 47(1):178–188
Hakim C (1996b) Key Issues in Women's Work: Female Heterogeneity and the Polarization of Women's Employment. Athlone, London
Kelkar G, Nathan D (2002) Gender relations and technological change in Asia. Curr Sociol 50(3):427–441
Kelkar G, Shreshtha G, Veena N (2002) IT industry and women's agency. Gend Technol Dev 6(1):63–84
Krishna A, Brihmadesam V (2006) What does it take to be a software professional? Economic and Political Weekly. 19 July 2002, pp 3307–3314
Lee JC (2004) Access, self-image and empowerment: computer training for women entrepreneurs in Costa Rica. Gend Technol Dev 6(1):63–84
Mitter S, Rowbotham S (eds) (1995) Women encounter technology: changing patterns of employment in the Third World. Routledge, New York
NASSCOM and Deloitte (2008) Indian IT/ITES industry: impacting Indian economy and society 2007–2008. NASSCOM, New Delhi
Noronha E, D'Cruz P (2008) The dynamics of teleworking. Gend Technol Dev 12(2):157–183
Proctor I, Padefield M (1998) Young Adult Women, Work and Family: Living a contradiction. Mansell, London
Ramsey N, McCorduck P (2005) Where are the women in information technology? Report prepared by Anita Borg Institute for Women and Technology for National Center for & Information Technology. Un. of Colarado, Boulder
Rothboeck S, Vijaybhaskar M, Gayatri V (2001) Labor in the Indian economy: the case of the Indian software labor market. ILO, New Delhi
Sen AK (1993) Economics and the family. In: Uberoi P (ed) Family, kinship and marriage in India. Oxford University Press, New Delhi, pp 452–463
Shanker D (2008) Gender relations in IT companies. Gend Technol Dev 12(2):185–202
Simon HA (1955) A behavioural model of rational choice. Q J Econ LXIX(1):99–118
Upadhya C (2005) Gender issues in the Indian software outsourcing industry. In: Gurumurthy A, Singh PJ, Mundkur A, Swamy M (eds) Gender in the information society: emerging issues. UNDP-AIDP & Elsevier, New Delhi, pp 74–84
Upadhya C (2007) Employment, exclusion and merit in the Indian IT industry. Economic & Political Weekly. 19 May 2007, pp 1863–1868
Wajcman J, Lobb LAP (2007) The gender relations of software work in Vietnam. Gend Technol Dev 11(1):1–26

Chapter 14
Dualism in the Informal Economy: Exploring the Indian Informal Manufacturing Sector

Rajesh Bhattacharya, Snehashish Bhattacharya and Kalyan Kumar Sanyal

Introduction

Over the past four decades, the informal economy has become central to understanding poverty in developing countries and devising welfare policies. The informal economy has emerged as an important economic space for state interventions in poverty alleviation. This is reflected in the explosion of official statistics on the informal sector at the national level. Yet, "informal economy" remains one of the fuzziest concepts in circulation, yielding to every contextual demand on its meaning. As Maloney (2004, p. 1159) observes, "[t]hree decades of research have not yielded consensus either on the definition of the informal sector or its "*razon de ser*"."

We feel that to understand the current conjuncture of postcolonial development in countries like India, it is imperative to move beyond some of the dominant theories and the conventional debates on the informal economy, and cast a fresh look at conditions of existence, reproduction and survival of the informal sector. Moreover, to comprehend the current growth process in developing countries that have experienced rapid economic growth led by the formal sector but a simultaneous expansion of the informal sector, it has become critically important to analyze the relationship between the growth process in the formal sector and the process of reproduction of the informal sector in a new light.

R. Bhattacharya (✉) · S. Bhattacharya
South Asian University, New Delhi, India
e-mail: rajesh2466@gmail.com

S. Bhattacharya
e-mail: b.snehashish@gmail.com

K. K. Sanyal
University of Calcutta, Kolkata, India

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_14,

In this essay, we present a set of arguments that departs radically from prevalent theories of the informal sector, and try to empirically substantiate our claims by looking at informal manufacturing enterprises in India and their relationships with the formal sector. In particular, we show a marked dualism within the informal sector in terms of the dominant production relations in the informal enterprises. Given this dualism, the natures and effects of the relationships between various types of formal and informal enterprises are not homogeneous—they have differential impacts on the performance and growth of the informal enterprises. We then propose some broad contours of a fresh research program on the informal economy and its relation to contemporary postcolonial capitalist development based on the above perspective that would push and move beyond the limits of the dominant discourse on this subject.

In the following section ("Problematizing the Informal Economy"), we set the context of the chapter by briefly summarizing the key relevant issues in major debates on the informal economy, and then proceed to outline our core arguments. In the next section ("Heterogeneity in the Informal Manufacturing Sector in India: Rethinking Dualism"), we use this perspective to analyze the disaggregated National Sample Survey (NSSO) data on the Indian informal manufacturing sector. In the section "Beyond Old Debates", we point out some serious lacks in the existing scholarship on the informal economy which become more visible once we adopt the theoretical lens outlined below. If these lacks can be filled up through further research within this theoretical paradigm, it can throw open new avenues for analyzing the pattern and trajectory of capitalist development in the postcolonial economies. Finally, in the concluding section we summarize very briefly the major arguments of this chapter.

Problematizing the Informal Economy

Debates on informal economy continue to rage on two major issues. First, there are two competing views in understanding the informal sector—the continuist and the dualist views. According to the former perspective, the informal sector in developing economies is treated as a micro-enterprise sector which exhibits characteristics of self-employment in the developed countries—i.e., dynamic, creative, risk-taking entrepreneurial activities. Hernando De Soto (1989, 2000) popularized the view that informal entrepreneurs are as dynamic as any in the formal sector, but they are victims of a legal system that favor larger corporate firms in the formal sector. Fajnzylber et al. (2009, p. 414), based on empirical evidence on Mexico, argue that "microenterprises in Mexico show dynamic patterns consistent with the entrepreneurial risk-taking view usually applied to firm dynamics in the industrial world".

In the dualist view, the informal sector is understood as the parking lot for the poor—an economic space brought into existence by the workforce rejected or rendered surplus by the formal sector. Instead of resembling the modern industry, the informal sector resembles traditional agriculture as the domain of disguised

unemployment. Studies like Mandelman and Montes-Rojas (2009), on Argentinean labor market, suggest that self-employment is more often the refuge for urban unemployed and freshly arrived rural migrants than voluntary entrepreneurial activity. La Porta and Shleifer (2008) finds evidence for the dualistic view based on data on a large set of countries provided by World Bank Enterprise Surveys, Informal Surveys and Micro Surveys.

> From the perspective of economic growth, one should not expect much from the unofficial economy, with its millions of entrepreneurs, except to hope that it disappears over time (p. 347).

The second axis of debate is the nature of formal–informal linkage (Tokman 1978). There are two issues of importance here—(i) whether the informal sector is autonomous or integrated with the rest of the economy, and (ii) whether the linkage is beneficial or harmful for the informal sector.

According to one view (presented in ILO 1972; Oshima 1971; Sethuraman 1975; etc.), informal economy is an autonomous segment supplying goods and services to lower income groups of the urban population and thereby forming a self-reproducing closed totality. The autonomous informal sector produces with maximum input of labor, little capital (often second-hand equipment) and technology rendered obsolete in the formal sector. Although there is absence of linkage between the formal and the informal economy in production, consumption and technology, the relation between the two is considered a benign one. However, in the Latin American marginalist school (e.g., Quijano and Westwell 1983; Nun 2000), the informal sector or the "marginal pole", though autonomous, is subordinated to the formal sector or the "hegemonic pole", in two ways—in terms of a labor surplus that the informal sector has to carry and which is excluded by the formal sector, and also in terms of restricted access to input and output markets, which are monopolized by large firms in the formal sector. This implies that the informal sector is characterized by surplus labor subsisting on a residual resource base. According to another view, the informal sector is not autonomous, but linked to the formal sector and the linkage subjects informal firms to exploitation by formal sector firms. Informal firms are dependent on formal firms for inputs and markets in subcontracting relations and the surplus produced in the former is extracted by the latter through the pricing mechanism. Moreover, formal sector firms, in an attempt to cut labor costs, use home workers, sweatshops, etc., in the informal sector. These people appear to be self-employed, but in reality are "disguised workers" stripped of legal protection and benefits of formal employment (Portes and Walton 1981).

Finally, some authors (e.g., Ranis and Stewart 1999; Arimah 2001; ILO 1985; House 1984) consider the linkage beneficial, at least under certain conditions. Ranis and Stewart (1999) make a distinction between a dynamic and a stagnant component of the informal economy and argue that in the presence of a dynamic, competitive formal sector, with high linkage ratio with the informal sector, economic growth in the formal sector will lead to the expansion of the dynamic component of the informal sector displacing the traditional/stagnant component.

Thus, through increased subcontracting relations, formal sector growth has a favorable outcome for the informal sector. Challenging the story of exploitative relations between the formal and the informal sectors through the process of subcontracting, Maloney (2004) contends that only a minority of informal sector enterprises are subcontracted to the formal enterprises and, at least in the case of Mexico in the 1990s, the incidence of such subcontracting has declined over time. Further, the subcontracted informal firms generally have higher earnings (adjusting for human capital differences) than firms that are outside contractual relations. Hence, subcontracting relations might be beneficial to both the firms in contract as it allows them flexibility in labor allocation.

Globalization has added a new dimension to the debates on informal sector. Global dispersion and relocation of production, the rising intensity of subcontracting, casualization of labor, erosion of labor laws, and other social regulations have blurred the boundary between formal and informal sectors. In recent times, therefore, debates on informal sector have been increasingly related to the debates on globalization and structural adjustment. Some authors, however, have maintained a distinction between *informalization* within the capitalist circuits in the formal sector and the *informal economy* (Sanyal 2007; Sainz 1998). While the former represents a dispersion of the formal capitalist circuits, the latter constitutes an "outside" of the formal capitalist space.

In our previous work, one of the authors, Sanyal (2007), presented an understanding of the informal economy, reminding in some ways the Latin American marginalist view, but extending much beyond the marginalist formulation. Sanyal argued that the informal sector can be seen as a noncapitalist "outside" to the space of capital. This economic space cannot be considered as a remnant of *pre*capitalism that will vanish with development of the capitalist formal sector. Rather, the development—even expansion—of the noncapitalist "outside" can happen simultaneously with the development of the capitalist sector. The emphasis is more on the economic significance of that component of the informal sector that is outside the production space of capital, rather than on the linkages between the two. The classic paradigm of economic growth in developing countries was based on the presupposition that the modern (capitalist) economy would expand by breaking up traditional (precapitalist) economies, transferring both economic resources and laborers from the traditional to the modern economy. Yet, the experience of economic growth in developing countries shows that while the capitalist economy did expand by breaking up traditional economies, it did so by transferring resources, not laborers. As a result, a "surplus" labor force emerged in developing countries consisting of dispossessed producers whose traditional livelihoods were destroyed but who were not absorbed in the modern sector.

> Wherever the "surplus" labor force settled, the following economic characteristics emerged—a clear preponderance of self-employment largely assisted by family labor, the household as a major site of production—particularly in case of non-agricultural activities—and community or kinship networks involving trust and reciprocity in place of impersonal exchange relations (Sanyal and Bhattacharya 2009, p. 36).

Enterprises with these characteristics—which distinguish them from public and private corporate enterprises in the formal sector—constitute the informal economy. A part of the informal economy is integrated to global or domestic capital via processes of subcontracting and putting-out, but an even larger part of it constitutes a non-capitalist production space populated by the surplus labor force and engaged in entrepreneurial activity primarily to secure subsistence; it constitutes an "outside" to the space of capital where accumulation structures social outcomes. Sanyal (2007) referred to the space of capital as the "accumulation-economy" and its "outside" as the "need-economy".[1]

Bhattacharya (2010) presented a Marxian analysis of class qua surplus in the informal manufacturing sector in India. Using NSS data, he finds a high degree of class heterogeneity within the informal manufacturing sector. Thus both capitalist as well as noncapitalist (self-appropriative and feudal) production processes are prevalent in the informal sector, with the self-appropriative petty commodity production (often using family labor) being the predominant form. Moreover, the capitalist and noncapitalist spaces *within* the informal sector can be clearly demarcated in terms of differential abilities of enterprises in the two spaces to produce surplus and to retain a part of the surplus to accumulate. Contrary to some popular perceptions of the informal economy, it is argued that both capitalist and noncapitalist informal enterprises can and do produce surplus, out of which they make several payments like rent, interest, merchant fees as well as bribes, hafta, contributions to political parties, etc. However, the volume of surplus produced in the noncapitalist enterprises is substantially low, both in comparison to the surplus produced in the capitalist enterprises as well as in absolute terms. The ability to produce a relatively high volume of surplus allows the informal capitalist

[1] Bardhan (2009), in his critique of Sanyal and Bhattacharya (2009), has argued that exclusion may not be inherent in the process of capitalist development—he cites China as an example of a country that has possibly exhausted its reserve of surplus labor through a process of rapid capital accumulation. He argues that exclusion results from policies (not limited only to the oft-cited labor laws) that obstruct the pace and diffusion of capitalist accumulation. Within India, Bardhan argues, there are regional variations in the rate of capitalist accumulation and they can be attributed to the state-level policy differences. Noting the peculiarities of the Indian society, Bardhan concludes that "[t]he Indian transition is thus bound to be rather tortuous, though in the long run inexorable, and its narrative will be more complex than usual" (ibid: p.35). In the following section, on the basis of NSSO data on informal manufacturing sector in India, we argue that the informal sector, in terms of production relations, is largely constituted by a non-capitalist "outside" of capital. However, based on the limited empirical analysis in the space of this chapter, and the available data, we have no way to throw light on the processes (e.g. "exclusion" in Sanyal (2007) and Sanyal and Bhattacharya (2009)) behind persistence of such a non-capitalist "outside of capital". Ultimately we differ with Bardhan on the ontology of capital that grounds our respective positions. Bardhan argues that a full transition to capitalism is inevitable, if only delayed due to policy interventions; we argue that policy interventions are endogenous to the process of capitalist development, being products of capital-labor contradictions within the circuits of capital. Hence, even if we accept that policy interventions are behind the persistence of informality, it only reinforces our argument that the nature of capitalist development itself renders its fullness impossible.

enterprises to retain a part of the surplus for accumulation, leading to expanded reproduction. On the other hand, noncapitalist enterprises can generate some surplus only by pushing down the subsistence levels of the working owners and other family workers much below the customary standard of living for wage-workers in the formal sector in India. These enterprises can reproduce themselves at a given level of operation only because they can minimize much of the firm-level expenditures by using household (non-market) resources like family labor (thus minimizing wages), using the household space as the space of production (thus minimizing rent), using basic tools and machineries gathered from the household (thus minimizing interest) and avoiding input costs through illegal means (like, e.g., "hooking" electricity connections), and so on. Under these conditions, any net surplus that these households can retain are likely to be channelled into their consumption funds to satisfy their "needs" (or to attain some notional or customary standard of living) rather than to be used for accumulation.

In the next section, using data on informal manufacturing sector in India, we compare productivity and household income between firms within and outside sub-contracting relations and between capitalist and noncapitalist firms. We investigate the extent of *informalization of capital* (taking incidence of subcontracting as a proxy for that) as compared to the size of the *informal sector*. Further, we try to illustrate the size of the non-capitalist "outside" by disaggregating the informal sector in terms of its production relations, and then explore the difference in productivity and incomes between capitalist firms and non-capitalist firms; i.e., we try to empirically ascertain a dualism *within* the informal sector. Finally, we try to uncover the pattern of productivity and income differences between firms within and outside contract to see if evidence points to beneficial or harmful impact of subcontracting on informal firms.

Heterogeneity in the Informal Manufacturing Sector in India: Rethinking Dualism

Although more than 90 % of the Indian work force is in the informal sector, there is no time-series data encompassing the entire informal sector. Most researchers rely on quinquennial surveys on unorganized manufacturing sector conducted by NSSO. There has been only one enterprise survey on the informal sector (employing the International Conference of Labour Statisticians (ICLS) 1993 definition) in 1999–2000—conducted parallelly with the household survey for employment–unemployment surveys. Subsequently, two reports on informal sector and conditions of employment (using ICLS 2003 distinction between informal sector and informal employment) were brought out by NSSO in 2004–2005 and 2009–2010. However, none of these surveys repeated the enterprise survey of 1999–2000. More importantly, the quinquennial NSSO surveys on the unorganized sector leave out agriculture and services—where informality is more predominant

than in the manufacturing sector—with the sole exception of the NSSO survey on unorganized services sector in 2001–2002.

In this section, we explore the incidence of subcontracting in the informal manufacturing sector and compare productivity and income of capitalist and non-capitalist enterprises within and outside subcontracting relations. For that purpose, unit-level data from the NSSO survey on unorganized manufacture is useful. However, from the data, it is not possible to ascertain to what extent such subcontracting relations constitute formal–informal linkages. There are many possibilities—formal sector manufacturing firms may subcontract or put-out parts of their production process to informal manufacturing firms; formal as well as informal merchants may enter into contractual relations with informal manufacturers; or larger informal manufacturing enterprises may enter into subcontracting or putting-out relations with smaller informal manufacturing enterprises. The data only identifies contract relations with respect to master enterprises or contractors. The figures for subcontracting are, therefore, an upper bound for formal—informal linkage and no precise conclusions pertaining to formal—informal linkages can be derived from this type of unit-level secondary data. The National Commission for Enterprises in the Unorganized Sector (NCEUS) has recently provided a standardized definition of the informal sector in the Indian context: "The informal sector consists of all unincorporated private enterprises owned by individuals or households engaged in the sale and production of goods and services operated on a proprietary or partnership basis and with less than ten total workers" (NCEUS 2008, p. 41). Here, we adopt the NCEUS definition and use unit-level disaggregated data from the 62nd Round Survey of unorganized manufacturing enterprises in India carried out by NSSO in 2005–2006 on a sample of 82897 unorganized manufacturing enterprises. We have used a subset of 77365 enterprises from this data-set (about 93 % of the sample enterprises) that conform to the above definition of informal enterprises, as well as have adequate data on all the major characteristics of the enterprises. The analysis is carried out in terms of population estimates from this sample, with the total number of informal enterprises in India estimated to be 16,798,700.[2] In subsequent analyses, by "informal

[2] The survey included 2260 urban enterprises that were part of the "list frame" of the survey, comprising of "very big" non-ASI enterprises. These enterprises are included in the unorganized sector since they are not covered in the ASI, but their economic characteristics are very different from the other unorganized sector enterprises (e.g., the gross value of output for these enterprises in 2001 were more than six times the average value of output of the enterprises in the small-scale industry in the urban areas [NSSO 2007a, b: B-3]). These enterprises were not considered for the analytical purposes of this chapter. The remaining 80,637 enterprises (42,050 rural and 38,587 urban enterprises) from the "area frame" of the survey, covering 4,798 villages and 5,125 urban blocks from all the 29 states and 6 union territories, were considered for analysis. However, some of these enterprises do not conform to the definition of the informal sector (though they were part of the unorganized sector) and, hence, are not part of the present analysis. Of the 80,637 enterprises, it was found that 377 enterprises did not operate on a proprietary or partnership basis. Among the rest, 2,870 enterprises had 10 or more total worker, and an additional 25 enterprises did not have proper information on the total number of workers. These 3,272 enterprises were not

manufacturing", we refer to this particular subset, while referring to the full set as "unorganized manufacturing". Therefore, our figures might differ from other analyses which focus on the entire unorganized sector. Further, our figures may differ substantially from other analyses because here we use *median* rather than *mean* values of the variables, while the latter is used almost universally in other analyses. Given the heterogeneity of the informal sector, mean and median values for most variables may differ substantially, with extremely high standard deviations that are several times larger than the mean values. So, to get a proper measure of central tendency, we use the *median* values instead of mean values throughout the following analysis.

The heterogeneity within the informal sector is recognized in the official classification of unorganized manufacturing sector enterprises as OAMEs (Own-account manufacturing enterprises—without any hired worker), NDMEs (Non-directory manufacturing enterprises—employing less than six workers (household and hired workers taken together) with at least one hired worker) and DMEs (Directory Manufacturing Enterprises—employing 6–9 total workers, both hired and household, with at least one hired worker). In more theoretical terms, we characterize DMEs as predominantly capitalist and OAMEs as non-capitalist (self-exploitative) enterprises, with the NDMEs exhibiting characteristics of both in substantial degrees.[3] In the theoretical literature on the informal sector (e.g. Ray

(Footnote 2 continued)
considered for this analysis. Thus the present analysis is based on the remaining 77,365 sample enterprises, a subset of the unorganized manufacturing sector surveyed in the 62nd round (about 93 % of the original sample), representing an estimated 16,798,700 enterprises (about 98 % percent of the estimated population of unorganized enterprises) that conform to the international as well as the Indian definition of an informal enterprise. These informal sector enterprises engage an estimated 32,331,802 workers (i.e., about 88.7 % of the estimated 36,442,799 unorganized manufacturing sector workers; and about 71 % of all workers who are engaged in manufacturing activities, organized and unorganized). The population estimates from the sample for all analyses in this chapter are calculated by using the weights provided by the NSSO along with the unit-level data.

[3] Resnick and Wolff (1987) characterize a class structure in terms of fundamental and subsumed class processes. The former refers to the production and appropriation of surplus labor produced by direct producers; the latter refers to the processes by which the appropriators of surplus labor distribute it to those who provide conditions of existence of the fundamental class process. The identification of a class structure as capitalist or non-capitalist (self-appropriative, feudal, slave, communist, etc.) depends on specifying the mode of appropriation of surplus, the particular political, cultural, economic, and natural conditions of existence of that mode and hence the subsumed classes responsible for providing those conditions. See Bhattacharya (2010) for the details and the implications of such classifications for the informal economy. In Sanyal (2007), the formal and informal economies are not class-specified, but are rather conceptualized around the logic of production—production for consumption/need in the "need-economy" and production for accumulation in the "accumulation-economy". The "need-economy" and "accumulation-economy" can each accommodate both capitalist and non-capitalist enterprises. While both the above approaches are important in the context of understanding the informal economy, our purpose in this essay is limited. Given that the NSSO data does not allow us to precisely identify the mode of appropriation of surplus and its conditions of existence, nor does it

1998; Ranis and Stewart 1999), we have seen a distinction between a dynamic informal sub-sector which can be understood as competitive small-scale capitalist sector and which is integrated with the formal sector, and a stagnant informal subsector that survives on the basis of income-sharing, a characteristic the latter shares with subsistence agriculture. Therefore, a part of the informal subsector is more like industrial capitalism, while another part resembles more traditional noncapitalist peasant agriculture. Our analysis attempts (to the extent we can do so based on the NSSO data) to highlight this dualism *within* the informal economy, even at the smallest scale. In the informal manufacturing sector, OAMEs, NDMEs, and DMEs accounted for 86.9, 10.5 and 2.7 % of the total enterprises in 2005–2006. For DMEs, mean numbers of full-time hired and family workers are 5.1 and 1.7, respectively, whereas for NDMEs, the corresponding figures are 1.7 and 1.4, respectively.[4] The informal sector is clearly dominated by family enterprises, with wage-relations largely absent in the overwhelming majority of firms constituting a large non-capitalist space within the Indian economy.

In the rest of this section, we limit our focus to subcontracting relations prevalent in the informal sector, which enables us to emphasize the heterogeneity within the sub-contracted informal manufacturing firms in terms of productivities and incomes. In 2005–2006, of all informal enterprises, 31.2 % were within contractual relations, to various degrees, with master enterprise/contractor. Within OAMEs, NDMEs, and DMEs, the figures are 32.5, 25.5 and 33.6 %, respectively. For the entire unorganized manufacturing sector, incidence of subcontracting has slightly fallen between 2000–2001 and 2005–2006 for NDMEs (from 28.9 to 25.4 %) and DMEs (from 34.7 to 30.4 %), and marginally risen for OAMEs (from 30.7 to 32.5 %) (Sahu 2010). Therefore, the incidence of subcontracting is limited to less than one-third of informal/unorganized enterprises in the aggregate, as well as for noncapitalist and capitalist enterprises within the informal sector. Since these figures on incidence of subcontracting provide the upper bound to the extent of formal–informal linkage, it is evident that subcontracting by the formal sector cannot provide the *raison d'être* of the informal economy. Among DMEs, subcontracting is concentrated among smaller sized firms—30.4 % of "unorganized"

(Footnote 3 continued)
allow us to identify the objective of production (need or accumulation), we treat enterprises with hired labor as capitalist enterprises and enterprises run on own and family labor as non-capitalist (self-appropriative) enterprises. DMEs and NDMEs have at least one hired laborer combined often with own and family labor, while OAMES are run soley on own and family labor. We understand DMEs and OAMEs as predominantly capitalist and non-capitalist (self-appropriative) in nature respectively, while NDMEs partake the character of both to substantial degrees. Of course, in identifying this way, we ignore other non-capitalist class processes involving family labor (e.g. feudal) in all of these enterprises. However, given the nature of the data such finer theoretical distinctions are hardly empirically tractable, though their theoretical significance needs to be acknowledged. In our analysis, predominance of a particular class process in an enterprise simply means the quantitative predominance of hired labor versus own plus family labor, as calculated from unit-level data.

[4] Here we consider the positions of two part-time workers as one full-time position.

Table 14.1 Distribution of enterprises within contract by contract type

Contract type	OAME	NDME	DME
Solely for the master enterprise/contractor	88	63	69
Mainly for the master enterprise/contractor but also for other customers	6	19	24
But mainly for customers	6	18	7
Total	100	100	100

DMEs (which includes larger sized DMEs) as compared to 33.6 % of "informal" DMEs (DMEs with up to ten total workers including family labor) in 2005–2006 were operating on contract.

From the following (Table 14.1), we get a more disaggregated picture of contractual relations in the informal economy. It is seen that 88 % of subcontracted non-capitalist enterprises work *solely* on contract, and almost two-third of subcontracted capitalist enterprises do so. Thus, vast majority of subcontracted informal manufacturing enterprises works solely for master enterprise/contractor, signifying a high level of dependence on the subcontractor, but such dependence is more pronounced in case of OAMEs.

The extent of dependence of the subcontracted firms on the master enterprises/contractors will have to be ascertained from other components of the survey. It is found that almost 87 % of all subcontracted enterprises receive raw materials from master enterprise/contractor (89 % of OAMEs, 70 % of NDMEs, and 74 % of DMEs within contract). Also, more than 80 % of all subcontracted firms use their own equipment (81 % of OAMEs, 90 % of NDMEs, and 88 % of DMEs within contract). Third, almost 95 % of the subcontracted enterprises (with almost no variation among OAMEs, NDMEs, and DMEs) have designs specified by the master enterprise/contractor. From above, it can be inferred that most of the subcontracted enterprises work with their own tools and machines (hired tools and equipments constitute an insignificant component of the total for all types of enterprises) and they are primarily dependent on the master enterprises/contractors for their raw materials, for access to market and for assessing market demand through design specifications. So, subcontracted firms are dependent on the master enterprise/contractor for both conception and execution of work. The subcontracted firms are "independent" when it comes to technology (as they own most of the equipments); however, given the low level of technology for informal firms, this independence may signal absence of technology transfer from formal to informal manufacturing enterprises. This casts doubt on any beneficial linkage operating through technological dynamism of the informal enterprises.

Further, it is found that that among OAMEs, 93 % of the subcontracted firms organize production within household premises, while 75 % of non-subcontracted firms do so. The corresponding figures for NDMEs are 42 and 23 %, and for DMEs are 27 and 15 %. While it is intuitively obvious why among larger sized firms a higher proportion of firms organize their production outside their household premises, it is interesting that the proportion is lower for firms within contract. Thus, subcontracting mostly involves outsourcing to home-workers, who have

been traditionally referred to as disguised workers for capitalist enterprises. This point to the phenomenon of *informalization* of capitalist production relations. In the previous section, we have distinguished between the informal sector and the informalization of the formal sector. Empirical evidence, presented above, justifies such a theoretical distinction; it points to the (i) existence of a large noncapitalist economy outside subcontracting and (ii) the predominance of home-based production among subcontracted firms.

If we look at productivity and income by enterprise type within and outside subcontracting relationships, we do not find any clear evidence in support of the view that formal-informal linkages benefit the informal enterprises (of course, with the caveat, as noted before, that incidence of contract is not limited to formal–informal linkage). What data shows is that firms within contract are in general worse-off in terms of productivity and incomes than firms outside contract. It could be that worse-off firms tend to enter into contract more than relatively better-off firms or that contractual relations negatively affect productivity and income for the subcontracted firms. In the absence of panel data, the causality cannot be ascertained.

Before we discuss productivity figures, we note a curious pattern from Table 14.2—the standard deviation is markedly higher for firms outside contract than for firms within contract, and it increases substantially with firm size. When we truncate each group by 5 % at both ends of its distribution of GVA (Gross Value Added)/worker, the general picture remains the same, i.e., the standard deviation increases with firm size; however, the difference in standard deviations between subcontracted and non-subcontracted firms become comparable (in particular, for NDMEs and DMEs, the gap in standard deviation between subcontracted and non-subcontracted firms shrinks drastically from full to truncated subsamples). The table shows that dispersion in productivity, and hence heterogeneity within firms, is more pronounced in case of firms outside contract than for firms within contract. That is, subcontracted firms are much less differentiated than non-subcontracted firms in general, and within DMEs in particular. If we think that the better-off firms are more likely to graduate from the informal to the formal sector, then more such firms are outside contract than within contract. This appears to contradict the view that formal–informal linkage opens up doors to the formal sector for more productive informal firms.

Table 14.2 Annual GVA/worker (in Rs.) by enterprise type and by incidence of contract

Annual GVA/worker	OAME		NDME		DME	
	Contract	No contract	Contract	No contract	Contract	No contract
Full groups						
Median	6,570	9,240	25,290	28,470	32,213	31,560
Mean	9,525	14,070	29,639	36,775	39,837	45,616
SD	14,862	17,538	24,255	63,913	41,506	2,74,352
Truncated groups						
Median	6,570	9,240	25,350	28,470	32,213	31,937
Mean	8,181	12,032	27,775	30,926	36,031	34,903
SD	5,640	9,665	14,235	16,773	20,903	22,905

Table 14.3 Median Total assets (in Rs.) (owned and hired) per enterprise by enterprise type and incidence of contract (mean values in parentheses)

	OAME			NDME			DME		
	Contract	No contract	Total	Contract	No contract	Total	Contract	No contract	Total
Total fixed assets	7,500 (20,774)	15,030 (38,371)	11,610 (32,647)	1,35,600 (2,26,549)	1,30,000 (2,55,088)	1,31,100 (2,47,824)	3,16,500 (5,54,932)	3,10,400 (5,96,603)	3,15,000 (5,82,595)

In subsequent analyses, we use the non-truncated data. From Table 14.2, we see that the median values for GVA/worker are substantially higher for the capitalist enterprises (DMEs and NDMEs) compared to noncapitalist enterprises (OAMEs). For OAMEs and NDMEs, the median values of GVA/worker are higher for firms outside contract compared to those within contract (41 and 12.6 % higher, respectively), while for DMEs, it is 2 % lower for firms outside contract compared to those within contract. Thus, the difference in productivity of firms within and outside contract decreases with firm size. For OAMEs, firms outside contract are substantially better-off in terms of productivity than those within. It can be that the less productive non-capitalist firms are forced to enter into relationships of subordination via contract to reproduce their conditions of existence (access to raw materials, access to markets, etc.).

This argument can be further strengthened by looking at the difference in total assets (both owned and hired) of firms within and outside contracts (see Table 14.3). For OAMEs outside contract, the market value of total assets per enterprise is double that of OAMEs within contract. Hence, only asset-poor OAMES tend to enter contract relations, while asset-rich OAMES tend to operate independently. No such relation is found for NDMEs and DMEs, but expectedly, these enterprises are substantially more asset-rich than the OAMEs in general.

Table 14.4 gives the annual household income per participating family worker, including the working owner. We first calculate net surplus per enterprise (net surplus = GVA minus rent, interest and wages) and then divide it by the number of family workers (full-time equivalents) active in the enterprise. Since we have not imputed wages to family labor, interest to productive assets owned by the family or rent to household premises where it is used for production, this figure does not measure economic profit. However, in the absence of market opportunities for employment of these family owned resources elsewhere, this is a close approximation of the household income gain from the enterprise.

From Table 14.4, we can see that OAMEs outside contract are better off than OAMES within contract. For NDMEs, too, we get the same picture, though the difference is smaller, while there is almost no difference between DMEs within and outside contract. Interestingly, the difference in productivity between NDMEs and DMEs is much less pronounced compared to the difference in income. Assuming that family labor is equally productive as hired labor, the explanation must lie in the difference between productivity and wages of hired workers in DMEs, and the larger size of hired workforce in DMEs compared to NDMEs.

Table 14.4 Annual household income (in Rs.) per family worker by enterprise types

	OAME		NDME		DME	
	Contract	No contract	Contract	No contract	Contract	No contract
Annual household income per participating family worker (median)	6,520	9,024	33,480	37,500	59,184	59,310

Table 14.5 Growth status for by enterprise type and by incidence of contract for enterprises at least three years in existence (in per cent)

Growth status	OAME			NDME			DME			All enterprises
	Contract	No contract	Total	Contract	No contract	Total	Contract	No contract	Total	
Expanding	24.8	18.1	20.2	24.2	29.3	28.0	37.6	36.3	36.7	21.4
Stagnant	58.6	64.3	62.5	54.7	52.1	52.7	48.0	49.3	48.8	61.2
Contracting	16.6	17.6	17.3	21.2	18.6	19.2	14.4	14.5	14.4	17.4

As we move from NDME to DME, we see surplus emerging out of the difference between average productivity and wages of hired workers; thus the income of the household increases partly due to rise in productivity, but mostly due to profit made from hired workers.[5]

OAMEs within contract may experience higher growth compared to those outside contract, because firms under contract are less productive firms, for whom contract ensures a more secure supply of raw materials and access to markets, enabling them to generate more income for the household. This appears to be borne out by the reported growth perception of OAMEs that have been in existence for at least 3 years (see Table 14.5). For subcontracted OAMEs, approximately 25 % report as expanding, 59 % report as stagnant, and 17 % as contracting over the previous 3 years before the survey. The corresponding figures for OAMEs outside contract are 18, 64, and 18 %, respectively. For NDMEs, a similar pattern is observed, while for DMEs the difference in growth status for firms within and outside contract is insubstantial. In aggregate, for all enterprises, approximately 21 % report as expanding, 61 % as stagnant, and 17 % as contracting, i.e., for the informal manufacturing sector as a whole, it appears that 79 % of the enterprises did not accumulate and expand over the 3 years before the survey. As may be expected, proportion of expanding firms increases with firm-size (20 % of OAMEs, 28 % of NDMEs, and 37 % for DMEs).

The ability of an enterprise to accumulate and grow depends to a large extent on the surplus generated within it, especially for informal enterprises that generally do not have much access to formal credit. From Table 14.4, it can be calculated that the median subcontracted OAME generates Rs. 543 as monthly income for each family worker, and the median non-subcontracted OAME generates Rs. 752 per family worker per month. These are almost close to poverty line incomes, indicating a large proportion of OAMEs generate monthly income per household worker that is below or close poverty line income. The corresponding figures for the NDMEs are Rs. 2,790 and Rs. 3,125, respectively. While these figures are substantially higher than those for OAMEs, they are still hardly large enough to enable accumulation, after meeting consumption needs of the household. For the median DME, the monthly income per household worker is little less than

[5] In Marxian terms, volume of surplus increases with the number of exploitable wage workers.

Rs. 5,000 for firms both within and outside contract. This is a substantial earning in the context of the Indian informal economy, and therefore indicates possibility of internal accumulation. From the above figures, it can be understood why the informal economy is characterized as largely a "need-economy". It can even be argued that many of these informal enterprises are economically able to reproduce themselves at a given scale of operation—or even expand—by suppressing household consumption below the customary standard of living.

Next we try to see whether and how the general broad picture varies between high-income a low-income states. It might be argued that the incidence of low productive OAMEs should be less in high income states than in low income states, and linkages between informal and formal enterprises should be more prevalent in high income states. We take 10 among the top 12 states in terms of share of informal enterprises, and divide them into two categories in terms of their 2005–2006 per capita net state domestic product (NSDP)—high income states whose per capita NSDPs are above Rs. 30,000 (Maharashtra, Gujarat, Kerala, Tamil Nadu, Karnataka) and low income states whose per capita NSDPs are around or below Rs. 20,000 (Rajasthan, Orissa, Madhya Pradesh, Uttar Pradesh, and Bihar). We find that 82 % of informal enterprises in high income states are OAMEs and 91 % of informal enterprises in low income states are OAMEs. We also find that higher percentages of OAMEs and NDMEs are under subcontracting in high income states than those in low income states, whereas percentage of DMEs under subcontracting is lower in high income states than in low income states (see Table 14.6). From this, it can be seen that, first, the predominance of OAMEs in the informal sector is not just a typical feature of poorer states. Second, compared to low income states, smaller sized firms—especially the OAMES—are more likely to enter into subcontracting relations than larger firms in high income states.

Further, Table 14.7 shows that OAMEs and NDMEs are less productive within subcontracting relations compared to those outside such relations in both high income and low income states. The difference is more substantial in case of high income states. In case of DMEs, the pattern varies between high income and low

Table 14.6 Percentage of enterprises within contract by enterprise type

	High income states	Low income states
OAME	43.3	23
NDME	25.8	21.6
DME	29.5	35.3

Table 14.7 Annual GVA/worker by enterprise type in high and low income states

		High-income states		Low-income states	
		Contract	No contract	Contract	No contract
Annual GVA/worker	OAME	7,104	12,252	6,072	8,564
	NDME	23,221	30,510	21,720	23,225
	DME	38,910	23,603	24,347	28,429

income states. DMEs within contract are more productive than those outside contracts in high income states, but the opposite is true for low income states. Taking all these information together, it may be inferred that although the total incidence of subcontracting is higher in high income states compared to that in low income states, the phenomenon of subcontracting is more concentrated in less productive firms among the former. This shows that higher per capita income is not necessarily associated with more dynamic linkages between informal and formal sector enterprises. This lends further support to our previous arguments that there is no picture in the aggregate to suggest that linkages with the formal sector necessarily have beneficial effects on the productivity of informal firms. What data shows is that more productive firms are more likely to be outside contract than less productive firms in high income states.

In this section, we have used NSSO unit-level data to argue that certain major axes of debate in the informal sector need to be revisited. In particular, we have tried to show that:

- There is considerable heterogeneity within the informal sector. The heterogeneity is best captured in terms of production relations, making a distinction between non-capitalist family enterprises and capitalist enterprises. This enables us to point to a dualism within the informal economy.
- We have argued that subcontracting is limited to less than one-third of informal enterprises. Subcontracting relations mostly involve home-based production, subordinated to the master enterprise/contractor. This illustrates the process of informalization of capitalist production. However, informal economy cannot be reduced to informalization, since two-thirds of the enterprises are outside subcontracting relations.
- We have shown that subcontracting is more concentrated among less productive firms and this suggests that more productive firms may not perceive subcontracting relations to be beneficial for productivity growth.

In the next section, we discuss certain agenda for future research.

Beyond Old Debates

The sheer size of the informal economy in developing economies continues to pose theoretical challenges for policy-makers as well as for economists working on these countries. Such dogged persistence, and even expansion, of the informal economy in growing economies like India, calls for policy responses that have to balance the objective of maximizing growth, i.e., accumulation, within the capitalist sector, against the objective of ensuring subsistence for the overwhelming majority of the workforce in the noncapitalist informal sector. This necessitates a dualism in the policy stance. For example, in India, labor markets are deregulated to promote growth on one hand, while, on the other hand, policies like NREGA effectively act as minimum wage legislation in rural areas. This dualism in policy, however, hardly

finds its counterpart in theoretical development economics. Although the voluminous literature on informal sector is animated by debates on dualism, the theoretical literature continues to rely majorly on the assumption of oneness of the economy. We argue that research on developing countries like India should incorporate the twin economic problems of accumulation and subsistence in modeling the developing economy. In this section, however, we limit ourselves to offering what we perceive as some of the blind spots in research on informal sector.

From Informal Production to Informal Organization of Production

Research on informal economy focuses majorly on enterprise level characteristics, without any attention to the social organization of informal production. This is partly due to concerns over poverty and vulnerability of the workforce within the informal sector, which lead researchers to focus more on outcomes (income, asset, productivity, etc.) than organizational structures that support such outcomes. The informal sector is generally understood in terms of its lack of "organization", i.e., absence of institutions that provide conditions of existence of informal production space—e.g., lack of formal credit, lack of formal schooling, lack of social security, etc. However, the sheer size and persistence of the informal sector calls for more research on the social reproduction of the informal economy, which has to come from a perspective that sees the informal economy as having an organization different from that of formal sector, rather than as having no organization (i.e., being unorganized). What is required is a study of the institutional base of the informal economy, i.e., locally specific institutions of regulations that structure the informal economy. For example, what is conspicuously absent is an investigation into modes of regulation of local entry into specific branches of informal activity. It is often assumed that the informal economy is fiercely competitive. We need to investigate the logic of such competition. In conditions of surplus labor, there is overcrowding in all branches of informal activity. Therefore, for households using their own resources like land, labor and capital in any branch of activity, there is limited scope for more profitable employment of these resources in other branches. Further, for much of the informal economy, there is hardly any distinction between consumption and investment goods. The residence (i.e., land and building) is also used as the workshop; bicycles, for example, are as much a durable consumption good of the household as a fixed asset of the enterprise transport. In such a context, non-labor resources of informal production cannot be employed separately from household labor resources. Therefore, the former cannot be withdrawn from overcrowded branches of activity to less crowded ones in search for higher returns. In such a condition, the only logic of competition could be return to labor—i.e., subsistence. The study of informal economy should focus on how competition is regulated, which may involve erection of entry barriers so as to ensure conditions

for social reproduction of informal households at the subsistence level. If households earn more than subsistence, competitive pressure forces household income to subsistence level. Such regulation of competition is often politically imposed, through active intervention of political parties and explicit local trade associations, which enforce the regulation on informal enterprises. Inclusion and exclusion of firms in the informal economy is locally regulated and not governed by wider goods market conditions.

In short, conventional models of market competition are not adequate for the study of the informal sector. Research should also focus on non-market regulation of competitions.

The Role of "Commons"

The provision of basic subsistence for the households within the informal economy is critically predicated on the availability of adequate resources that enable the households to enter the space of commodity production. However, given the low levels of initial endowments and lack of funds for procuring resources through market, they are forced to innovate ways of accessing these resources through non-market relations outside the commodity space. Official data show that the median informal firm incurs negligible expenses on many necessary inputs and productive assets. The absence of such explicit costs enables the firms to retain most of value-added as household income to meet subsistence needs of the household. Thus, one important mechanism by which many informal enterprises and the households that depend on those enterprises manage to survive is by avoiding explicit costs of production. Further research is needed to explore how such explicit costs are avoided—through encroachment, illegal sharing of privately supplied services to the formal sector, sharing of common assets (like land as in the slums) within the informal economy, interest-free borrowing of assets from coproducers in the informal economy, etc. In the absence of such information, official data presents an almost improbable picture of informal production.

A fundamental condition of sustenance of the surplus population is their ability to create "commons" by undermining the regime of existing private property rights. It is well known that informality overlaps illegality to a large extent. This overlap has been generally a concern from the perspective of governance. What is less understood is the functional role of illegality in creation of commons. A dominant view of development holds that private property rights expand with development. The informal economy provides, on the contrary, an example of how commons may be an integral part of the development trajectory. The act of illegality appears to be appropriative from one angle (encroachment of private property) and productive from the other (constitution of commons). There is a vibrant tradition of research on local and informal governance of commons (see, e.g., Ostrom 1990; Ostrom et al. 1999). It would be useful to harness that tradition to research on informal sector.

The Household as a Site of Production

The informal economy is now understood as including those enterprises for which household and firm enterprises cannot be separated. The processes of consumption and production are intertwined, not only because both occur at the same site (within households), but also because there is no clear demarcation between time devoted to production and consumption within a normal working day. However, available surveys on the informal sector focus on enterprise characteristics, isolating them from the sphere of consumption of household. We have earlier noted production is oriented towards provision of subsistence to households that run the enterprises. The standard surveys, by separating household consumption surveys from informal enterprise surveys, miss the close link between consumption and production. This calls for a unified household-enterprise survey framework to study the informal economy. In individual research work, such surveys are easier to design and may prove more useful than existing survey frameworks. Another interesting survey methodology, widely used in the OECD countries, particularly for studying homework and carework by women in households, is time-use survey. Time-use surveys can be creatively used for the study of home-based informal production (see, e.g., Hirway 1999).

Production and Use of Surplus

While it is generally understood (and as is evident from the foregoing analysis of the informal economy in India) that much of informal economy struggles to generate income at the subsistence level, the notion of subsistence itself is not problematized in the existing literature. Since most researchers look at informal economy from the poverty perspective, subsistence is often equated with poverty or deprivation. However, subsistence has an alternative meaning too—one that can be discerned in the tradition of classical political economy—in which subsistence is understood as customary standard of living. In Marx, the value of labor power, i.e., value of subsistence goods of the worker, is determined by the social context, including "a historical and moral element". Therefore, it is possible to think of subsistence as the full basket of commodities deemed socially necessary for reproduction of life in a given historical context, and not just a basket of goods that secure biologically determined minimum calorie intake or some common minimum basic needs (Bhattacharya and Sanyal 2011). Within the informal economy there is considerable heterogeneity in household incomes generated per enterprise. Similarly, there is considerable difference in household income between informal and formal sector manufacturing workers. In such a context, the current standard of living of majority of informal workers may well be below their customary standard of living.

Once the notion of subsistence is problematized, the meaning of surplus and its accumulation also gets transformed. What appears as surplus may in fact be deferred consumption for those households whose current standard of living is below the customary standard of living. These households may save and invest a part of the current income to accumulate and grow; however, this accumulation may be motivated by the desire of these households to catch up with the customary standard of living. Moreover, the notion of subsistence, overdetermined as it is by the entire social context, is a dynamic concept. As the social context changes (in its political, economic, cultural and natural dimensions), the meaning of life, along with what is necessary for reproducing life (in particular, labor power) changes. The concept of subsistence too changes along with it. For example, the evolution of the capitalist economy may change the notion of subsistence not only for its own workers, but also for the self-employed informal producer outside it.[6] In this sense, "accumulation" may itself redefine "need" as one of its social consequences.

In that case, accumulation in informal enterprises is more about concrete historically determined consumption needs of the household than market-determined competitive rates of return to capital. If this is true, then accumulation in informal enterprises has a meaning different from that of accumulation in formal enterprises. While the NSSO surveys have questions on the growth perceptions of enterprise owners and provide data on net additions to fixed assets, it is impossible to determine (i) how is such enterprise growth motivated–competitive pressure to accumulate in order to maintain rates of return or the desire to close the gap between current and customary standards of living; and (ii) how much of the net addition to assets constitutes household investment and how much of it constitutes business fixed investment.

Future research should focus on how surplus is used by informal enterprises—whether for additional consumption or investment in enterprises. Further, household-enterprise combined surveys should also investigate the perceived difference in current and customary standard of living and the motive for accumulation in enterprises—whether it is competitive pressure or household's consumption needs.

[6] Think of health and education, both of which are normally accepted necessities for reproduction of labor power in the formal sector. Once socially established as a norm, it enters the subsistence basket of the informal producers too, although the ability to access such basic services in a privatized healthcare and education system is markedly different for the average formal and informal sector worker. In this case, the evolution of the capitalist economy—with its requirement of more productive, educated, and healthy labor force—influences the notion of subsistence in the noncapitalist economy. Just by looking at rising income and productivity in the noncapitalist informal economy, one cannot infer that accumulation is taking place, without making a prior assumption regarding the level of subsistence.

Sources of Informal Entrepreneurs

There is little work on enterprise migration from informal to formal sector. Of course, the possibility applies to more dynamic and more productive informal enterprises. One line of research could be to explore industry-wide migration of enterprises from informal to formal sector. Enterprise surveys on formal sector could be used to that purpose. In the legalist approach to the informal economy, it is argued that the legal structure creates incentive for firms to avoid formal regulations. Thus, the costs of staying formal expand the informal sector. However, it is well known that the informal entrepreneurs face a different set of costs in running their enterprise—costs that include paying hafta to local mafia, bribes to police, monetary and non-monetary contributions to political parties, etc. So, there are costs of staying informal. Moreover, there are limits to the possibilities of accumulation and growth within the informal sector, given credit constraints and constraints on number of workers that can be employed without formal registration under the Factory Act. One angle of research could be to ascertain to what extent the accumulating firm weighs the costs of informality against costs of formality in making decisions to stay informal or migrate to the formal sector.

On the other hand, it would also be worthwhile to explore to what extent formal sector "feeds" the informal sector—i.e., to what extent competitive pressure on small firms leads to a process of differentiation such that those losing out in the game migrate from formal to informal sector. On the other hand, formal sector workers losing their employment might be forced to set up informal enterprises to secure their livelihoods. These sources of informal entrepreneurs should be identified along with the more traditional source—migration out of agriculture.

Commodity-Chains in the Informal Economy

With globalization, there has been a marked change in the distribution of income and the social patterns of consumption. With increasing inequality of income and an aspirational consumption-driven "middle class", it is possible that the consumption space of the poor and the rich may drift apart. This raises the interesting question–do poor people produce for poor people? This issue calls for research to ascertain to what extent the informal manufacturing sector produces for the informal workforce. One way to explore the issue is to trace the commodity chains in the informal economy–from resources to production to consumption, i.e., to investigate to what extent the informal manufacturing sector "manages by itself"—in terms of resources they need for production and in terms of market for their produced output. In the early literature on urban informal economy—a distinction was drawn between an "upper circuit" and a "lower circuit" of the urban economy in third world cities (McGee 1973). Further research is needed to understand to what extent such circuits are closed circuits. It is difficult to do this

from official surveys. But case studies and field surveys by researchers can provide invaluable insights on this question.

Capitalocentrism and the Discourse on Informality

The discourse on informal sector is firmly grounded in a capitalocentric vision, which sees the informal economy as either following or subsumed under the logic of capital. While the mainstream discourse argues that an informal enterprise, faced with a given set of constraints, takes its decisions regarding the production process in the same way as any formal capitalist enterprise, the radical discourse generally posits the informal economy as functional to the needs of capital accumulation in the formal (capitalist) sector. Thus, in neo-Marxian theories of underdevelopment, laborers are categorized either as (a) wage-workers whom capital exploits in the formal capitalist sector and (b) the reserve army of laborers in the informal (non-capitalist) economy which enhances class-exploitation in the former and whose surplus is extracted by capital via subcontracting or outsourcing (Portes and Walton 1981; Moser 1978; Gerry 1987). Thus, even if the economy is seen to be fundamentally fractured between capitalist and non-capitalist spaces, the entire existence and reproduction of the non-capitalist (informal) space is theorized to be ultimately determined by the dynamics of the capitalist space. While this discourse sheds important light on the capitalist development process, it also has the effect of suppressing the specificities and possibilities of a non-capitalist dynamics of the informal economy that can challenge the logic of capital—in the end denigrating the potentiality and agency of a non-capitalist alternative to capital's rule. In this regard, it will be more fertile to produce a theoretical narrative that sees capital and its "outside" as spaces structured by alternative logics of production. This will produce a fresh perspective on the informal economy—and a new line of research—that questions the universality of capital, challenges the capitalocentric understanding of the economy, and begins to imagine alternative paths to development.

Conclusion

In this chapter, we have argued that the informal economy requires a more nuanced perspective than that offered by the ones dominant in the literature. In particular, we held that the informal enterprises constitute a production space that is organized around provisioning of economic needs rather than accumulation of economic surplus. Further, we have distinguished between processes of informalization and exclusion. Informalization refers, on one hand, to a reorganization of capital-labor relations (e.g., effects of labor market reforms like casualization of job contracts) and, on the other hand, to dispersion of production within the circuit

of capital (e.g., subcontracting and outsourcing). Exclusion refers to a discontinuity in the labor market, between wage-laborers, disguised wage-laborers in putting-out relations and unemployed wage-laborers on one hand, and petty producers on the other. The former derive their subsistence directly or indirectly from the economy organized around accumulation. The latter secure their subsistence in an economy organized around the economic motive of needs. While globalization has exacerbated the processes of informalization, processes of exclusion are integral to the workings of capital and cuts through policy regimes. The emphasis on outsourcing and subcontracting in the literature on formal-informal linkages is misplaced. A vast majority of the informal enterprises does not have any contractual relations with the formal sector. Using NSSO data, we have shown that, first, contractual relation between the formal and the informal sectors is not a major characteristic, and, second, empirical evidence suggests that contractual relations between enterprises belonging to the two different sectors need not necessarily be beneficial to the informal enterprises. Finally, we have proposed that a fresh research program needs to emerge that recognizes the limitations of the dominant formulations, and explores the emerging perspective based on the dynamic interaction of need and accumulation as organizational principles of production.

References

Arimah BC (2001) Nature and determinants of the linkages between informal and formal sector enterprises in Nigeria. Afr Dev Rev 13(1):114–144

Bardhan P (2009) Notes on the political economy of India's tortuous transition. Econ Politi Wkly 44(49):31–36

Bhattacharya R, Sanyal KK (2011) Bypassing the squalor: new towns, immaterial labour and exclusion in post-colonial urbanisation. Econ Politi Wkly XLVI(31):41–48

Bhattacharya S (2010) Non/capital, class and development: The case of informal manufacturing in India. PhD Thesis, University of Notre Dame. Available at: http://etd.nd.edu/ETD-db/theses/available/etd-07232010-064338/unrestricted/BhattacharyaS072010.pdf. Accessed 4 March 2012

De Soto H (1989) The other path: the invisible revolution in the third world. Harper and Row, New York

De Soto H (2000) The mystery of capital: why capitalism succeeds in the west and fails everywhere else. Basic Books, New York

Fajnzylber P, Maloney W, Rojas GM (2009) Microenterprise dynamics in developing countries: how similar are they to those in the industrialized world? evidence from Mexico. World Bank Econ Rev 20(3):389–420

Gerry C (1987) Developing economies and the informal sector in historical perspective, Annals American Acad Pol Soc Sc, 493:100–119

Hirway I (1999) Estimating work force using time use statistics in india and its implications for employment policies. Paper presented at the international seminar on time use studies, Ahmedabad, India, 7–10 Dec. Available at: http://hdrc.undp.org.in/resoruces/gnrl/ThmticResrce/gndr/Indira_Estimate.pdf. Accessed on 27 May 2005

House WJ (1984) Nairobi's informal sector: dynamic entrepreneurs or surplus labor? Econ Dev Cult Chang 32(2):277–302

International Labour Office (1972) Employment, incomes and equality. A strategy for increasing productive employment in Kenya. ILO, Geneva
International Labour Office (1985) Informal sector in Africa. Job and skills programme for Africa (JASPA). ILO, Addis Ababa
La Porta R, Shleifer A (2008) The unofficial economy and economic development. Brookings Pap Econ Act 39(2):275–363
Maloney W (2004) Informality revisited. World Dev 32(7):1159–1178
Mandelman FS, Montes-Rojas GV (2009) Is self-employment and microentrepreneurship a desired outcome? World Dev 37(12):1914–1925
McGee TG (1973) Peasants in the cities: a paradox, a paradox, a most ingenious paradox. Hum Organ 32(2):135–142
Moser C (1978) Informal sector or petty commodity production? dualism or dependence in urban development? World Dev 6 :9–10
NCEUS (National Commission for Enterprises in Unorganized Sector) (2008) Definitional and statistical issues: task force report. NCEUS, Government of India, New Delhi
NSSO (National Sample Survey Organization) (2007a) Operational characteristics of unorganized manufacturing enterprises in India, Round 62nd (2005-06) Report No. 524. Ministry of Statistics and Program Implementation. Government of India, New Delhi
NSSO (National Sample Survey Organization) (2007b) Unorganized manufacturing sector in India: input, output and value-added. Round 62nd (2005–06) Report No. 526. Ministry of Statistics and Program Implementation. Government of India, New Delhi
Nun J (2000) The end of work and the marginal mass thesis. Lat Am Perspect 27(1):6–32
Oshima HT (1971) Labour force explosion and the labour intensive sector in Asian growth. Econ Dev Cult Change 19(2):161–183
Ostrom E (1990) Governing the commons: the evolution of institutions for collective action. Cambridge University Press, New York
Ostrom E, Burger J, Field CB, Norgaard RB, Policansky D (1999) Sustainability: revisiting the commons: local lessons, global challenges. Science 284(5412):278–282
Portes A, Walton J (1981) Labor, class, and the international system. Academic Press, New York
Quijano A, Westwell P (1983) Imperialism and marginality in latin America. Lat Am Perspect 10(2/3):76–83
Ranis G, Stewart F (1999) V-goods and the role of the urban informal sector in development. Econ Dev Cult Change 47(2):259–288
Resnick SA, Wolff RD (1987) Knowledge and class. University of Chicago Press Ltd, Chicago
Ray D (1998) Development economics. Princeton University Press, Princeton
Sainz JPP (1998) The new faces of informality in central America. J Lat Am Stud 30(1):157–179
Sahu PP (2010) Subcontracting in India's unorganised manufacturing sector: a mode of adoption or exploitation? J S Asian Dev 5(1):53–83
Sanyal KK (2007) Rethinking capitalist development: primitive accumulation, governmentality and post-colonial capitalism. Routledge, New Delhi and New York
Sanyal KK, Bhattacharya R (2009) Beyond the factory: globalisation, informalisation of production and the new locations of labour. Econ Politi Wkly 44(22):35–44
Sethuraman SV (1975) Urbanization and employment: a case study of Djakarta. Int Labour Rev 112(2–3):191–206
Tokman V (1978) An exploration into the nature of informal–formal sector relationships. World Dev 6(9/10):1065–1075

Chapter 15
Indian Gems and Jewelry Industry: An Enquiry into the Nature of Competitive Advantage

Sudeshna Chattopadhyay and Sarmila Banerjee

Introduction

In the initial stage of globalization the competitiveness of developing countries is often grounded in inefficiency-driven factors like cheap labor, weak environmental norms, etc. Rapid expansion of production and exports of a number of labor-intensive low technology industries like textiles, footwear, etc., in developing countries like India and China has been largely due to the incorporation of the firms from these countries in national and global value chains following rapid expansion of subcontracting activities. Globalization and improvement of transport and communication technology has led to decoupling of labor-intensive manufacturing activities from marketing and retailing activities and relocation of the labor-intensive stages of production to developing countries with low wages and weak social standards. In case of apparel production, for example, most firms with brand labels do not take part directly in any manufacturing activity but focus on design, marketing, and retailing. The labor-intensive manufacturing activities are mostly outsourced to firms in the developing countries to take advantage of the cheap labor and weak social regulation in these countries. Specialization in these labor-intensive low value added services has no doubt led to growth of jobs and foreign earnings in the less-developed countries but the consequent increase in employment has mostly taken place in the informal sector which is marked by low wages, unskilled work, and sweatshop conditions of employment (Christerson and Appelbaum 1995; Scott 2006).

These pseudo bases of comparative advantages are not sustainable in the long-run. If these initial bottlenecks can be overcome and genuine competitiveness can be

S. Chattopadhyay (✉)
Bidhannagar College, Kolkata, India
e-mail: sudeshna.chattopadhyay@gmail.com

S. Banerjee
Department ofEconomics, University of Calcutta, Kolkata 700050, India
e-mail: sarmila_b@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_15,

achieved through innovation and technical progress then only there will be eventual takeoff on the road of development (Porter 1992). In this case expansion of export from these sustainable sectors will not only lead to higher growth but can also ensure improved living standard. The gems and jewelry industry has witnessed spectacular export growth in the post-liberalization period and has emerged as one of India's leading foreign exchange earners (www.rbi.org, www.ibef.org). The question that we would like to address in this paper is whether this competitiveness is based on in-built weaknesses like existence of lax social standards or representing genuine comparative advantage of the industry in the international market.

The rest of the paper is organized as follows: In section II, (Trend and Productivity at the Industry Level), an analysis of trend, growth, and productivity performance of the Gems and Jewelry sector of India is reported by using the Annual Survey of Industries (ASI) data between 1984–1985 and 2004–2005. In section III, (Technical Efficiency at the Firm Level), firm level ASI data have been used to carry out an analysis of the changing pattern of technical efficiency of the firms across different policy regimes. To study the linkages among production, R&D, and export we have identified 20 major gems and jewelry exporting firms from the PROWESS database of CMIE and results are reported in section IV, (Profile of Major Exporting Firms). In the following section V, (Morphology of Production Nexus), the special organizational characteristics of gems and jewelry manufacturing units in India have been discussed and in section VI (Primary Survey) the findings of a primary survey conducted on the gold and silver ornament manufacturing units of Kolkata during 2008–2009 has been reported to get first-hand feel of the ground reality. Section VII (Concluding Observations) concludes the paper.

Trend and Productivity at the Industry Level

An analysis of trend, growth rate as well as total factor productivity of the gems and jewelry manufacturing sector has been carried out over a 21-year period from 1984–1985 to 2004–2005. The entire period has been divided into two sub-periods: 1984–1985 to 1994–1995 and 1995–1996 to 2004–2005, where the first sub-period corresponds to pre-reform, pre-WTO situation and the second sub-period represents the post-globalization state. To study the trend we regressed the annual gross value added (Y_t) expressed at 1993–1994 prices, against time (t) and the growth rate is obtained by regressing the natural logarithm of gross value added ($\ln Y_t$) against time (t). The regression coefficients (β's) are representing deterministic time trends and growth rates, respectively.

The results reported in the Table 15.1 indicate an increase in trend value with a decline in growth rate in the later sub-period. To check the hypothesis of structural

Table 15.1 Trend and growth of gems and jewelry production: 1984–1985 to 2004–2005

Equation	Sub-period	β-Coefficient	t-Value	p-Value	$\bar{R}^2$	df
$Y_t = \alpha + \beta t$	I	2601.00	8.30	0.00	0.87	9
	II	7109.17	2.81	0.02	0.43	8
	I + II	5915.61	9.41	0.00	0.81	19
$\ln Y_t = \alpha + \beta t$	I	0.28	9.20	0.00	0.89	9
	II	0.10	14.53	0.00	0.51	8
	I + II	0.21	14.53	0.00	0.91	19

Source Calculated from ASI data

break a Chow test has been carried out[1] where the F-value (2.84) turned out to be insignificant at 5 % level implying a stable structure across the two regimes.

At the next stage an attempt has been made to verify the evidence of technological progress, if any, over this entire time period. With an implicit assumption of a CRS production technology, the Solow Index[2] helps to decompose this growth of gross value added into the contribution of specific factors (L_t, K_t) as well as that of the total factor productivity A_t. So, $\frac{\Delta Y_t}{Y_t} = \frac{\Delta A_t}{A_t} + w_t \frac{\Delta L_t}{L_t} + r_t \frac{\Delta K_t}{K_t}$ where $\frac{\Delta Y_t}{Y_t}$ =rate of change of gross value added; $\frac{\Delta L_t}{L_t}$ = rate of change of labor force; $\frac{\Delta K_t}{K_t}$ = rate of change of fixed capital; w_t = share of labor in gross value added; r_t = share of capital in gross value added; and $\frac{\Delta A_t}{A_t}$ rate of change of total factor productivity.

From the ASI data information on all variables except A_t can be gathered. L_t is represented by *total wages and salaries paid to employees* and K_t by *gross investment of the year* where this gross investment is defined as follows: $I_t = \frac{1}{P_t}[(K_t - K_{t-1}) + D_t]$, where I_t is the gross investment, K_t is the capital stock of period t, $D_t$is the *depreciation,* and P_t is an appropriate *deflator*. To estimate K_t, the perpetual inventory method[3] has been used and since the gems and jewelry production is reported in a residual industrial category along with a number of other non-specific products, hence, the GDP deflator is used as the relevant deflator to convert the series in constant prices. The series of $A_t$values would follow as a residual category.

By Solow index $\frac{\Delta A_t}{A_t} = \frac{\Delta Y_t}{Y_t} - \left[w_t \frac{\Delta L_t}{L_t} + r_t \frac{\Delta K_t}{K_t}\right]$ and$A_{t+1} = A_t\left[1 + \frac{\Delta A_t}{A_t}\right]$.

If $A_t$at $t = 0$ is taken as unity, then the series for A_t can be generated numerically. The estimated trend of the A_t series is obtained as: $A_t = \underset{(-9.56*)}{-0.86} - \underset{(-5.71*)}{0.04}\ t$, $\bar{R}^2 = 0.67$, $df = 19$.

---

[1] $F = \left\{[RSS_p - (RSS_1 + RSS_2)]/k \Big/ [(RSS_1 + RSS_2)/(n - k)\right\} \sim F_{k,(n-k)}$, and $F_{(2,17),0.5} = 19.4$.

[2] Growth accounting approach.

[3] K_t is obtained by the perpetual inventory method as follows:

$K_t = K_0 + \sum_{i=1}^{t-1} I_{t-i}$, where K_o is the capital stock in the benchmark year. Here we have selected 1973–1974 as the benchmark year and 1993–1994 as the reference period for GDP-deflator.

The trend is negative and statistically significant indicating that there is no sign of temporal improvement in total factor productivity. Thus, the possibility of overall technological up gradation in the production front is not confirmed by the available data.

Technical Efficiency at the Firm Level

To examine whether there has been any significant change in the output-oriented technical efficiency in the sector we have derived the firm level technical efficiency scores using Data envelopment analysis (DEA).[4] The analysis has been done for the years 1989–1990, 1993–1994, 2000–2001, and 2004–2005 with unit level ASI data at 5-digit level of disaggregation. The number of firms for which all relevant information was available for the year 1989–1990 was 42 whereas the number increased to 188 by 2004–2005.[5] Table 15.2 presents the technical efficiency scores for the selected years in terms of size classes defined with 0.10 class widths. Table 15.3 reports the corresponding descriptive statistics. Figure 15.1 depicts the stacked bars showing the change in relative frequency of firms with technical efficiency scores below 0.50 for these selected years.

There has been an overall decline in the output-oriented technical efficiency of the sector over the period of our study. A number of Fisher's t tests were conducted for different pairs of samples to check whether the sample means in each of the years under consideration is significantly different from that of the preceding years. In all cases the difference turned out to be highly significant, both under the assumptions of equal as well as unequal variances. The share of most inefficient firms with technical efficiency score less than 0.2 has gone up from 2.38 to 65.5 % between 1989–1990 and 2004–2005. This is evident from Fig. 15.1 also. The relative frequency of the firms with technical efficiency score below 0.5 has steadily increased over time. The skewness of the distributions have consistently changed from −1.92 to +1.78 and the median value remains always below mean in more recent years, confirming this concentration at the lower end of the distribution in later period.

In spite of fabulous export performance this fall in technical efficiency is indicating an interesting possibility: after reform the scope of becoming cost-competitive in the international market has been enhanced for the relatively inefficient new entrants also, who are mostly organizing production through procuring export-contracts, importing raw materials, and out-sourcing the manufacturing process to the informal units to make quick money from the entire transaction. This possibility, if confirmed, will endorse perpetuation of low

[4] Ray (2004) and Coelli (1998) present comprehensive discussion of DEA technique.

[5] A firm is retained for DEA analysis provided information on all relevant variables like input-used, output-produced, etc., is available on it. In ASI unit level data, sometimes some of the relevant information on the reported firms is missing.

Table 15.2 Distribution of output oriented technical efficiency scores

Class boundary	No. of firms				Relative frequency (%)			
	1989–1990	1993–1994	2000–2001	2004–2005	1989–1990	1993–1994	2000–2001	2004–2005
0.0–0.1	0	7	32	90	0.00	8.75	31.07	47.87
0.1–0.2	1	15	16	33	2.38	18.75	15.53	17.55
0.2–0.3	1	9	20	23	2.38	11.25	19.42	12.23
0.3–0.4	0	15	9	10	0.00	18.75	8.74	5.32
0.4–0.5	1	8	7	7	2.38	10.00	6.80	3.72
0.5–0.6	0	4	1	0	0.00	5.00	0.97	0.00
0.6–0.7	5	2	1	3	11.90	2.50	0.97	1.60
0.7–0.8	5	5	1	1	11.90	6.25	0.97	0.53
0.8–0.9	8	2	1	1	19.05	2.50	0.97	0.53
0.9–1.0	21	13	15	20	50.00	16.25	14.56	10.64
Total	42	80	103	188	100	100	100	100

Source Calculated from ASI data

Table 15.3 Descriptive statistics of frequency distribution of output-oriented technical efficiency scores

Summary statistics	1989–1990	1993–1994	2000–2001	2004–2005
Mean	0.843	0.422	0.320	0.235
Median	0.902	0.357	0.225	0.105
Standard deviation	0.203	0.311	0.321	0.300
Coefficient of variation	4.89	22.92	32.20	38.30
Skewness	−1.92	0.72	1.25	1.78
Count	42	80	103	188

Source Calculated from ASI data

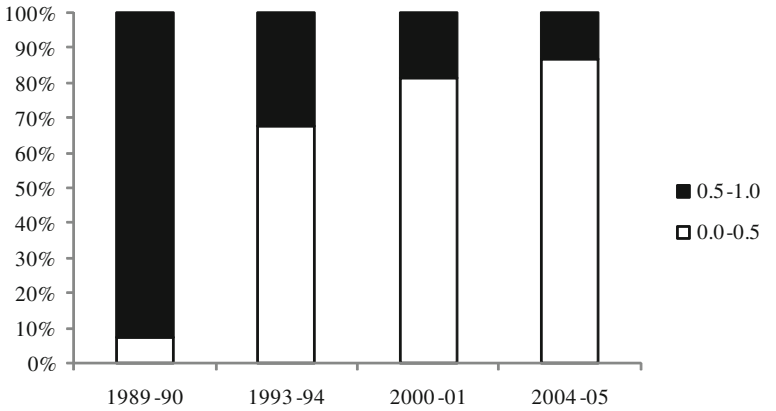


Fig. 15.1 Stacked bars of output-oriented relative technical efficiency scores. *Source* Authors' estimation

technical efficiency and the long-run sustainability of the recently experienced fortune of the gems and jewelry industry will be questionable.

However, the ASI database is not suitable for verifying this line of explanation as here no export data are reported at the firm level and the information on imported inputs and R&D expenditures are not always adequate. To overcome this limitation we have utilized the database on listed gems and jewelry making companies consolidated by the Centre for Monitoring Indian Economy (CMIE) in its PROWESS database for the post-reform period.

Profile of Major Exporting Firms

The CMIE-PROWESS database provides us the company level data from their balance sheet for listed and other companies.[6] To carry out the analysis the following variables have been defined[7]: **Output,** material **Inputs**, **Labor, Capital, Export** of final product, **Import** of raw material, and **R&D**. To express the variables in real terms to make them inter-temporally comparable, each variable is deflated by using the GDP deflator with 1993–1994 as the base year. We have analyzed the overall performance of the firms reported in the CMIE firm level database for four selected years, viz., 1994 (the year immediately preceding the introduction of WTO), 1996 (to study the early effects of the emergence of this supra-national institution on this industry), 2000 (to capture the effects of newer policies like SPS and TBT agreements), and 2005 (when the effect of the new world economic order has been prominently felt on all important transactions). The number of firms under gems and jewelry industries reported in the CMIE database has increased steadily from 21 in 1994 to 80 in 2005. There has been more than 50 % increase in the average value of output of the firms reported in the database over this period, indicating that the sector has made commendable progress in the post-globalization period (Table 15.4).

As is evident from Table 15.4 the total value of export of the firms reported in the database increased by more than nine times while the average value of export has been more than doubled confirming the importance of gems and jewelry sector as a major foreign exchange earner. The average export intensity (i.e., export as a percentage of total output) of the firms reported here is close to 60 % or more for all the years under consideration (Table 15.5); however, the investment in R&D is very low. In fact, the average expenditure on R&D never exceeded 0.4 % of the value of output during this period (Table 15.5). So, the sector did not benefit much in terms of improved production technology. The production process being highly labor intensive in nature it is natural to expect that the benefit of buoyant market would be reaped by labor. However, contrary to our expectation the average value of compensation to employees has actually declined between 1994 and 2005 even though the average value of output of the industry has witnessed remarkable growth (Table 15.4). The average value of net fixed assets has also declined.

---

[6] CMIE reports information on all companies traded in India's major stock exchanges and several others including the central public sector enterprises.

[7] **Output** of the firm is the value of total output defined as the total sales of the goods of the firms plus the change in the stock of finished goods measured in terms of the opening stock minus the closing stock. The inputs are: Intermediate **Inputs**, measured in terms of the companies' expenditure for raw material plus energy inputs (given by the expenditure on power and fuel), **Labor,** measured in terms of compensation to employees, and **Capital** as the value of net fixed assets. In addition **Export** has been taken as the value of goods export and **Import** is the value of raw material import. **R&D** represents the total research and development expenditure incurred by the relevant firm.

Table 15.4 Year-wise total and average value of selected variables (Rs. in crores)

Years	Indicators	Output	Input	Labor	Capital	R&D	Export	Import
1994	No. of firms	21	21	21	21	21	21	21
	Total value	1140.08	656.67	70.98	327.91	4.08	815.84	529.74
	Average value	54.29	31.27	3.38	15.61	0.19	38.85	25.23
1996	No. of firms	33	33	33	33	33	33	33
	Total value	2194.52	1133.66	111.11	497.88	3.73	1314.22	872.1
	Average value	66.5	34.35	3.37	15.09	0.11	39.82	26.43
2000	No. of firms	43	43	43	43	43	43	43
	Total value	3825.12	2542.65	196.57	690.31	5.16	2769.17	1855.99
	Average value	88.96	59.13	4.57	16.05	0.12	64.4	43.16
2005	No. of firms	80	80	80	80	80	80	80
	Total value	6704.4	4315.47	208.72	1034.6	5.22	7571.01	2575.34
	Average value	83.81	53.94	2.61	12.93	0.07	94.64	32.19

Source Calculated from CMIE-PROWESS database

Table 15.5 Year-wise share of selected variables

Indicator	Indicators	1994	1996	2000	2005
R&D	No. of firms	1	2	3	2
	Firms (%)	4.76	6.06	6.98	2.50
	Value of output (%)	0.36	0.17	0.13	0.08
Export	No. of firms	19	27	33	54
	Firms (%)	90.48	81.82	76.74	67.5
	Value of output (%)	71.56	59.89	72.39	112.93[a]
Import of Raw Material	No. of firms	18	22	28	43
	Firms (%)	85.71	66.67	65.12	53.75
	Expenditure on inputs (%)	80.67	76.93	72.99	59.68

Source: Calculated from CMIE-PROWESS Database

[a] Export as a percentage of total value of output exceeds 100% as the current export figures includes the value of goods exported from the unsold stock of output produced in an earlier financial year

Thus, the magnificent growth of the sector failed to create a positive impact on the share of labor or capital. The gems and jewelry industry reveals dependence on the imported inputs in the form of good quality gold and uncut precious and semiprecious stones. So, the average share of imported raw materials in total expenditure on intermediate inputs is greater than 50 % for all the years under consideration (Table 15.5).

With the expansion in the domestic market followed by high growth rate of GDP in the recent decade the use of domestic supply of intermediate input has gone up significantly in this sector (from about 20 % in 1994 to more than 40 % in 2005) indicating the emergence of a new group of firms who are mostly catering to the needs of the domestic market. Those who are participating in the global market for them the dependence on imported raw materials is still around 70 % (Table 15.6).

Table 15.6 Year-wise import dependence of exporting firms

Exporting firm with	Indicators	1994	1996	2005	2006
Use of imported inputs	No. of firms	17	21	26	41
	Firms (%)	89.47	77.78	78.79	75.93
	Expenditure on inputs (%)	84.88	82.81	74.78	68.09
	Average value of export (Rs. in crores)	47.43	62.44	102.64	110.38

Source Calculated from CMIE-PROWESS database

Thus, inadequate investment in R&D, meager share of labor in total output and heavy dependence on imported raw materials appeared to be the essential features of the gems and jewelry exporting firms. These observations endorse our earlier surmise of the presence of pseudo comparative advantage for this sector in a globalized world coming in the guise of weak regulatory control for social standards. At the next step, a panel regression will be carried out to verify this possibility on a stronger statistical footing.

Twenty listed firms have been identified from the CMIE database on which all relevant information[8] like total output (OUT), total exports (EXP), expenditure on intermediate inputs (INP), expenditure on labor measured in terms of compensation to employees (W), expenditure on capital (K), and import of raw materials (IMR) was available for three selected years, viz., 1996, 2000, and 2005. This helped us to generate a panel of 60 observations. Since the within group variations for all the selected variables are small while the between group variations are large, our panel appears to be balanced and consistent.

In the regression analysis OUT and EXP are taken as two dependent variables and only for OUT time dummies are incorporated to capture the temporal effects across different policy regimes, if any. For the regression of OUT (i.e., regression 1), EXP, INP, W, K, and IMR are taken as independent variables. Regression 2 is same as regression 1 with two time dummies as additional regressors. In regression 3 the dependent variable is EXP and the independent variables are the different components of cost. The results are reported in Table 15.7. For all regressions, both fixed effect and random effect models are estimated and on the basis of Hauseman (χ^2) test[9] the appropriate method of estimation has been selected.

For OUT, (regression 1), in terms of Hauseman's χ^2test the Fixed Effect model has been accepted. This means there exists reasonable degree of heterogeneity among the constituent firms of the sector in terms of their input–output

[8] As we have discussed earlier, expenditure on R&D is extremely low and very few firms reported any expenditure on R&D during the years under consideration. Hence, it is apparent that R&D has not exercised any significant influence on the performance of the sector. So, R&D has not been included as an explanatory variable in our regression analysis.

[9] with Random Effect as the Null Hypothesis.

Table 15.7 Regression analysis of firm level performance

Y-variable	OUT		EXP
Regression	1	2	3
X-variables	*Fixed effect*	*LSDV (Time Dummy)*	*Random effect*
INP	1.033*	0.623*	0.083
W	−0.773	4.297	−1.078
K	0.011	1.065	0.270
EXP	1.727***	1.140***	–
IMR	−2.206**	−1.003	1.308***
D_1	–	−11.178	–
D_2	–	13.488	–
$\bar{R}^2$	–	0.890	–
Hausman- χ^2	12.09** (FE)	–	4.03 (RE)
ρ	0.813	–	0.418
F: [H_0: $u_i = 0$]	0.70	–	–
F: [H_0: $\beta = 0$]	–	60.25***	–
Wald-χ^2	–	–	618.76***
BPLM-χ^2	–	–	9.67***
No. of groups	20	20	20
No. of time Periods	3	3	3
No. of observations	60	60	60

Source Calculated from CMIE-PROWESS database
*Significant at 1% level;
** Significant at 5% level;
*** Significant at 10% level

correspondence and therefore, the firm-specific effects are dominantly present in the system. Here, explanatory variables like INP, IMR, and EXP are important factors; however, W and K do not appear to be statistically significant. The performance of OUT has been strongly influenced by EXP. So, at least, a part of the growth of the sector is export-led.

To explore the impact of policy changes on output of the sector the dummy variable regression is reported (regression 2). Here, IMR lost its statistical significance. However, export still turns out to be statistically important. None of the time dummies are significant. This result is re-endorsing our earlier analysis and it may be concluded that globalization did not impart any significant influence on the trend, growth rate of output, productivity, and technical efficiency of this sector.

Finally, in the regression of EXP on the different cost components (regression 3) only IMR turns out to be statistically significant. Neither W nor K or INP appears to be statistically significant. So, the export performance is neither explained by improved machineries nor by the improved fortune of the workers.

Here the Random Effect model turned out to be the more appropriate one.[10] It seems to indicate that the more export-oriented production unit is importing raw material, out-sourcing the follow-up processes to the informal sector and re-exporting the finished product. Since in a putting out system the intermediary firms are tied to the lead firms and organize production strictly according to the guidelines specified by the lead firm, the intermediary firm has little incentive or freedom to experiment and innovate. This possibility could become the reality provided the production nexus of the sector would reveal close connectivity between formal and informal units. So, in the following section we will look into the relevant literature to investigate whether the formal and informal sector firms in the gems and jewelry industry are interlinked through the subcontracting chain, identify the factors which might form the basis of such linkage and examine whether the inter connectedness between the formal and the informal sector explain the pseudo competitiveness of the industry in international market.

Morphology of Production Nexus

An interesting feature of the gems and jewelry manufacturing industry in India is that despite its huge size, the sector is largely unorganized.[11] The jewelry that finds a market across the country and is exported to other parts of the world is often crafted in tiny informal workshops. These informal jewelry firms carry out different stages of the manufacturing process on behalf of the large formal sector firms in return of *making charges*. Thus, an important feature of the gems and jewelry industry in India is the interconnectedness between the formal and the informal sector firms through the production chain. Though it is difficult to get exact figures due to the informal nature of the units, the contribution of these small informal sector units in total export of jewelry is quite high. As per figures from the Gem and Jewelry Export Promotion Council, during 2000–2001, exports from the Domestic Tariff Area (DTA) in Mumbai amounted to $266 million and a major share of which was from Zaveri Bazaar area where the small informal jewelry manufacturing units of Mumbai are concentrated (Katakam 2001). Mumbai is no exception. In fact, in all the regions of India which specialize in crafting of jewelry

[10] It is likely that the exporting firms function in a putting out system and work as a part of global production chain where the main function of these exporting intermediary firms is to import raw materials, get the production stages done by the informal sector units and re-export the finished product. For example, the diamond processing firms import rough diamonds and after processing re-export the processed diamonds. In fact 85 % of the world's diamonds are processed in India and the diamond cutting firms in India has strong production links with the leading diamond retailers and manufacturers like De Beers, Alrosa, BHP Billiton (www.afdiamonds.com.,http://catchef.ft.com).

[11] Organized players such as Tata with its Tanishq brand constitute only 4 % of the market (www.commodityonline.com/news).

or polishing of precious and semiprecious stones there is huge concentration of small informal sector firms accounting for major portion of employment, production, and export. For example in Jaipur, which is the second most important gem exporting center in India and number one in terms of sales to the foreign tourists, approximately 90 % of the 0.2 million workers employed in gem polishing industry work in unorganized units (www.lac.org.hk). The informal artisans not only get meager payments for their work but are also exposed to serious occupational health hazards as in most of the informal units labor safety and pollution prevention norms are not followed.[12] In the precious and semiprecious stone cutting and polishing industry in different parts of India a large number of workers and their families are affected by deadly silicosis. A study carried out by National Institute of Occupational Health (NIOH), India in 2003 confirmed the presence high level of silica dust in the gem processing units in Jaipur (NIOH 2003; Patel 2006, 2007). Another recent NIOH study (2007) conducted to assess the health status of the workers manufacturing diamonds in Surat raise serious concern over the health hazards faced by the workers in the flourishing diamond processing industry (www.nioh.org).

From the above discussion it is apparent that the two main features of the gems and jewelry manufacturing industry in India that may be held responsible for its spectacular success in the international market are extensive interlinkage between the formal and informal sector through the production chain and weak regulatory pressure of compliance to social standards related to labor and environmental norms. Hence, to get further insights into the organizational setup of the industry a primary survey on the jewelry making units of Kolkata has been conducted in 2008–2009. The findings of that survey are reported in the following section.

Primary Survey

We have chosen Kolkata as our area of survey[13] as a large number of both registered and unregistered gems and jewelry producing firms are located here. Moreover, jewelry exports from the eastern region have grown at over 21 %

[12] The most distressing feature of this industry is that a large percentage of the workers employed in the industry especially in the gem cutting and polishing segment are children. The children are preferred here because of their lower wage and tender fingers especially suitable for cutting and polishing the smaller stones. In the flourishing gems polishing industry at Jaipur nearly 20 % of the workforce are children (Burra 1988a; Kruijtbosch 1996). One of the main reasons why India is able to compete in the international market in spite of not getting the best quality raw material is this availability of cheap labor who can be subjected to inhuman working conditions. While the artisans get meager payments and exposed to serious health hazards the profit that the traders and the exporters make on the fine craftsmanship is over a thousand percent (Burra 1988b).

[13] The primary survey has been conducted as a part of the research project titled "Environmental Regulation and Process Sub-Contracting: A Study of Gems and Jewelry Manufacturing Units in

compared to an all India growth rate of 6 % and Kolkata accounts for almost all of the jewelry exports from the eastern region (www.wbidc.com). Kolkata mainly manufactures plain gold and silver jewelry and most of jewelry produced here is crafted in the small informal manufacturing units. Here most of the goldsmiths are concentrated in a few localities: Bowbazar & Taltala in Central Kolkata, Garanhata, Sovabazar & Sinthi in North Kolkata and Bhowanipur-Kalighat, Ballygaunge-Dhakuria & Behala in South Kolkata. We have selected the three sites for our field survey; Bowbazar from Central and Garanhata and Sinthi from North.[14] Bowbazar represents the central business district where majority of transaction takes place. Garanhata is the oldest location where the trade concentrated even in pre-independence days and Sinthi in the North has become a major production center over the past quarter of the century, especially after the influx of refugees from Bangladesh. A number of the immigrants were connected with gold smithies in Bangladesh and most of these families settled down along the railway network connected through Dumdum junction. Sinthi being close to Dumdum, flourished over recent time. There are approximately 20,000 goldsmiths in Bowbwzar.[15] In Garanhata there are about 3,000 while in Sinthi there are about 25,000 manufacturing units.

The gold jewelry manufacturing process can be divided into the following major stages:

(a) *Refining*. If the gold (silver) is of inferior quality or if the gold (silver) is *recycled one* from old jewelry then it has to be refined before it can be used for ornament making.

(b) *Crafting or making of jewelry*. The Indian goldsmiths are famous for their jewelry crafting skills. Most of the jewelleries produced in India are meticulously hand crafted.

(c) *Finishing*. Finishing includes polishing of jewelry and *chila* work.

(d) *Galai or recycling*. Some amount of gold is wasted during the crafting, polishing, and refining processes. During the crafting or polishing of gold jewelry fine gold particles come out from the ornament and gather on the floor-dust of the workshop. Some amount of the gold also gets dissolved in the acid that is used for polishing ornaments and refining gold. The shop owners collect the gold particles mixed with dust and other impurities and take them to some special units referred to as *Gold Galai* units where the artisans are especially trained to separate the gold from the impurities.

(Footnote 13 continued)
Kolkata", UGC Sanction No: PWH 007/06–07(ERO) sponsored by the University Grants Commission (UGC) under the UGC Minor Research Project scheme.

[14] No unit from South-zone has been surveyed as we had to abandon the survey abruptly because of active resistance from the local Gold Traders associations.

[15] The statistics on the number of goldsmiths in different locations in Kolkata has been provided by Bangiya Swarna Silpi Samiti which is a major association of gold and silver jewelry manufacturers in West Bengal.

The initial plan was to collect information on 300 units each, from both Bowbazar and Sinthi and another 200 units from Garanhata. Though the goldsmiths were initially reluctant to give interviews, eventually they cooperated and in Bowbazar and Garanhata no major problem was faced in conducting the survey as per schedule. However, in Sinthi due to extreme reluctance and active resistance from the goldsmiths information could be collected from 85 units[16] only before abandoning the survey formally.

The salient features of the majority of units surveyed are their informal status and extensive interlinkage through the production chain within themselves and with the larger formal sector firms. Other important features relate to the poor compliance of the informal units to social norms and lack of investment in improvement of infrastructure, technology, and human capital.

Interlinkage Between the Formal–Informal Units

A unit to be labeled as a formal unit, must be registered with a number of official bodies and obtain a host of licenses like the Trade license, Manufacturing license, etc. Some of these licenses are obtained at an individual level (required for both firm-owners and self-employed persons) and some others are at the firm level. Most of the units surveyed did not have all the required licenses.

It is interesting to note that out of the total 586 units surveyed as many as 416 units (71 %) either did not have any name or were reluctant to disclose it. Out of these 416 units, 286 have either NOC, i.e., Consent to Operate license from the Pollution Control Board or manufacturing (*Karigari*) license from Kolkata Municipal Corporation (KMC) or both and the remaining 130 do not have any license at all (Table 15.8). Hence, these 130 are purely informal unregistered units. In fact only 33 manufacturing units claimed to have all kinds of licenses, and, therefore, formal in nature. However, it will be shown later that even these units are employing labor on an informal basis.

Since the jewelry making process is divided into a number of operations, each demanding some specialized craftsmanship, most of the units surveyed concentrate in a sub-group of operations. Table 15.9 reports the nature of specialization of the survey units. Out of 586 units surveyed 58.2 % are specializing in any particular stage of production (refining, making, finishing, or recycling) and the rest is

[16] The reluctance of the goldsmiths to face the interviewers might be due to the fact that most of the units did not have the necessary licenses and did not follow the laws especially those related to labor and environmental standards. The goldsmiths were apprehensive that due to increased public attention through media coverage they might be subjected to serious regulatory monitoring. The resistance from the goldsmiths was so hostile that we were forced to abandon our survey after interviewing only 85 units there. Sinthi is reportedly infamous for clandestine cross border trade in gold. This might be the reason why we had to face extraordinarily active resistance from the goldsmiths.

Table 15.8 License status of the survey units

Name of the license(s)	No. of units in Bowbazar (%)	No. of units in Garanhata (%)	No of units in Sinthi (%)	Total no. of units (%)
NOC	91 (30.2)	51 (25.5)	4 (1.3)	146 (24.9)
Trade license	292 (97.0)	199 (99.5)	83 (99.8)	574 (97.9)
Karigari license	108 (35.8)	47 (23.4)	19 (22.4)	174 (29.7)
Health License	190 (63.1)	104 (52.0)	27 (31.8)	321 (54.8)
NOC + Trade license	90 (29.9)	51 (25.5)	4 (1.3)	145 (24.7)
NOC + Karigari license	31 (10.2)	3 (1.5)	0 (0.0)	34 (5.8)
NOC + Trade +Karigari + Health	30 (10.0)	3 (1.5)	0 (0.0)	33 (5.6)
Total	301 (100)	200 (100)	85 (100)	586 (100)

Source Primary survey

Table 15.9 Nature of process specialization of the survey units

Name of the processes	Numbers	Percentage
All	24	4.1
Refining(R)	55	9.4
Making(M)	163	27.8
Finishing(F)	117	20.0
R + M	148	25.3
R + F	4	0.7
M + F	69	11.8
Galai(recycle)	6	1.0
Total	586	100

Source Primary survey

involved in more than one stage. However, only 4 % are carrying out the entire process *from cradle to grave* under the same roof. Majority of units (69 %) are involved in crafting the ornaments while refining and finishing jobs are concentrated in fewer specialized hands. The small informal units get manufacturing orders from the large (registered) jewelry shops on the basis of a putting out type contract. The basic metal is generally supplied by the registered firms. The formal sector firms procure orders, organize production through sub-contracting chains, and also take care of the marketing and related promotional activities. Often, the dealers act as intermediaries between the large firms and the small informal firms. The dealer takes the order from the large firm and then gets the entire process done from the different units specialized in the respective stages. Thus, the small units are interconnected with one another and the larger firms through a long subcontracting chain.

Out of the total 586 firms surveyed only 42 firms, i.e., 7.2 % were not dependent on the larger firms and sold their manufactured items directly and

Table 15.10 Mode of disposal of liquid waste

Mode of liquid waste disposal	No. of units in Bowbazar	No. of units in Garanhata	No. of units in Sinthi	Total no. of units
Niharawala	271	168	69	508
Drain	227	158	49	434
Niharawala + Drain	204	131	34	369
None reported	7	5	1	13

Source Primary Survey

exclusively to the customers. The majority of the firms (73.7 %) handed over all their products produced on a putting out basis to other larger firms or the dealers and thus had no direct transaction with the customers. In fact, the remaining 19.1 % of firms are engaged in both types of dealings.

Environmental Compliance

The solid waste from these units consists of traces of acid, iron particles, etc., and the liquid waste consists of acid water. The gaseous wastes consist mainly of the acid and cadmium fumes emitted during production. Of the 586 units covered in our survey not a single one has any provision to treat waste. 434 units (i.e., 74.06 %) admitted that they dump their liquid wastes[17] in the municipal drain and 508 units (86.69 %) admitted that they sell the sediment from the liquid waste to the *Niharawalas*[18] (Table 15.10). Even when the firms are not admitting the accusation of direct dumping they cannot produce any concrete evidence of waste management. This allegation holds equally even for the firms claimed to have NOC from the Pollution Control Authority.

To check air pollution the firms use chimneys or exhaust fans. However, these measures are often inadequate. The acid fumes are so toxic that ordinary chimneys or exhaust are not sufficient to absorb the fumes. The air inside these tiny workshops (even those with exhaust fans and chimneys) are often filled with pungent smell of the acid, making it difficult to breathe freely. The WBPCB has designed

[17] A survey conducted by Kolkata Metropolitan Development Authority (KMDA) back in 1999 in the advent of setting up of Manikanchan SEZ in Kolkata revealed that the liquid waste, especially the acid from gold manufacturing which is being disposed through Kolkata Municipal Corporation (KMC) drain in huge quantities without any processing is slowly seeping through and is getting mixed with underground water to become a source of health hazard (KMDA, February 2003).

[18] The *Niharawalas* are a class of people who earn their living by extracting gold from the wastes that they buy from jewelry manufacturing units The *Niharawalas* do not possess any kind of license and extract gold from the wastes through a very polluting process. After extracting the gold from the residues, they dump the remaining solid waste in open vat and the liquid in public drains.

Table 15.11 Means of controlling air pollution in the survey units

Means of controlling air pollution	No. of units in Bowbazar	No. of units in Garanhata	No. of units in Sinthi	Total No. of units
Exhaust Fan	233	122	44	399
Chimney	49	40	18	107
Exhaust fan + Chimney	27	10	3	40
None	46	48	26	120

Source Primary Survey

special equipment fitted with special scrubbers to absorb the toxic acid fumes. However, not a single unit used the equipment. Among the units surveyed, 120 (20.5 %) units did not have even a chimney or an exhaust fan (Table 15.11).

To control pollution from these units the West Bengal Pollution Control Board (WBPCB) has specified that the units located within the residential areas should not use more than 1 L of sulfuric/nitric acid per month. However, as there are innumerable small sized informal units in this sector, it is almost impossible for WBPCB to monitor input use and identify units that violate the norm. Moreover, a restriction on input use with no reference to output is not a good example of a regulatory control. Surprisingly, even some of the units producing excess acid–water claimed to have NOC clearance.

Another source of pollution in gold manufacturing emanates from the use of Cadmium as soldering agent. According to the World Gold Council, Cadmium fumes are highly toxic. There is another safer alternative agent like indium soldering, but the artisans are reluctant to switch to other soldering agents to retain their cost-competitiveness. The Bureau of Indian Standards has introduced hallmarking and there is a rule which forbids issuance of hallmark to the jewelry with Cadmium. However, there are umpteen instances of evasion in this case (Basu 2007). In our survey we have seen that out of 398 firms involved in the making of the ornament about 98 % use Cadmium. Moreover, the small manufacturers have a general misconception and belief that use of Cadmium is a must in obtaining the Hallmark stamp. So, not only they are flouting norms, they are bearing wrong impression regarding the recommended practices to follow.

The environmental laws are routinely flouted by the informal units exposing the surrounding population and the goldsmiths to serious health hazards. The firms have no incentive to undertake any abatement activity as there is intense competition among the firms and they fear that the investment in abatement will increase their costs. Out of 586 units surveyed 451 (i.e., about 76.96 %) units reported that they would have to shutdown if they were forced to follow the pollution control norms.

Social Security Provisions

The small informal units are generally owned by one or more goldsmiths. Tiny ill ventilated rooms generally hired on a monthly rental basis serve as the workshop. The owner goldsmiths hire workers to work under him. In most cases the owners also participate in the manufacturing process. Often the workers are not paid in cash. Instead, the worker gets a certain percentage of the gold that is wasted during the manufacturing process. Out of the 586 firms surveyed we were able to collect information on wages from 495 firms. The average wage of the worker is about Rs. 2,000 per month. Thus, the workers on an average are forced to live on less than Rs. 67 per day which is lower than the national floor of minimum daily wage of Rs. 100 (www.paycheck.in). The workers are not only poorly paid they are also denied the benefits of statutory provident fund and employees state insurance schemes. Not a single firm reported to have these facilities for the workers. There are at least two major Trade Unions of the goldsmiths in West Bengal with their Head Offices in Kolkata, viz., *Bangiya Swarna Silpi Samiti* and *Akhil Bharatiya Swarnakar Sangha*. But the members of these organizations are generally the owner goldsmiths and so it seems that the interests of the worker engaged in gold smithies do not get enough attention.[19] Information on the monthly income of the owners of the informal units was collected from 543 firms. The average monthly income of the owners is stated to be about Rs. 4,800.00 per month the range being Rs. 1,500 to Rs. 12,000. While the income of the owners is generally higher than the workers it is much lower than the profit made by the dealers or the larger formal sector firms. Thus, in comparison to the huge profit that the large formal sector jewelry firms and exporters make out of their craft both the small owners and workers involved in this trade are ill paid and are often exposed to serious pollution in return of their hard work and exquisite artistic talent.

State of Physical and Human Capital

Improvement and up gradation of technology, infrastructure, and human capital is essential for accelerated growth of any sector. In the informal units surveyed, however, investments in all these areas were found to be abysmally low. None of the firms reported to make any investment in infrastructure development in the last 10 years. Only 77 out of the 586 units' surveyed reported purchase of some machineries like digital weighing machine, newer variety of polishing machine popularly referred to as *magnet machine*, etc., over the last 5 years. None of the firms reported any significant change in the ornament-making techniques followed

[19] During our survey most of the owners were reluctant to let their workers face the interviewers independently and most of the survey questionnaires were filled up on the basis of the interviews given by the owners alone.

by them over the decades. The only change seemed to be the introduction of cadmium as a soldering agent in place of the traditional metal solders. However, given the serious health hazard due to excessive exposure to cadmium fume, it can hardly be recorded as an improvement.

Although the artisans form the backbone of the gems and jewelry industry the manufacturing units not only show reluctance to invest in physical capital but in human capital as well. Very few of the artisans working in the informal units have completed school education and none of them has received any formal training in jewelry making. A new entrant to the jewelry making profession generally has to work as an apprentice in the workshop of a goldsmith for a few months before he starts working as a full fledged craftsman. In fact, apprenticeship in the workshop of a fellow goldsmith offers the only scope for training. As a result the goldsmiths miss the opportunity to supplement their traditional skills with modern knowledge. There are few institutes like Shingar Institute in Howrah, Jesani Institute in Mumbai, National Institute of Fashion Technology with branches in Delhi and Kolkata, which offer jewelry designing and making courses but most of the informal artisans do not satisfy the minimum educational qualification necessary for being eligible to enroll for those professional courses.

Concluding Observations

It is shown in the paper that in spite of the commendable performance of the gems and jewelry sector at the domestic and international level the capital or total factor productivity of the sector has not shown any improvement in the post-liberalization period.[20] The sector has not experienced any discernible technological up gradation during our period of study. Our analysis also revealed that the technical efficiency of the sector has declined over the period 1989–1990 to 2004–2005. Our primary survey, suggested that the formal and the informal sector firms are interlinked in this sector through an extensive subcontracting chain. The results of our survey also revealed that the informal jewelry making firms have all the features of a subsistence sector like low wage, low investment in physical and human capital formation, poor working conditions, and lack of arrangement for worker safety. Thus, through interlinkage the formal sector is getting access to the cheap labor of the informal sector and gaining competitiveness by flouting all labor and environmental norms without showing any sign of technological up gradation.

According to Porter a country has to pass through several stages before it can acquire long-run competitiveness. In the first stage, competitiveness is essentially

[20] In fact, the Indices of revealed comparative advantage show very high RCA values of gems and jewelry exports from India amounting to 8.39 for Jewelry, goldsmiths' and silversmiths' wares, etc. (SITC897) and 16.49 for Pearls, precious and semi-precious stones (SITC667) for the year 2002(Chattopadhyay 2010).

cost-based competitiveness and as Sanyal (1993) explains, in case of developing countries, competitiveness at this stage is based on factors like surplus labor or absence of environmental concern. However, such competitiveness cannot be sustained in the long run as it makes the country dependent on low productivity activities and, hence, cannot ensure sustained increases in standard of living. Our primary survey revealed that even though the artisans are the driving force behind the gems and jewelry industry little regard is paid toward their health and safety and a vast majority of them lack even basic primary education. It is important to recognize that a healthy and educated labor force is also more productive. Hence, in a labor-intensive industry like gems and jewelry investment in human capital is an essential prerequisite to encourage sustained improvements in labor productivity and to ensure favorable pre-conditions for long-run growth of the industry as a whole.

Our analysis suggests that the competitiveness in gems and jewelry manufacturing industry is still in the early stage of development as the source of this competitiveness is lying in the existence of extensive formal informal linkage in production creating scope for regulatory leakages. If the sector continues to grow only on the basis of temporary advantages derived from 'weaknesses' of the economic structure then such a growth process would merely lead to expansion of low productivity employment in the informal sector with eventual deterioration of the quality of life. Moreover, as Christerson and Appelbaum (1995) rightly pointed out, the competitive advantage of low technology, labor-intensive industrial clusters are also easily contestable as these clusters are under constant threat of being out-competed by newer cheaper wage locations. Hence, technological and qualitative up gradation of the gems and jewelry sector is essential pre-condition to ensure sustainable growth of the sector.

References

Basu R (2007) Jewellers at Toxic Risk. The Telegraph, 7 November 2007, p 25

Burra N (1988a) Exploitation of children in Jaipur gem industry I: structure of industry. Econ Polit Wkly 23(3):75–79, 16 January

Burra N (1988b) Exploitation of children in Jaipur gem industry II: health hazards of gem polishing. Econ Polit Wkly 23(4):131–138, 23 January

Coelli T (1998) A multi-stage methodology for the solution of orientated DEA models, CEPA working paper no: 1/98. Department of Econometrics, University of New England, Arimdale, Australia

Chattopadhyay S (2010) Globalization and environment: a study on gems and jewellery units of Kolkata, unpublished Ph.D dissertation, University of Calcutta

Christerson B, Appelbaum RP (1995) Global and local subcontracting: space, ethnicity, and the organization of apparel production. World Dev 23(8):1363–1374

Katakam A (2001) Sweat and gold. Frontline 18(22), November 9. http://www.frontlineonnet.com

KMDA (2003) Final Report on Identifying & Mapping of Industries of KMA Volume -1, February 2003; (Project sponsored by West Bengal Pollution Control Board under Industrial

Pollution Control Project (1D/P/105) of Japan Bank for International Co-operation and executed by Kolkata Metropolitan Development Authority)
Kruijtbosch M (1996) Child and adult labor in export-oriented garment and gem polishing industry of India with case studies from Tirupur, Bangalore, Jaipur and Trichy. Indian Committee of the Netherlands, Utrecht, 2 November 1996. www.indianet.nl/chadlab.html
National Institute of Occupational Health (NIOH) (2003) Occupational health problems, evaluation and control: A study in gem polishing industries at Jaipur. ENVIS-NIOH Newsletter, Vol. 1, No. 3, National Institute of Occupational Health. www.envisnioh.org
National Institute of Occupational Health and its Regional Centers (2007) Assessment of the health status of the diamond manufacturing workers. Highlights of the work done 1 April 2006 to 31 March 2007. www.nioh.org
Patel J (2006) A short paper on the problem of silicosis in gemstone-processing industry and legal loopholes in the implementation of the prevention of occupational disease law. Labor Action China, www.jewelrycampaign.net/eng/problemofsilicosis.doc
Patel J (2007) Indian gems & jewelry industry. Labor Action China, www.lac.org.hk
Porter ME (1992) The competitive advantage of nations. Macmillan, London
Ray SC (2004) Data envelopment analysis: Theory and Techniques for Economics and Operations Research. Cambridge University Press
Sanyal K (1993) Paradox of competitiveness and globalization of underdevelopment. Econ Polit Wkly 28(25):1326–1330, June 19
Scott AJ (2006) The changing global geography of low-technology, labor-intensive industry: clothing, footwear and furniture. World Dev 34(9):1517:1536

Chapter 16
On Assessment of Women Empowerment at Individual Level: An Analytical Exposition

Joysankar Bhattacharya, Sarmila Banerjee and Montu Bose

Introduction

Ever since gender issues entered into the domain of policy analysis, efforts have been made to monitor the progress of interventions through two major indices suggested by UNDP, viz., the gender-related development index (GDI) and the gender empowerment measure (GEM). These two indices have been constructed to assess the extent of gender inequality at a global scale and make them comparable over time and across space. Two important methodological concerns were to arrive at the inequality adjusted human development index (HDI) by introducing the concept of equally distributed equivalent (*ede*) measures and to select the indicators to be included in the indices to represent gender biases. The inequality between men and women are considered in a dimension where the potentials of the two groups are assumed to be the same and all observed differences in achievements are ascribed to difference in social, economic, cultural, political and legal factors. If X_f and X_m represent the value of the indicator X for female and male, respectively, then the social worth of (X_f,X_m) and that of its equally distributed equivalent $(X_{\text{ede}}, X_{\text{ede}})$ would be the same and X_{ede} would be obtained as a harmonic mean of the distribution instead of the simple arithmetic mean where the former is known to be distribution sensitive and the latter is not. The GDI is obtained as the equally distributed equivalence of the HDI where three relevant

J. Bhattacharya (✉)
Central University of Bihar, Patna, India
e-mail: joysankar.bhattacharya@gmail.com

S. Banerjee
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: sarmila_b@yahoo.co.in

M. Bose
DRS Project Fellow, Department of Economics, University of Calcutta,
Kolkata 700050, India
e-mail: monbose@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_16,

indicators are life expectancy at birth to represent health status, adult literacy rate and gross enrolment rate to represent the status of education and index of income share to represent economic status. The indicators included in GEM are female and male share of parliamentary seats to represent inequalities in political participation and decision making, female and male share in administrative, professional and technical jobs to represent inequalities in economic participation and decision making and female and male estimated earned income to represent inequalities in power over economic resources. In 2010, by taking indicators from both GDI and GEM and incorporating a few more indicators of women vulnerability the gender inequality index (GII) has been constructed by UNDP where health vulnerability has been represented in terms of reproductive health (maternal mortality rate and adolescent fertility rate), empowerment has been represented in terms of parliamentary participation and (higher level) educational attainment and economic vulnerability has been assessed in terms of labour force participation rate. Thus, initially the gender sensitive indices are defined as achievement indices at a macro level and a few quantifiable indicators are selected to represent the (unobserved) differences in empowerment. However, the focus is gradually shifting towards participation measures where instead of the status of achievement the vulnerability/capability to participate in the socioeconomic sphere is enjoying greater importance.

The gender sensitised development indices are measuring the extent to which women have access to certain levels of power; by power we mean oppressive power, challenging power as well as creative power. The process of women empowerment is related to all three dimensions: exposing the oppressive power of the existing gender relations, critically challenging them, and creatively trying to shape different social relations. So, at individual level the extent of conscious participation in the decision making process to protect and enhance ones own well being would be a definite measure of empowerment. If policies and programmes are taken up to ensure more fair development through mitigating gender disparity and addressing women empowerment then evaluation of policy impact at the individual level would be difficult in the absence of any specific methodology to map any capability improvement to a gain in empowerment.

In Bhattacharya and Banerjee (2012) an attempt has been made to focus on this latent nature of empowerment, manifested through capability enhancement. A critique of the idea of considering autonomy as the sole indicator of empowerment has been presented and an attempt has been made to supplement autonomy with other dimensions like health and knowledge in shaping empowerment. But the capabilities by definition cannot be directly measured. What can be measured, however, are the *functionings* namely the achievements in each dimension. These achievements are generally identified by proper indicators reflecting the performance in the associated dimension. There could either be one indicator or, as is more often the case, a whole range of indicators available for each capability dimension. Moreover, the observed level of achievement is not only contingent on the level of intrinsic capability but it is highly contextual and gets affected by different socio-economic preconditions.

If we view empowerment latent and manifested through capability-assessment via functioning,[1] construction of an empowerment indicator requires a specific type of modelling which can incorporate the possibility of measurement error in both independent as well as dependent variables (in our case capabilities and hence empowerment). A research methodology has been developed by using multiple indicator multiple cause (MIMIC) type of structural equation modelling (SEM) modelling) that follows three major steps: (1) measuring latent capabilities through observed functionings, (2) assessing the contribution of different socio-economic factors in influencing capabilities and (3) deriving measure for empowerment determined by different levels of capability. The method has been applied on the National Family Health Survey (NFHS) third round dataset, which is a large-scale, multi-round survey conducted on a representative sample of households throughout India. As an illustration we have considered only the state of West Bengal and analysed the difference in the average level of empowerment across rural and urban areas. The rest of the paper has been organised as follows: "Latent Variable Modelling" discusses the theoretical challenges faced in modelling latent variables, "Structural Equation Modelling and MIMIC" reports the measurement techniques adopted, "Variables and Data" defines variables and discuss the NFHS-3 dataset, "Data Analysis" presents data analysis and "Concluding Observations" concludes the paper.

Latent Variable Modelling

Latent random variables are unobserved or unmeasured variables and factors. The observed indicator(s) of a latent variable contain random or systematic measurement errors even when the latent variable is free of these. Since all latent variables correspond to concepts, they are hypothetical variables though there is definite variation in their degree of abstractness. Measurement is the process by which a concept is linked to one or more latent variables, and these are linked to observed variables. Once a concept is selected or devised, the four steps in the measurement process are (1) give the meaning of the concept, (2) identify the dimensions and latent variables to represent it, (3) form measures, and (4) specify the relation between the measures and the latent variables (Bollen 1989).

The first step is accomplished by developing a theoretical definition. A theoretical definition provides the meaning of the concept, links a term to a specific concept, identifies its dimensions and the number of latent variables, and sets a standard by which to select measures. The next step in measurement is to arrive at

[1] In spite of extensive controversy in the literature related to the suitability of defining capability in terms of any observable achievement indicator, we are proposing a measurement of empowerment in terms of capability enhancement as our purpose is primarily lying in developing and illustrating a methodology where a latent variable can be measured as a function of another latent variable along with a bunch of other observable indicators.

an operational definition. The operational definition describes the procedures to follow to form measures of the latent variables that represent a concept. This operational definition or measure is appropriate to the extent that it leads to an observed variable that corresponds to the meaning assigned to a concept. But often the information is not complete and a considerable element of judgement must enter classification decisions. These and other factors lead to random and non-random errors in measures. Virtually all measures contain such errors and to formalise these types of errors a model is formed to capture the relation between the measure and latent variables.

At this point it is interesting to note that the historical emphasis in the discipline of econometrics is placed on models without measurement error in the variables and instead with stochastic "shocks" in the equations. For a classical single-equation regression model, measurement error in the dependent variable causes no particular problem because it can be subsumed within the equation's disturbance term. The presence of measurement error in the independent variable does not alter the properties of least squares estimates of regression coefficients. But the variance of the measurement error remains hopelessly entangled with that of the disturbance term. The measurement error in the dependent variable remain unaddressed since "the reader will usually be convinced of the futility of consistent parameter estimation in such instances unless repeated observations on dependent variables are available at each data point or strong a priori information be employed" (Wansbeek and Kapteyn 1984, 1813).

Structural equation modelling can deal with measurement error in both endogenous (dependent) and exogenous (independent) variables. A part of the endogenous variables is in itself unobservable (latent), but can be imperfectly measured by a range of indicator variables. The full model corresponds to a regression model where both dependent and independent variables may be measured with error. We consider a particular form of structural equation modelling, called (MIMIC) model which is useful in addressing all the three measurement issues : (1) it works within the multidimensional framework and considers a mixture of both aggregative and non-aggregative strategies,[2] (2) it maintains the simultaneous and interdependent structure by its very nature: and (3) considering capability as a latent variable, manifested through a set of observed achieved functioning and the influence of several exogenous causes affecting capabilities pays proper attention to the analytical structure of the concept of empowerment .

[2] Even if achieved functionings have been considered in a non-aggregative fashion, their contribution in constituting the latent capability leaves room for aggregation in the space of capabilities.

Structural Equation Modelling and MIMIC

Linear measurement models decomposing manifest variables into hypothesised unmeasured variables have been well-studied in econometrics (Lawley and Maxwell 1971). When the unmeasured variables are linear combinations of manifest variables, such as principal components, they are measurable, but when the unmeasured variables span a space of greater dimensionality than the space of manifest variables, as in factor analysis, they are not measurable or determinate (Williams 1978). In that case, they are appropriately called latent variables (Bentler 1980). Models that combine simultaneous equations and measurement models into interdependent multivariate linear relations have been only recently proposed and applied, but their impact on psychology and other social sciences promises to be substantial (Aigner and Goldberger 1977; Bieby and Hauser 1977; Jöreskog 1978). In these developments, no consideration has been given to measurement models where other exogenous variables that *cause* and influence the latent factor(s) may also be accommodated. The MIMIC model (Jöreskog and Goldberger 1975) presents such extensions. In this approach each of the indicators is assumed to be a component of functioning and capabilities, as latent variables, are linked to the observed indicators. In separating causes from indicators, the MIMIC approach brings more structure to bear on the problem than do the comparatively unstructured methods like scaling, factor analysis, principal components etc.

The method of scaling, as exemplified in the construction of HDI, represents an individual's functioning as a proportion of the maximum achievable functioning level. It is obtained by deducting the minimum possible score from the individual's achieved score and by dividing the difference by the range of the score. However, the method of scaling suffers from two main limitations. First, scaling can address functionings and not capabilities. Second, in scaling, the problem of measurement error is treated in an ad hoc way. If the analyst believes that a certain indicator does not appropriately represent a functioning, several indicators are combined to represent a functioning under the assumption that the measurement errors in the single indicators cancel each other out. For example, in the calculation of the (HDI), the functioning of being well-educated is represented by the arithmetic average of the adult literacy rate and the school enrolment rate. Usually, the weights are chosen in accordance to the analysts' values and do not use any statistical information contained in the data.[3]

[3] The method of scaling assumes perfect substitutability between the functionings: an individual can trade off her welfare in terms of health and education with an infinite elasticity of substitution. Fuzzy sets theory as applied in the empirical capability literature is an extension of the method of scaling. It was pioneered in this area by Chiappero Martinetti (2000) and has been applied by Lelli (2001) and Qizilbash (2002). It extends the method of scaling in two respects. First, it the indicator variable into a 0-1 interval by allowing for nonlinear projection functions such as a sigmoid function. By allowing for different weighting schemes, the use of fuzzy sets provides more flexible substitution patterns between functionings. For example the arithmetic

In both principal component analysis and factor analysis capabilities, understood as latent variables, are estimated as linear combinations of observed indicators. Instead of assigning arbitrary weights to indicators here the weights are derived from the variance of the indicators themselves. If the variance of each variable is interpreted as its information content with respect to the latent variable, the combination based on weights derived from this variance may be interpreted as statistically optimal factor analysis (or principal component analysis) and can also be applied to aggregate several functionings to a composite welfare index. In this case the weights are not arbitrarily chosen by the analyst, but determined by the data.

The structural equation models retain the strength of factor analysis and can incorporate other exogenous influences representing different socio-economic contexts. They allow us to measure latent capabilities (η_1, η_2, η_3, etc.) using observed indicators (Y_1, Y_2, Y_3, etc.) and here the weights are derived from the system by considering the entire variance-covariance matrix of measurement errors. Moreover, a MIMIC-type structural equation model can incorporate exogenous causes (X_1, X_2, X_3, etc.) that also influence these latent variables. Thus the model comprising three parts: a structural equation shows how the latent variable is estimated through the observed indicators and a measurement equation takes into account the causal link among the latent variables and the exogenous factors and finally perceived outcome as a function of estimated latent variables. To estimate the achieved level of empowerment (perceived outcome) for individual agents three capabilities have been identified as the key determinants, viz., health capability (η_1), knowledge capability (η_2) and autonomy (η_3).

Variables and Data

The National Family Health Survey (NFHS) is a large-scale, multi-round survey conducted in a representative sample of households throughout India. Since 1992–1993 three rounds of the survey have been conducted. Each NFHS round has had two specific goals:

(a) To provide essential data on health and family welfare needed for policy and programme purposes, and
(b) To provide information on important emerging health and family welfare issues (Fig. 16.1).

NFHS provides state and national information for India on fertility, infant and child mortality, the practise of family planning, maternal and child health, reproductive health, nutrition, anaemia, utilisation and quality of health and family

(Footnote 3 continued)
average is sometimes replaced by a Liontief Function (Kuklys 2005) and in this case as the assumed elasticity of substitution between the functionings is zero, no trade-off takes place between functionings.

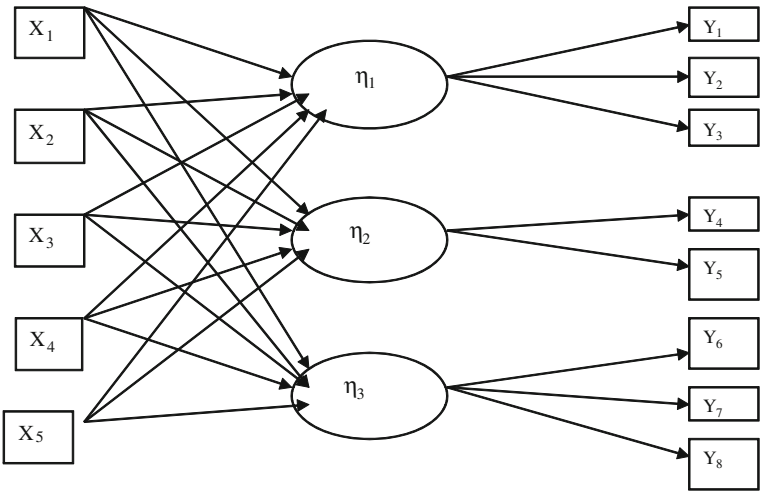


Fig. 16.1 Path diagram of MIMIC model

planning services. Fieldwork for NFHS-3 was conducted in two phases from November 2005 to August 2006. The urban and rural samples within each state were drawn separately and the sample within each state was allocated proportionally to the size of the state's urban and rural populations.[4] NFHS-3 used three types of questionnaire: Household, Women's and Men's questionnaire. Our study concentrates mainly on women and household questionnaire to capture adult women's capability-enhancement (and hence empowerment) along with their relative positioning within the household. We have 3,066 observations on Rural West Bengal, 3,357 observations on Urban West Bengal. for which complete information was available in NFHS-3 dataset. Appendix A gives a description of the variables under consideration.

Estimation of Capabilities

In this case the specification of MIMIC model considers three latent capabilities, viz., Health (η_1), Knowledge (η_2) and Autonomy (η_3) where η_1 is reflected through three observed indicators like *Food Intake* (INTAKE$[Y_1]$), *Anaemia*

[4] In each state, the rural sample was selected in two stages and in urban areas, a three-stage procedure was followed. In rural areas, with Probability Proportional to Population Size (PPS) at the first stage, Primary Sampling Units (PSUs) were selected. Here census villages are the PSUs. In the second stage, households were randomly selected within each PSU. In urban areas, in the first stage, wards were selected with PPS sampling. At the next stage, one Census Enumeration Block (CEB) was randomly selected from each sample ward. In final stage, households were selected randomly within each selected CEB.

(ANAEMIA[Y_2]) and the BMI (BMI[Y_3]), η_2 is reflected through *Educational attainment* (EDU[Y_4]), *Functional literacy* (FLIT[Y_5]) and the application of awareness/knowledge in the context of *Family planning* (CONTRALIT[Y_6]); η_3 is reflected through extent of participation in *Actual decision making* (ACTUAL[Y_7]), *Perceived freedom* (PRCVD[Y_8]) and the extent one is allowed to enjoy an extended space in terms of *Permission regarding mobility*[5] (PERMISS[Y_9]).

As has already been mentioned that besides the set of indicators (Y's) there is some intrinsic interdependence among latent capabilities (η's). Thus $Y_p = Y_p(\eta_i, \eta^{-i})$, where η^{-i} represent all η's except η_i. Thus, Y_p, the p-th indicator variable associated with η_i is influenced by both η_i and $,\eta^{-i}$ and confirms the interdependent nature of capabilities in an ideal set-up. However, MIMIC model, while successful in relating the observed functionings (Y_p's) with unobserved capabilities (η_i), fails to capture the inter-capability interdependence. Here $Y_p = Y_p(\eta_i)$. The combined effect of all unobservable η_i's would be culminated into another unobservable measure, viz., Women empowerment (EMPI). The influence of each η_i on the final outcome EMPI is likely to be contingent on different socio-economic and cultural factors. Hence, to posit the social agenda related to the enhancement of *women's empowerment* in an appropriate policy context, one needs to study the influence of such factors (X_i's) on the η_i's and the analytical frame has to be extended beyond measurement model accordingly. This will take us to the estimation of structural part where seven common causes have been identified as *marital status* of the respondent (MSTAT[X_1]), *religion* (RELGN[X_2]), *caste*(CASTE[X_3]), *relationship with the household head* (RELNHH[X_4]), *housing condition* (Hsc[X_5]),[6] *wealth index* (WINDEX[X_6]) and *age* of the respondent (AGE[X_7]), which are expected to influence the capabilities.

Estimation of Empowerment

We propose a structural equation (measurement) model to estimate the Empowerment Index, an unobserved phenomena in terms of estimated capability scores, obtained from the MIMIC model, specified above. The regression coefficients of the measurement part of the MIMIC model work as weights in constructing the estimated capability scores as weighted averages of indicators. The constituents of empowerment index are the estimated capability score of each category, Health, Knowledge and Autonomy and we consider the path diagram of a structural

[5] PERMISS is contributing negatively towards the attainment of autonomy.

[6] Housing condition is a latent variable which is measured in terms of an index constructed by extracting principal component(s) and factor scores from a group of related variables like type of wall material, type of roof material, cooking fuel type, toilet type, source of drinking water, source of lighting, etc. Appendix B

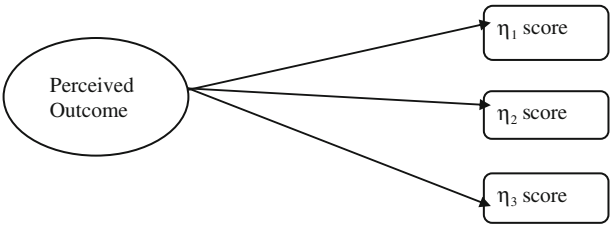


Fig. 16.2 Path diagram to estimate final outcome

equation (measurement) model to bring out the influence of each capability score on this latent empowerment index (Fig. 16.2).

Data Analysis

The analysis is reported in three parts: (a) the measurement part and (b) the structural part of the MIMIC model and finally (c) an estimation of the unobserved empowerment index in terms of estimated capability scores, so obtained, by utilising the Linear Interdependent Structural Relationship (LISREL) software to map unobservable into observables.

The Measurement Part of the MIMIC Model

The regression results are presented in Table 16.1. The regression coefficients portray the manifestation of latent, unobservable capabilities through observed indicators.[7] It is interesting to note that in rural West Bengal INTAKE and ANAEMIA play more important role in shaping health capability than in the urban parts whereas for the latter only BMI has significant statistical influence. EDU and FLIT the main constituents of the latent Knowledge capability and the relative values of the regression coefficients are always higher in the urban West Bengal compared to its rural counterpart indicating higher use-value of formal training in day-to-day urban living. PRCVD, the perceived autonomy (measured as inversely related to the incidence of domestic violence) appears to be an important determinant of Autonomy capability with higher effect in rural areas. However, it is interesting to note that permission-requirement from husband or other members of the family (PERMISS) for women to move out for economic and social activities

[7] Given the nature of available NFHS-3 data, the LISREL programme fails to complete loop for two indicators, viz., CONTRALIT and AUT. Hence, here no regression weight could be obtained and these dimensions are eventually dropped from the relevant sets of functioning.

Table 16.1 Regression coefficients of the measurement equation measurement equation

Capabilities	Functionings	West Bengal	Rural Bengal	Urban Bengal
Health	INTAKE	0.57*	1.34**	1.01
	BMI	0.42*	0.87	0.41*
	ANAEMIA	0.30*	0.36*	0.22
Knowledge	EDU	1.26**	0.51*	1.03*
	FLIT	1.25**	0.47*	1.30*
	CONTRALIT	–	–	–
Autonomy	AUT	–	–	–
	PRCVD	1.17*	1.15*	0.69*
	PERMISS	4.89**	4.85**	3.56**

** significant at $p \leq 0.05$, * significant at $p \leq 0.10$

plays the most crucial role in determining latent autonomy capability. The need to ask for permission makes one's relative position in familial hierarchy more visible and influences one's own sense of autonomy within the system.

The Structural Part of the MIMIC Model

A number of interesting patterns can be noted from Table 16.2. In both Rural and all West Bengal level Health capability decreases with AGE and Autonomy increases; however, the situation in Urban is just the reverse. With improvement in MSTATUS health capability deteriorates everywhere and autonomy improves, and the loss in terms of the former is the lowest in the Urban area whereas the gain in

Table 16.2 Regression coefficients of the structural equation

Capability	AGE	MSTATUS	RELGN	CASTE	RELNHH	WINDEX	HSC
All West Bengal							
Health	−0.24*	−0.20*	0.07	−0.34	0.05	0.72**	0.56**
Knowledge	−0.25**	−0.01	−0.28**	−0.21**	0.14**	0.29**	0.34**
Autonomy	0.06*	0.29	0.01*	−0.01	0.02**	0.08*	0.02*
Rmsea 0.08							
Rural West Bengal							
Health	−0.05*	−0.20*	0.11	−0.05	0.05	0.36**	0.07**
Knowledge	−0.26**	−0.43	−0.64*	−0.25**	0.185**	0.44**	0.80**
Autonomy	0.10*	0.25*	0.01*	0.14	0.06**	0.06*	0.15**
Rmsea 0.07							
Urban West Bengal							
Health	0.12*	−0.06*	−0.02	−0.08	0.10	0.20*	0.09*
Knowledge	0.07**	−0.07	−0.34*	−0.32**	0.61**	0.81**	0.53**
Autonomy	−0.06*	0.90*	−0.07*	0.02	−0.06*	−0.14*	−0.09*
Rmsea 0.06							

** significant at $p \leq 0.05$, * significant at $p \leq 0.10$

Table 16.3 Latent empowerment and its measurable determinants: SEM estimation

Capabilities	West Bengal		Rural		Urban	
	coefficient	t-value	coefficient	t-value	coefficient	t-value
Health	0.12	1.34	0.19	1.94*	0.09	1.80
Knowledge	0.17	2.14**	0.15	2.01**	0.22	2.23**
Autonomy	0.06	1.96*	0.04	1.90*	0.06	1.90*

** significant at $p \leq 0.05$, * significant at $p \leq 0.10$

terms of the latter is the highest there. Both RELGN and CASTE have most significant influence on knowledge capability everywhere. Finally, WINDEX and HSC, the economic status is important for all capabilities everywhere and only in Urban area this significant impact on autonomy is negative. The women from richer family in Urban West Bengal enjoy lesser autonomy.

The SEM to Estimate Empowerment

We will use the regression coefficients of the measurement part of the MIMIC model as weights to derive capability scores of each category, viz. Health, Knowledge and Autonomy and the capability scores will be weighted averages of the respective indicators. The extent of enhancement of women empowerment depends on these capability-scores and in what follows the measurement coefficients will be used to construct this Empowerment index by following the simple rule of weighted average. The regression coefficients of the measurement model of latent empowerment index are presented in Table 16.3.

In both the areas of West Bengal, rural and urban, knowledge capability score contributes more to the attainment of empowerment and health capability contributes more in the rural areas. It is interesting to note that when dimensions like Health and Knowledge along with autonomy in constituting empowerment index, the coefficients (factor loading) for health or knowledge are greater than that of autonomy. This is not to undermine the importance of autonomy, rather to supplement it with the dimensions like health and knowledge so as to get a comprehensive measure for empowerment index.[8] In the following section the paper will be concluded by showing the sensitivity of the average value of this empowerment index across groups.

[8] We demonstrated this phenomenon before in our earlier work on a primary data set (Bhattacharya and Banerjee 2012). A critique of the idea of considering autonomy as the sole indicator of empowerment was presented and an attempt was made to supplement autonomy with other dimensions like health and knowledge in shaping empowerment. The results demonstrated the fact that high autonomy along with high attainment in other capabilities definitely improves the empowerment index, but considerable empowerment-attainment may be observed even with low autonomy but with higher achievements in other capabilities and vice versa.

Table 16.4 Average value of empowerment index across socio-economic groups

Code*	Rural		Urban		Code*	Rural		Urban	
	X_{Median}	#	X_{Median}	#		X_{Median}	#	X_{Median}	#
AGE					*WINDEX*				
1	1.99	789	2.14	692	1	1.94	2,024	2.14	165
2	2.08	1,137	2.15	1,130	2	2.09	608	2.14	423
3	2.09	1,240	2.14	1,535	3	2.09	434	2.14	2,769
MSTATUS					*RELNHH*				
1	2.09	2,447	2.14	2,272	1	2.09	1,893	2.14	1,666
2	1.88	68	2.21	67	2	2.03	541	2.14	872
3	1.93	551	2.14	1,018	3	1.99	632	2.14	819
RELGN					*CASTE*				
1	2.09	2,041	2.14	2,579	1	2.05	1,920	2.14	2,601
2	2.03	1,001	2.14	709	2	2.09	801	2.14	631
3	1.96	24	2.14	69	3	1.94	345	2.14	125

* Codes are explained in Appendix A; #: number of observations

The median values of the empowerment index (EMPI) for a number of socio-economic groups across different regions have been reported in Table 16.4. It appears from the table that the average empowerment status is more stable in Urban West Bengal where a concentration around a median value of 2.14 in a 3-point scale is generally observed. However, the same does not hold for Rural West Bengal. Here sufficient sensitivity is observed across different levels of these exogenous factors. Here MSTATUS is revealing highest sensitivity followed by AGE and WINDEX; relationship with head of the household [RELNHH], religion [RELGN] and caste [CASTE] are suggesting same pattern.

Concluding Observations

This paper tries to offer a comprehensive measure for empowerment where empowerment is viewed as capability-enhancement. Three capabilities have been identified as health, knowledge and autonomy. A quantitative measure for empowerment constituted of capability scores on all these three dimensions. A particular form of structural equation modelling, called (MIMIC) model has been used to estimate capabilities and the empowerment index has been constructed as weighted average of the scores of Health, Knowledge and Autonomy. The method has been applied on NFHS-3 data collected from adult women of West Bengal and the results demonstrated the fact that the empowerment status at the individual level is fairly sensitive with respect to the variations in different contextual factors and the application of MIMIC type model helps a great deal in reducing the measurement error intrinsic to the latent nature of *empowerment* by definition.

Appendix A

Variable-description

Variable	Description	Categories	Variable Name
	Observed causes		
Age (v012)	Women of age group 15–25 yrs	1	AGE
	Women of age group 26–35 yrs	2	
	Women of age group 36–49 yrs	3	
Religion (v130)	Hindu	1	RELGN
	Muslim	2	
	Christ, Buddhist, Sikh, Jain, Others	3	
Caste (s46)	General	1	CASTE
	SC	2	
	ST, OBC, Others, don't know	3	
Relationship with head of the household (v150)	Head, wife	1	RELNHH
	Daughter, daughter-in-law, sister	2	
	Others	3	
Marital status (v501)	Married	1	MSTATUS
	Never Married	2	
	Divorced, Widowed, not living together	3	
Wealth index (v190)	Poorest, Poorer	1	WINDEX
	Middle Income	2	
	Richer, Richest	3	
	Observed causes: Housing condition (HSC) (index constructed)		
Drinking water (v113)	Piped into dwelling, piped to yard, public tap, tube well, protected well, unprotected well, river/canal/ pond etc., tanker, bottled water, others		
Toilet type (v116)	Flush to piped sewer, flush to septic tank, flush to pit, flush to somewhere else, flush, do no't know where, ventilated improved pit latrine (VIP), pit latrine with slab, pit latrine without slab, no facility, dry toilet, others		
Sources of lighting (v119)	Electricity (present/absent)		
Cooking fuel type (v161)	Electricity, LPG, biogas, kerosene, coal, charcoal, wood, straw/grass, agricultural crops, animal dung, others		
Type of wall material (v128)	No wall, cane/palm/trunks, mud, grass, bamboo with mud, plywood, unburnt brick, reused wood, cement/concrete, stone, burnt brick, cement block, wood planks, gi/metal, others		

(continued)

(continued)

Variable	Description	Categories	Variable Name
Type of roof material (vv129)	Mud/clay, sand, dung, raw wood planks, palm/ bamboo, brick, stone, vinyl, ceramic tiles, cement, carpet		
	Indicators for HEALTH		
BMI* (v445)	Varies between 12.10 to 20.00	1	BMI
	Varies between 20.00 to 30.00	2	
	Varies between 30.00 to 48.68	3	
Anaemia** (v453)	Varies between 3.3 to 5.00	1	ANAEMIA
	Varies between 5.00 to 10.00	2	
	Varies between 10.00 to 16.6	3	
Food Intake (s558a-g) Consumption of food items (Milk or Card, Pulses or beans, dark green leafy vegetables, fruits, eggs, fish, chicken or meat):	Daily	1	INTAKE
	Weekly	2	
	Occasionally, never	3	
	Indicators for Education		
Educational achievement (v149)	No education, incomplete primary	1	EDU
	Complete primary, incomplete secondary	2	
	Complete secondary, higher	3	
Ability to read (v155)	Can't read at all	1	FLIT
	Able to read only parts of sentence	2	
	Able to read whole sentence	3	
Knowledge of Contraceptive Method (v301)	Knows no method	1	CONTRALIT
	Knows only traditional method	2	
	Knows modern method	3	
	Indicators for Autonomy		
Decision taken by (v632, v739, w124, v743a) (How to spend money, contraception use, access to health care facility, money for own use)	Respondent alone	1	AUT
	Jointly with husband/ partner, other	2	
	Mainly husband/ partner	3	

(continued)

(continued)

Variable	Description	Categories	Variable Name
Domestic violence (d104, d105a-j, d109, d110a-e, d115b-y) (ever pushed shook, throw something, slapped, punched, kicked, dragged, strangled or burn, attacked with knife, gun, forced sex, twisted arm or hair)	No, often during last 12 months	1	PRCVD
	Sometimes during last 12 months	2	
	Not in last 12 months, yes, but currently a widow	3	
Permission Requirement (s824a-c) (market, other household, health facility)	alone	1	PERMISS
	With someone else only	2	
	Not at all	3	

Appendix B

Construction of Housing Condition Index

Principal component analysis method (PCA) has been used to calculate the housing condition index. Here six variables indicating one's housing condition have been used, viz., drinking water (drwtr), toilet type (toilet), lighting source (litng), fuel use (fuel), wall material (wall) and roof material (roof).

The output is as follows:

KMO and Bartlett's Test

Kaiser–Meyer–Olkin.	Measure of sampling adequacy	0.844
Bartlett's test of sphericity	Approx. Chi Square	14501.688
	df	15
	Sig.	0.000

Here, the value Kaiser–Meyer–Olkin measure of sampling adequacy is 0.844 which is far beyond the acceptable minimum of 0.6. On the other hand, Bartlett's test of Sphericity is significant here. Therefore, it is possible to conduct a PCA with the present dataset.

Total Variance Explained

Component	Initial eigen values			Extraction sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.201	53.354	53.354	3.201	53.354	53.354
2	1.009	16.811	70.166	1.009	16.811	70.166
3	0.618	10.297	80.463			
4	0.424	7.063	87.526			
5	0.391	6.518	94.044			
6	0.357	5.956	100.000			

Extraction Method Principal Component Analysis.

In the above table, the eigen values which are the variance of the components are available for all six variables. From the result it's clear that there are two regression factor score for the housing condition index. The weighted average of these two factor score forms the required housing condition index, where weights are 3.201 and 1.009, respectively.

References

Aigner DJ, Goldberger AS (1977) Latent variables in socio-economic models. North-Holland, Amsterdam

Bentler PM (1980) Multivariate analysis with latent variables: causal modelling. Annu Rev Psychol 31:419–456

Bielby WT, Hauser RM (1977) Response error in earning function for non- black males. Sociol Methods Res 6:241–280

Bhattacharya J, Banerjee S (2012) Women empowerment as multidimensional capability enhancement: an application of structural equation modelling. Poverty Public Policy 4(3):79–98

Bollen KA (1989) Structural equation with latent variables. Wiely- Interscience Publication, New York

Chiappero-Martinetti E (2000) A multidimensional assessment of well-being based on Sen's functioning approach. In: Rivista Internazionale di Scienze Social, CVIII 2, pp 207–239

Jöreskog KG (1978) Structural analysis of covariance and correlation matrices. Psychometrika 43:443–473

Jöreskog K, Goldberger A (1975) Estimation of a model with multiple indicators and multiple causes of a single latent variable. J Am Stat Assoc 70(351):631–639

Kuklys W (2005) Amartya Sen's capability approach, Theoretical insights and empirical applications. Springer, New York

Lawley DN, Maxwell AE (1971) Factor analysis as a statistical method, 2nd edn. Butterworths, London

Lelli S (2001) Factor analysis vs. fuzzy sets theory: assessing the influence of different techniques on Sen's functioning approach, public economics working paper series 121. Center for Economic Studies, Belgium

Qizilbash M (2002) Vagueness and the measurement of poverty, The economics research centre, school of economic and social studies, university of East Anglia, discussion paper N. 2000–2003

Wansbeek TJ, Kapteyn A (1984) Errors in variables: consistent adjusted least squares (CALS) estimation. Commun Stat A 13:1811–1837

Williams J (1978) A definition for the common-factor analysis model and the elimination of problems of factor score indeterminacy. Psychometrika, Springer, 43(3):293–306

Chapter 17
Sustainable Poverty Reduction: Credit for the Poor

Asis Kumar Banerjee

Introduction: Saving, Credit and Development

This chapter is a review of the research done over the past two decades or so on the role of credit supply in sustainable poverty reduction. Naturally, it is credit to the poor that will be the focus of our attention. We shall review both theoretical and empirical research. We shall also suggest possible lines of further research (which do not find mention in other reviews) and make a few policy recommendations (which, again, do not seem to appear in the existing literature) with reference to the Indian economy.

Two *limitations* of this review should be highlighted at the outset. *First*, sustainability here would mean *economic* sustainability. Obviously, poverty reduction is meaningful only if the reduction is sustainable rather than transitory. We shall, however, not enter into the discussion of the impact of poverty reduction measures on the sustainability of the *physical environment*. This should not be taken to be a denigration of the environmental question. The reason for our narrower focus is simply that, as we shall see, the analysis of the narrower question will be enough to keep our hands full. *Second*, we shall be confined to the matter of *credit* and shall not attempt to enter into other aspects of *finance*. (Credit is one type of financial service. There are others, for instance, insurance.)

We start by noting that one of the clearest lessons learnt from the theory of development relates to the importance of *investment*: economic development is primarily driven by investment.

Needless to say, it is *saving* (or *savings* i.e. the accumulated effect of previous acts of saving) which is the main source of funds for investment in an economy, leaving aside the matter of foreign aid. In most cases, however, savers themselves are not investors. Rather, producers who are in need of investible funds seek loans

A. K. Banerjee (✉)
Institute of Development Studies Kolkata, 27/D, Block DD, Salt Lake, Kolkata 700064, India
e-mail: asisbanerjee.cu@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_17, © Springer India 2013

from the savers. Again, in most cases the savers do not give loans directly to the investors. It is the *financial intermediaries* who perform the function of collecting the savings of the savers and lending them to the investors. The most important and the best known examples of financial intermediaries are, of course, banks. The crucial role played by financial intermediaries in economic development is a well-researched field in economics.

Since it is saving that fund investment, conventional approaches to economic development rightly identified the problem of increasing the rate of saving as the central issue of development. In recent decades, however, a somewhat different aspect of the matter has engaged the attention of economists. In many cases it is seen that there are lendable resources in the economy but these are not lent out for productive purposes. (For instance, in many backward economies landlords are often seen to be sitting over large amounts of accumulated or inherited wealth but refraining from lending them to producers.) Insufficient *demand* for credit cannot be the reason for this state of affairs. Indeed, even in backward economies, the credit market is, more often than not, beset with excess demand rather than excess supply.

This chapter starts by looking at the attempts to explain this paradox. While doing so we shall lump savers and financial intermediaries together and call them lenders. The transaction between lenders and borrowers is the focus of attention.

Asymmetric Information

Special Nature of Credit Markets

Attempts at explaining the paradox of idle loanable funds in the presence of excess demand have led to some very important advances in economics. It has been noticed that there are some markets which do not function in the way markets have been traditionally assumed to function in economic theory. By virtue of the work done by George Akerlof, Michael Spence and Joseph Stiglitz, the winners of the Nobel Prize in economics in 2001 we are now aware of the crucial role played by *asymmetric information* in these markets.

It is now generally recognised that asymmetric information problems are ubiquitous. However, there are markets in which these problems are particularly severe. The leading example of a market beset with informational asymmetry is the market for *credit*.

Adverse Selection

A lender is concerned with maximising the expected rate of return on loans (E, say). E depends on the (a) rate of interest r and (b) the probability p of the loan being repaid. The lender usually prefers a higher value of r to a lower one.

However, p itself may be a function of r. This may happen for *two* reasons. One of these is the following. (For the other see the discussion on "Repeated Lending Games".)

A borrower who is willing to pay high r is likely to be a bad borrower i.e. a borrower with a low probability of repayment [either because she simply does not have the intention to repay (the case of deliberate or *strategic* default) or because she would invest the loan in projects which are too risky and would go into *involuntary* default)]. Indeed, such borrowers are in a position to offer higher interest rates precisely because they perceive their repayment probability to be low. Hence, the higher is r, the lower may be p and the riskier may be the lender's total loan portfolio. This may have a negative effect on the lender's profits. For this reason, it is likely that as r increases, E increases at a lower rate and eventually falls when r becomes too high i.e. p becomes too small. In that case there would be a value of r (r^*, say) which maximises E i.e. which is optimal for the lender. The lender's *optimal supply* of loans is given by his supply curve for loans at the rate of interest r^*.

Now, the demand for and supply of loans in the market may be decreasing and increasing functions of the rate of interest respectively as usual; and there may exist a market-clearing value of r (r_0, say). It is possible, however, that $r^* \neq r_0$. Under realistic assumptions we would have $r^* < r_0$. There will be excess demand in the credit market. But the lender is not interested in increasing credit supply. Since it is r^* that maximises E, by definition it is not optimal to do so. There will be credit shortage in the market even if the lender has spare lendable resources. (See Stiglitz and Weiss (1981).

This has been called the problem of *adverse selection*. From the viewpoint of the lender the problem is one of sorting the "good" borrowers (i.e. those with high values of p) from the "bad" (i.e. those with low p). The essential problem is that p is known to the borrower but not to the lender (or, at least, the borrower can make a better guess about it than the lender). There is an *asymmetry of information*.

We can illustrate the problem with the help of a simple example used by Wydick (2008). Suppose that there are two investment projects, 1 and 2. Both have an investment cost of Re. 1 and both yield their outputs after a time lag 1 year. Project 1 is a safe project and yields a known amount of gross output (say, worth Rs. X). Assume that investment cost is the only cost of a project. The net return to the borrower with the safe project is then $X - 1 - r$. The net return to the lender is, of course, r.

Project 2 is a risky project. It yields gross output worth Rs. Y with probability p. Assume that $Y > X$. However, the project may flop. (The probability of this event is $1 - p$.) In that case the gross output is 0. The borrower then is unable to repay the loan. The net return to the borrower is, therefore, also 0. The lender, however, incurs a loss. Since Re. 1 was lent but is not returned, the net return for the lender is -1.

To ensure comparability, assume that two projects have same *expected* gross output. For Project 1, it is trivially, X. For project 2 it is $Y_p + 0(1 - p) = Y_p$. By assumption, therefore, $X = Y_p$.

Now consider a borrower who chooses between alternative investment projects on the basis of their expected net returns. The borrower's expected net return from the riskless Project 1 is $X - 1 - r$ while that from the risky Project 2 is $(Y - 1 - r)p + 0(1 - p)$. It is easily checked that since $X = Y_p$ and since p, being a probability, lies between 0 and 1, the expected net return from Project 2 is higher than that from Project 1. Therefore, the borrower prefers Project 2.

For the lender, however, the expected net return from Project 1 (which is r) is higher than that from Project 2 [which is $rp + (-1)(1 - p) = (1 + r)p - 1$], again because p is positive but less than 1.

Therefore, the lender would prefer the borrower to invest in Project 1. This borrower, however, prefers the risky Project 2. The problem is that even if there are other potential borrowers who are risk-averse and would rather choose Project 1, they will be outcompeted by this borrower if the usual market mechanism is allowed to operate. Because the borrower who prefers the risky Project 2 will be able to make higher interest offers to the lender: Since this borrower's expected net return is $(Y - 1 - r)p$, the maximum interest rate that he or she can offer, r_1 (say), is that value of r which would push down the value of this expression to zero. Thus, $r_1 = Y - 1$.

For the safe investor, however, the maximum possible interest offer, r_2 (say), is $X - 1$ (since the expected net return is $X - 1 - r$). Since $X = Y_p < Y$, $r_2 < r_1$. The safe investor will be outbidded.

In *developing economies* the adverse selection problem often takes a cruder form. In such economies systematic information on the credit history of borrowers is not available. It is, therefore, difficult for a lender to be sure about borrower's *intention to repay* a loan. Note that the nature of the problem here is somewhat different from the problem discussed in the preceding paragraphs. In fact, in developing economies it is frequently the case that the borrowers are poor people who are *risk-averse*. On the other hand, the basic problem is similar: it is difficult for a lender to judge the qualities of the loan applicants. This reduces credit supply to a level lower than under full information even when there is no dearth of lendable resources.

Moral Hazard

The adverse selection problem discussed above arises typically *before* a credit contract (whether written or unwritten) is entered into: the lender faces the problem of differentiating between good and bad borrowers. There is another channel through which asymmetry of information may work. This channel may become effective *after* a loan is given. Its effect, however, is the same as before:

an increase in r may affect the repayment probability p. The reason now is that such an increase may affect the borrower's *incentive to repay*. This problem has now come to be called the problem of *moral hazard*.

An increase in r decreases the return (to the borrower) of those projects which do succeed. From this, under realistic assumptions it can be shown that a rise in r induces the borrower to invest in the project which has a lower probability of success (resulting in a lower repayment probability p) but which brings higher return when it does succeed. For details of the proof see Stiglitz and Weiss (1981).

Again, therefore, as r rises, the expected return to the *lender* may first rise but will ultimately fall; again, there will be, for the lender, an optimal r (r^*) which may be less than the market clearing rate of interest.

Again, in a developing economy the moral hazard problem may take a cruder form. Here from the viewpoint of the lenders the problem is that, because the borrower is poor, there is the risk that he or she would divert the loan toward consumption. The basic problem, however, is again similar. The higher is r, the weaker will be the borrower's incentive to repay.

Screening and Monitoring

Adverse selection and moral hazard discourage lenders from lending and keep the investible resources of the country idle. The solution of these problems lies in *screening* and *monitoring* of the borrowers. Adverse selection can be avoided if there is a way to screen the loan applicants properly. Moral hazard can be avoided if there is a way to monitor what the borrowers are doing after getting loans.

In advanced economies the market mechanism has found a way to provide screening services to would-be lenders. There are companies which maintain systematic records of the credit histories of people (or at least of a large part of the population). Any lender can get access to the recorded credit history of a borrower for a small fee. This helps in screening out the bad borrowers and solves the problem of adverse selection (First-time borrowers without credit history face difficulties. This problem has not been fully resolved. In advanced economies, however, people buy consumer durables on loans. This information is used to screen application for business loans also. Those who do not have any previous credit history whatsoever are also sometimes given loans for short periods on an experimental basis if the amounts involved are small in proportion to their annual incomes.)

Moral hazard, as stated above, can be solved by monitoring the activities of the borrowers. This is a more difficult proposition. In case of large loans (which are typically required by companies) the lenders sometimes do make an attempt at monitoring. Sometimes there are representatives of the lenders (or lending institutions such as banks) on the boards of directors of the borrowing companies.

Collaterals

Sending representatives is not a cost effective way of monitoring unless the loan amount is large. A more dependable method has evolved over time and its use has now become the most common feature of loans. So far we have implicitly assumed that loans are unsecured. In practice, however, in most cases lenders demand *collaterals*. If the borrower fails to repay a loan, the collateral is forfeited. Care is taken to ensure that the market value of the collateral is not less than the amount of the loan.

It is obvious that use of collaterals solves the two problems of adverse selection and moral hazard at one stroke. It not only discourages the borrower from making an inefficient (or unintended) use of the loan. It also solves adverse selection: there is now no reason for a bad borrower to pretend to be a good one.

There is, however, an important qualification here. In order to offer collateral a borrower would have to prove that the he or she *owns* it. As Wydick (2008) points out, this presupposes the social institution of *property rights*. This way of solving the problems of adverse selection and moral hazard is, therefore, not available in a society in which people do not have property rights. Wydick refers to De Soto (2000) who argues that "property rights are the key that unlocks the power of capital markets" (See Wydick 2008, pp. 107–108).

Lending in Developing Economies

Difficulty of Screening and Monitoring in Developing Economies

Developing economies are in an especially disadvantageous position so far as credit problems are concerned. They do not have compiled credit histories of would-be borrowers. Hence, screening by this method is not feasible.

Monitoring by sending representatives is feasible only in case of large loans taken by companies. This is done fairly regularly.

Development, however, also requires that small and marginal producers get credit. These producers are usually poor and constitute a sizable portion of the population (if not the majority) in such economies. Since development, by definition, involves poverty reduction, it is possible to argue that putting lendable resources to productive use by providing credit to small and marginal producers is one of the central tasks of development.

The poor, however, are not in a position to provide collaterals as security. Thus, the usual way of getting around the problems of asymmetric information in credit markets is not available in developing economies.

Yet it is the poor who often need credit more desperately than the nonpoor. In developing economies, therefore, the coexistence of idle lendable resources and unsatisfied credit needs is particularly problematical.

Repeated Lending Games

How do lenders in backward economies solve their informational problems and manage to give credit even to some poor borrowers? Historically, there have been two ways of doing so. One of these can be described by the modern terminology of *repeated lending games*. The other is the phenomenon of *interlinked markets*. In this subsection we consider repeated lending games.

The credit market in such economies is actually fragmented into many local markets. In a given local market there are a very small number of lenders. Consider the frequently observed case where there is only one. In every period a borrower has to approach this particular lender if credit is needed.

The fact that a particular lender interacts with a borrower again and again in the credit market helps in tackling informational problems. Consider the lending game in which the lender first chooses between two *strategies*, 'Lend' and 'Do not lend'. If the 'Lend' strategy is adopted, then subsequently the borrower chooses between her two strategies, 'Repay' and 'Do not repay'. Suppose, for simplicity, that the borrower presents the proposal of a safe project which needs an investment of Re. 1 and brings a gross return of Rs. R. Let r be the rate of interest per rupee per period. If the lender takes the 'Do not lend' strategy, then the game ends there and both the lender and the borrower get zero net incomes. Supposing that it is the net incomes which are their concerns, we say that the *pay-off* of each *player* of the game is zero. The *pay-off vector* in which the first entry, say, is the pay-off to the lender and the second entry is that to the borrower is (0, 0). If the lender takes the 'Lend' strategy, however, the game continues. It ends when the borrower also has chosen a strategy. If 'Repay' is chosen then the borrower's net return is $R - 1 - r$ (Assume that this is positive, i.e. that $R > 1 + r$). The lender's net return is r. Hence, the pay-off vector is $(r, R - 1 - r)$. It is easily checked that if the borrower chooses 'Do not repay', then the pay-off vector will be $(-1, R)$.

Now, if this lending game is played only once, it is obvious that if the lender chooses the strategy 'Lend', then the borrower will choose the strategy 'Do not repay' because by doing so the borrower will maximise her pay-off or net return: Since r is positive, $R > R - 1 - r$.

Now note that the lender *knows* (or can infer) that if he or she chooses 'Lend', then the borrower will choose 'Do not repay' and that, therefore, the lender's pay-off from the 'Lend' strategy will ultimately be -1. Therefore, the lender will see that his or her optimal strategy is 'Do not lend' because the lender's pay-off from this strategy is 0 and $0 > -1$. Thus, the unique *solution* of this game is that the lender plays the strategy 'Do not lend'. The game ends there. The borrower does

not get the chance to choose a strategy. The solution pay-off vector is (0, 0). Nothing is invested although there are lendable funds in the hands of the lender.

The situation changes, however, if the game is played *repeatedly* period after period. Suppose that it is played an infinite number of times. The pay-off to any player (i.e. the lender or the borrower) is the discounted sum of the infinite series of the net returns of the different periods. The rate of time discount need not be the same for the lender and the borrower.

Assume that the lender has to choose one of the following two strategies:

Strategy 1 'Do not lend in any period'; and

Strategy 2 'Lend in the first period and, in all subsequent periods, lend if and only if the borrower has repaid the loan of the previous period'.

The borrower's choice, on the other hand, is assumed to be between the following two strategies;

Strategy A 'Repay in all periods'; and

Strategy B 'Do not repay in any period'.

(That these are the strategies available to the players is merely a *simplifying* assumption i.e. the results of the analysis will not change significantly if other reasonable strategies are admitted into the picture.)

Again, if the lender plays Strategy 1, then in all periods nothing is lent or invested. Hence, whatever the rates of time discount may be, the pay-off vector is (0, 0).

What happens if the lender plays strategy 2? If the borrower plays Strategy A, she repays in the first period and gets a net return of $R - 1 - r$. The lender gets a net return of r. Since the loan has been repaid in the first period, in the second period the lender lends again. Under Strategy A of the borrower, therefore, the net returns to the two players in the second period will be the same as those in the first period. This will happen period after period ad infinitum. Hence, the pay-off to the lender is:

$$r + r\{1/(1 + \alpha)\} + r\left\{1/(1 + \alpha)^2\right\} + \ldots = r\{(1 + \alpha)/\alpha\}$$

where α is the lender's rate of time preference i.e. $1/(1 + \alpha)$ is the time discount factor. α is assumed to be positive. Similarly, if the borrower's rate of time preference is $\beta > 0$, then the borrower's pay-off can be calculated. It is

$$(R - 1 - r)\{(1 + \beta)/\beta\}.$$

Since $R > 1 + r$, the last expression is positive. Hence, the vector of pay-offs of the lender and the borrower is:

$$[r\{(1 + \alpha)/\alpha\},\ (R - 1 - r)\{(1 + \beta)/\beta\}].$$

If, however, the borrower plays Strategy B, then, in the first period the lender's net return is -1 since the loan is not repaid; the borrower's net return is R.

Since the lender is playing Strategy 2, he or she refrains from lending in all subsequent periods. In all subsequent periods, therefore, both the players' net returns will be zero. In this case the vector of pay-offs is:

$$[-1, R].$$

The borrower will play Strategy A if her pay-off from that strategy exceeds that from Strategy B i.e. if $(R - 1 - r)\{(1 + \beta)/\beta\} > R$ i.e. if $\beta < [\{R/(1 + r)\} - 1\} = \mu$ (say). Thus, the borrower will always repay loans if her rate of time preference is low enough. Assume, for the time being, that this is the case.

Use, again, the fact that the lender *knows* that if she chooses Strategy 2 then the borrower will choose Strategy A i.e. that loans will always be repaid. With this knowledge the lender will choose Strategy 2 rather than Strategy 1 since $r\{(1 + \alpha)/\alpha\} > 0$ i.e. by doing so she maximises her payoff.

In all periods, therefore, credit is given and investment undertaken. In this way problems of informational asymmetry may be solved in local credit markets in backward economies.

Limitations: Repeated lending games, however, do not always explain how informational asymmetry is solved in credit markets. First of all, it depends on the assumption that there is only one lender in the picture. If, after defaulting on a loan from one lender, it is possible for a borrower to approach a second lender, the situation would be different. This particular limitation, however, may not be a major one. In local credit markets in backward economies the assumption of a monopolist lender may not always be unrealistic. Moreover, even if there are multiple lenders, their number may be small; and collusion, either implicit or explicit, seems quite likely.

There is, however, a more serious problem which does not seem to have been emphasised in the existing literature. In the argument above we assumed that the borrower has a *low* rate of time preference β. (In particular, it is less than μ.) We need to take a closer look at this assumption. It is generally realistic to assume that any person would prefer to receive a given income, Rs. X (say), in the present period rather than in some future period t i.e. X in period t should be *discounted* by a certain factor if it is to be equivalent to X in the present period. In terms of the symbols used above, the borrower considers $X\{1/(1 + \beta)^t\}$ amount of income in period t (which is something *less* than X since $\beta > 0$), to be equivalent to X in the first period.

Now, a person's rate of time preference is strongly influenced by his or her income or wealth. The poorer is the borrower, the *higher* will be β [so that the *lower* will be $1/(1 + \beta)$] i.e. the lower will be the discounted *present* value of a given X to be obtained in some future period t. Intuitively, the pressing needs for present consumption imply that a poor person is less willing to wait for the future than a rich person. Hence, the poorer is the borrower, the greater is the likelihood that the assumption that β is less than μ will be violated. Therefore, the problem of credit supply to the poor in a backward economy cannot be said to be satisfactorily solved by the repeated nature of the lending game.

Interlinked Markets

A more complete explanation of why lenders lend to the poor (at least in some cases) is given by the theory of interlinked markets (See, for instance, Bardhan (1984) and the references cited there).

The idea here is that in backward economies credit markets are often linked with other markets. This linkage provides lenders with ways of avoiding the problems posed by informational asymmetry. At the same time it increases their market power and enables them to extract a larger surplus from the borrowers.

For instance, a lender in such an economy is frequently also a landlord or a trader in agricultural commodities. A landlord money-lender lends usually to his tenant farmers (who are frequently share croppers) and farm workers. Familiarity with the would-be borrowers reduces the intensity of the adverse selection problem. More importantly, moral hazard is easier to deal with in such a setting even in the absence of collaterals. The landlord can threaten a tenant farmer with eviction or a worker with sacking in the event of a default. Moreover, in the case of genuine difficulties faced by the borrowers due to negative income shocks the landlord has other ways of extracting his dues, for instance, by reducing the wage payment to the worker or the share-cropper's share of the crop. It has also been shown that in such a setting the lender enjoys far greater market power than in a credit market without such linkage with other markets. They are able to extract a greater surplus from the borrowers by charging a higher *effective* interest rate. The case of the lender who is a trader in agricultural commodities is essentially similar.

In all such cases repeated interaction between the lender and the borrower and its consequences discussed in "Repeated Lending Games" also play a role. The *difference* here is that even if the borrower has a high rate of time preference, there is no way for him or her to escape the repayment obligation.

The theory of interlinked markets explains why credit is at all given to the poor. The same theory, however, also makes it clear why the poor remain poor and cannot use the credit to work their way out of poverty. The entire potential surplus in the hands of the borrower is extracted by the lender in the form of interest payments by using the enhanced market power bestowed by market inter-linkage.

Is there a way to solve the credit problem of the poor so that poor economies can come out of poverty by undertaking productive investment? This is the question we take up next.

Attempts at Self Reliance: ROSCAs

There is, in fact, a long history of the poor trying to rely on their own small savings to build up an investible pool which the neediest or the most eager among them can put to use. The type of institution that they have tried to build for this purpose has now come to be known as Rotating Saving and Credit Association (ROSCA).

Under different names and with small variations in matters of detail, ROSCAs have been operating for a long time in many countries of the world. In India ROSCAs are known as *chit funds*.

The idea of a chit fund is simple. A closely knit group of individuals consisting of extended family members and close friends agree to make contributions at regular intervals (usually monthly or weekly) into a kitty. Over a specified period of time (called *duration*) the group holds a series of meetings at regular intervals. Over this period it meets exactly as many times as there are members. At each meeting the contributions are made and the kitty is given to the member who makes the highest bid for it. The bid amount is divided equally among all the other members. It is called *dividend.* (Usually the dividend is given in the form of a discount applied on the future contributions of the other members.) A person winning the kitty at a meeting is not allowed to bid at any subsequent meeting. (Thus, any given member is sure to win the kitty at *some* meeting of the group.) But everybody must attend all meetings and make the specified contribution irrespective of whether he or she wins the kitty or not.

The chit fund is an ingenious way of solving the credit problem of the poor. This is particularly true of those members who win the kitty in the early rounds. Consider, for instance, a ROSCA of one-year duration set up by a group of 12 persons, each member agreeing to make a monthly contribution of Rs. 1,000 at each of the 12-monthly meetings that are to be held. Consider now the person who wins the kitty (of Rs. 12,000) at the very first meeting. Over the 1 year period he or she would make a total contribution of Rs. 12,000 which is the same as the value of the kitty. But the kitty is won now rather than later. In effect, this member obtains a credit of Rs. 12,000 at the beginning of the year. It is as if the member repays the *principal* amount of Rs. 12,000 by his or her own monthly contributions over the 1-year period. The bid amount which is paid additionally is essentially the *interest*.

It is true that those who win the kitty towards the end of the duration do not gain too much from it. For instance, the last member who gets the kitty has already paid up Rs, 12,000 by way of monthly contributions. Hence, the kitty does not bring a net gain on this count. Even for this person, however, there is a net earning by way of dividends obtained from bids by the early winners. Thus, even the last winner duly gets 'interest' on his or her 'savings'.

The two most commendable features of chit funds are the following: (1) The interest payments accrue to the members themselves. They do not go to money-lenders charging usurious rates of interest. One of the most important ways of siphoning off surplus from the hands of the poor (and thus hindering development of the economy) is rendered ineffective. (2) Who gets the greater advantage from the scheme by winning the kitty in an early round is decided by *bidding*. It seems reasonable to conclude that the early winners are precisely those members who need the money most urgently or can use it most productively. It seems to be a fair game.

[There is a type of ROSCA in which the winner is decided by a random mechanism, for example, a lottery, rather than by bidding. If the group is

homogeneous i.e. the need for the loan or the ability to use it productively is evenly distributed among the members, then this may be a more efficient allocation mechanism. For this and other issues relating to ROSCAs, see Besley et al. (1993) and Klonner (2003, 2008).]

It should be noted that the crucial point on which the success of a chit fund hinges is the strong social connection among the members of the group. This is what keeps adverse selection and moral hazard at bay. In the absence of this social cohesion those who win the kitties in the early rounds will simply *defect* i.e. refrain from making further contributions. The whole scheme will then break down. That this does *not* usually happen has been a surprising revelation for economists analysing the working of indigenous financial institutions. This has also been the starting point of the much too well-known *microcredit* movement, a topic to which we now turn.

Microcredit Movement: The First Phase

The Beginning

The pioneering role of Professor Muhammad Yunus and the Grameen Bank founded by him in Bangladesh in the microcredit movement is now well-known [Strictly speaking, however, two other microcredit experiments in Bangladesh (BRAC and ASA), the government-backed microcredit programme of Bank Rakyat Indonesia and the credit programme devised by the Self-employed Women's Association (SEWA), a trade union founded by Ela Bhatt in India, preceded the Grameen Bank].

Yunus concentrated on the task of reaching small loans to the very poor. He noted that repayment rates for such loans was surprisingly high, especially in the case of women borrowers and that the social network of the borrowers played an essential role in repayment. He laid stress on the formation of 'solidarity circles' or *kendras*. These were groups of a minimum of five women. Members of a group would help each other in repaying loans. Although 'joint liability' (i.e. the idea that if any member defaults, then all members of the group will be considered to be defaulters) was not formally required, it gradually came to be the implicitly agreed arrangement.

Although Yunus began experimenting with microcredit in the 1970s, the Grameen Bank was established formally in 1983 as an NGO. 40 % of its authorised capital base was contributed by its members, 40 % by the government and the rest by two state banks. It was owned by the members of the NGO but managed by Yunus and his close associates. The Bank flourished from the very beginning. Within a few years the members became the majority owners of the Bank.

Since then the model has been replicated with minor variations in a large number of countries of the world.

In course of time the lending policies employed by Grameen Bank and similar organisations underwent profound changes. In this Section, however, we briefly review the policies followed in the first phase of the microcredit movement. Roughly, this phase lasted till the end of the twentieth century or the early years of the twenty-first.

Group Lending with Joint Liability

The main feature of the lending policies of this phase was *group lending with joint liability*. Borrowers had to approach the lending institution in a group; individual loan applications were not entertained. Moreover, the borrowing group had joint liability. In the Grameen Bank, as noted above, this was an informally agreed arrangement. In many other cases the requirement was explicitly laid down.

Economic theorists analysed how this characteristic of the lending mechanism solved the problems of informational asymmetry. It so happened that the moral hazard problem was discussed first. Stiglitz (1990) pointed out that while traditional banks find it costly to monitor the use of borrowed capital, group lending with joint liability induce the members of a group to monitor each other's activities since a default by one member will disqualify the whole group from getting future loans (Note that the 'repeated lending' type of scenario is implicit in the argument). Rashid and Townsend (1994) noted that under joint liability a person who had to default due to reasons beyond her control would be helped out by the other members of the group in their own interests. Thus, joint liability creates a self-insurance against default. Besley and Coate (1995) laid stress on the strong social cohesion that is often seen to exist in traditional societies. A defaulter is socially ostracised because she puts the whole group to disrepute and inconvenience. All these considerations significantly mitigated the moral hazard problem and explained the high repayment rate of microloans.

Ghatak (1999, 2000) took up the question whether or to what extent joint liability solved adverse selection (i.e. the problem of deciding how to discriminate between good and bad borrowers *before* a loan has been given out). He argued that while the lender may not know the quality of the borrowers, there are situations where the borrowers are well-known to each other. If individuals are free to choose which group (if any) they would join, it is often the case that a good borrower will join other good borrowers and the group will keep bad borrowers at bay. Moreover, if a person is known to be a bad borrower, then it is in nobody's interest (not even in that of another bad borrower) to team up with her. As a result, bad borrowers find it difficult to form groups. All this implies that group lending can help in solving adverse selection: the very fact that a group has been formed signals that it is a group of good borrowers. The lender can then offer efficient credit contracts to the borrowers. Peer selection is the focus of the analysis here.

It may be mentioned here that in the theoretical literature there is a contrary view also. Sadoulet (1999) argues that borrowers may prefer to form

heterogeneous groups (containing both good and bad borrowers) in order to exploit the possibilities of mutual insurance. Empirical evidence, however, favours the hypothesis of homogeneous groups (See, for instance, Ahlin 2009).

Laffont (2003) emphasised that the crucial question here is whether a borrower's true type is revealed. He showed that in general some revelation mechanisms are superior to group lending. However, when *collusion* between borrowers is allowed, group lending is not only optimal within a class of simple revelation mechanism but also retains its usefulness in more general cases.

For a more detailed survey of the theory of group lending with joint liability see Ghatak and Guinnane (1999) and de Aghion and Morduch (2005).

These theoretical arguments have also been put to empirical tests. Among the traditional type of tests see Wydick (1999) who found in the context of Guatemala that peer monitoring a la Stiglitz did have a positive effect on repayment rates while the effect of social ties was ambiguous (because these sometimes work in the opposite direction i.e. people may feel embarrassed to press a close friend or relative for repayment).

Efforts have also been made to harness evidence from experimental games. In these approach potential borrowers play microfinance games in controlled experiments and their strategic choices regarding repayments are observed under various specifications. (See, for instance, Cassar et al. (2007) and Gine et al. (2010).) Not surprisingly, the evidence gathered from such experimental games seems to suggest that homogeneity of the borrowing groups, peer monitoring, personal trust and dynamic incentives (i.e. repeated games-type of considerations) all contribute toward explaining the high repayment rates of microloans.

In an innovative paper Bhattacharya et al. (2008) introduced the concept of group lending with *joint benefit* (rather than joint liability) under which repayment is rewarded with a benefit to the group (in the form of creation of social capital for the benefit of the group). It was shown that the intensity of the problem of moral hazard can be considerably mitigated by this system even in the absence of peer monitoring. The suggestion is yet to be taken up by microfinance practitioners and is, therefore yet to be empirically investigated.

Microcredit Movement: The Second Phase

Group Lending Without Joint Liability

As already noted, riding on the high repayment rates, the microcredit organisations started making rapid progress in business from the very beginning. In the first phase reviewed above most of these organisations were non-profit NGOs. Their success, however, soon attracted the attention of the more conventional profit-seeking companies. An added attraction for them was fact that, unlike conventional bank credit, microcredit was largely outside the control of central banks.

To this was added official patronage to plans for expansion of the microcredit movement. Since small loans to the poor were at the heart of the matter this movement was considered to be the gateway to global poverty eradication. The Microfinance Summit held in 1997 under the auspices of the Clinton government in the U.S.A, set the goal of reaching 100 million of the poor households of the world by 2005. This goal was, in fact, surpassed by the target date. Private philanthropic organisations as well as the public at large also came forward to help the expansion programme.

From the viewpoint of rapid expansion of microcredit, however, one of the key features of microcredit contracts in the first phase viz. *joint liability* was proving to be a hindrance. Although repayment rate was high, there was always a small percentage of default. Under joint liability the effect of this small percentage was *magnified* many times because when even one member of a group defaulted, the entire group was debarred from future loans. The expansion programme suffered.

Mainly under pressure from the large profit-seeking companies which were now eager to enter the field, joint liability was gradually disposed of. Fischer and Ghatak (2011) seem to understate the matter when they remark, "…..it (joint liability) is no longer the sole focus". In fact, in most of the countries of the world, it is no longer a focus. [India is an exception. See "The Indian Case" below.] Even the flagship institution of the microcredit movement, the Grameen Bank, launched the "Grameen II" project, a commercialised version of Grameen Bank. Joint liability which was so far the implicitly agreed arrangement though not explicitly spelled out was now explicitly banned. Lenders still insisted, however, that borrowers should form groups. Regular group meetings continued to be held in the presence of the lender's representatives. *Group lending without joint liability* was the new norm.

Many of the microcredit organisations that entered the field also dealt with other financial products. The older organisations too gradually expanded into these other areas. We shall refer to all of these organisations as microfinance institutions (MFIs) although we shall continue to concentrate on their micro*credit* operations.

Transaction Costs, Frequent Repayments and Sequential Lending

Transaction costs. How did the MFIs flourish even in the absence of joint liability? One explanation of the continued success is the traditional one in terms of transaction costs. Since group lending was retained, the usual difficulty faced by conventional banks in giving small loans was avoided. The difficulty is that the time and the administrative costs of assessing a loan application and maintaining a loan account are the same for a small loan as for a large loan. Per unit of these costs, however, the profit generated from a small loan is by far smaller than that from a large loan. Because of the high transaction costs per unit of loans banks are

often seen to be reluctant to give small loans. Group lending solves this problem for the MFIs by exploiting economies of scale. The average transaction cost of dealing with a borrower is much smaller when borrowers come in groups than when they come one by one.

Frequent repayments. The newer phase of the microcredit movement has also used newer mechanisms. Economists have also investigated how these have helped the lending institutions to ensure high repayment rates. One of the new features of microcredit is that the borrowers are now required to repay the loans in regular *instalments* (beginning soon after the loan is given) rather than at one go at the end of the specified period of the loan.

The advantage of this system of repayment has been explained in different ways. One common-sense explanation is that it inculcates "fiscal discipline" among the borrowers. Economists have also come up with more innovative explanations. For instance, Jain and Mansuri (2003) analyse the relationship of MFIs with the informal money-lenders. These informal lenders have been in the picture since time immemorial and know the borrowers intimately. Now, since an MFI starts collecting the repayment instalments shortly after giving a loan (i.e. *before* the borrower's investment project starts yielding output), the borrower typically has to take loans from the informal lenders in order to pay the instalments. In such cases the informational money-lender has a stake in ensuring that the borrower does not invest the microcredit in too risky a project. It is easier for the better-informed informal lenders to do the necessary monitoring than for the MFIs. In this way the system of frequent repayments enable the MFIs to mitigate the moral hazard problem by indirectly co-opting the informal lenders in the process.

Rai and Sjostrom (2004) point out another advantage of frequent repayments for the MFIs. The repayments are collected by their representatives at the group meetings. More frequent repayments mean more frequent meetings. Frequent meetings make it easier for the lender to extract information about the implementation of the projects even in the absence of joint liability. For instance, a borrower is often asked questions not only about the progress of her own individual project but also about that of the projects of others. In this way more frequent meetings result in better monitoring.

Fischer and Ghatak (2009) propose a theory which seems to be related to the "fiscal discipline" type of considerations. Assume that borrowers are present-biased (i.e. that they have a high rate of time preference). The immediate gain to a borrower from defaulting when a *large* repayment becomes due is very high. However, when the payments are in instalments, the burden of repayment at any time is smaller and thus less subject to the temptation to use it for other purposes such as current consumption. Although the analysis is actually more complex because of the presence of various other considerations which pull in different directions, the authors showed that when borrowers have *quasi-hyperbolic utility functions*, on balance frequent repayments make the scheduled repayments incentive-compatible for the borrowers. Moreover, the authors showed how an *optimal frequency* of repayments can be calculated by balancing the benefits

against the costs that frequent repayments impose on the borrowers (for instance, the opportunity cost of attending group meetings frequently) and the lenders (for instance, more expenditure on fees paid by the MFIs to their representatives for arranging frequent meetings).

Sequential lending. Yet another credit mechanism used by the MFIs is sequential lending (See, for instance, Roy Chowdhury 2005; Aniket 2009). If, for instance, there are two borrowers in a group, one of them is selected at random and given a loan first. The other borrower gets her loan only when the first borrower has repaid her loan. This would induce the second borrower to monitor the first borrower in her own interests (When the first borrower has repaid her loan and the second borrow has got hers, what would induce the second borrower to repay on time? In order to solve this problem, it is sometimes proposed that a part of the first borrower's income from her project would be withheld from her as a kind of collateral. It will be returned to her only when the second borrower has repaid her loan).

Empirical Investigation

Empirical researchers investigated whether the absence of joint liability had a negative effect on repayment rates and whether more frequent repayments had a positive effect. Gine and Karlan (2009) took up the former question. They carried out randomised control trials. The finding was that repayment behaviour did not differ significantly across borrowers for a large bank in Philippine when they were randomly assigned to either the joint liability or the individual liability type of loans.

While the paper is an important application of the randomised control approach, the basic question calls for further investigation. In this particular experiment all the borrowers who were given random assignments in this way were already borrowers under the joint liability system before the experiment was carried out. This means, as per the analysis in Ghatak (1999) discussed above, that they were "good" borrowers to start with. No wonder, therefore, that the presence or absence of joint liability did not make much difference for them. They would repay their loans in any case.

The empirical verdict on the efficacy of frequent repayments is also ambiguous. While microfinance practitioners seem to believe that more frequent instalments improve the loan repayment records, in practice the MFIs have had to revise the frequencies upwards and downwards repeatedly in view of the observed effects on defaults : more frequent repayment requirements sometimes increases the rate of default and sometimes decreases it. The few academic investigations that have been carried out also point toward ambiguous results. For instance, McIntosh (2008) reports that in his study borrowers of an already ongoing microcredit programme were given the choice to move over to biweekly repayments from the existing system of weekly payments. However, those that made the move recorded

slightly *better* repayment performance, contrary to expectation. Field and Pande (2008) carried out a randomised experiment by randomly assigning the borrowers of an MFI in India to weekly or monthly repayment schedules and found no evidence of a significant difference in repayment performance. Accumulation of further empirical evidence is necessary before deciding whether these findings are isolated cases or indicative of a more pervasive truth.

The Indian Case

The Banks-SHG Linkage Model and the Rise of the NBFC–MFIs

So far we have not mentioned the particular form that the microcredit movement has taken in India. This has been called the Banks-SHG linkage programme. The National Bank for Agricultural and Rural Development (NABARD) (which was set up in 1982 as an arm of the Reserve Bank of India, India's central bank, and replaced the Agricultural Refinance and Development Corporation) came up with this programme in 1992. Since then the scheme has been modified in several ways from time to time.

A self-help group (SHG) is a group of about 15 persons from a homogeneous social background usually formed under the initiative of a self-help promoting institution (SHPI). (An SHPI is usually an NGO though it is not required to be so.) With help and guidance from the SHPI, an SHG initially pools the savings of its members and lends them among its members. After it has familiarised itself with this basic activity and its accounting aspects, it becomes eligible to obtain loans from banks. (NABARD provides refinance to the banks for the purpose.) The loans are given to the members of the SHG and not to the SHG as such. Yet joint liability is enforced.

In most cases the concerned SHPI does not play a role in this transaction between the bank and the SHG. In some cases it plays a facilitating role. Only in a very small percentage of cases (about 6 %) is the bank lending actually done through the SHPI.

The major feature of this model of microcredit which sets it apart from the models currently prevailing in most other countries (and reviewed above in the Section on "Microcredit Movement: The Second Phase") is its continued reliance on joint liability. Since its initiation the Banks-SHG linkage programme has made rapid progress. While in 1993–1994 the number of SHGs linked to banks was 365 and the amount of new credit given by banks under this scheme was Rs. 0.36 crore, in 2004–2005 the corresponding figures were 5,39,365 and Rs. 2994.26 crore, respectively.

(Even in the 1990s the SHGs linked with the conventional banks were the dominant but not the only type of borrower groups in India. There have been some Grameen-Bank-type experiments. There is also the very successful SEWA Bank, a

bank founded by the trade union SEWA which is owned and operated by the members of the union and lays special emphasis on microloans.)

In the first decade of the new century, however, non-banking financial companies which are microfinance institutions (NBFC–MFIs or, simply, NBFCs) have appeared on the scene. Typically, these are large organisations with relatively low operational costs per unit of credit. They also lend to groups. So far these groups have also operated under joint liability and have been called joint liability groups (JLGs) to distinguish them from SHGs linked to the banks. Many of these NBFCs have experimented with frequent repayment schedules even in the presence of joint liability. The world-wide trend of moving away from joint liability, however, has also started being observed in India in case of the NBFCs.

Apart from the deposits garnered by these NBFCs in their roles as financial companies, loans from the conventional banks have been the major source of funds for them. (Indeed, for reasons related to transaction costs banks often show greater enthusiasm in lending in lumpy amounts to the NBFC–MFIs than in giving microloans to SHG members.) There has also been a significant increase in the inflow of private equity funds after some of these companies became large enough to be listed on the stock exchanges.

According to the Malegan Committee Report (see Reserve Bank of India 2011) the Banks-SHG linkage currently accounts for 58 % of the annual amount of microcredit given in India and the NBFCs account for 34 % with only 8 % coming from other microcredit organisations. The relative importance of the NBFCs is set to rise further in the future and may ultimately exceed that of the of the SHG linkage model.

Theoretical Research

There does not seem to have been much theoretical research motivated by the Indian microcredit scenario. One notable exception is Roy Chowdhury and Roy (2009) who consider a public–private partnership model of microcredit. NGOs help banks to channelise credit to the poor (i.e. the SHGs) both in borrower selection and in project implementation. It is argued that a distortion may arise because the private partner (i.e. the NGO) is a *motivated* agent. It is shown that, in the typical case where a borrower's project is neither too productive nor too unproductive, reducing such distortion requires delinking borrower selection from project implementation and involving the NGO in borrower selection only.

There is another line of research which investigates the effects of an increase in the number of microcredit providers on credit discipline. Since with the advent of the NBFC–MFIs there has been an increase in the degree of competitiveness in the market for microcredit in India, this type of research may have some relevance for the Indian case, at least indirectly. The point of interest in this type of research is that while in standard economic theory increased competition among producers usually benefits the consumer, this may not happen in the credit market. Consider,

for instance, the case of dynamic incentives for the borrowers to repay loans. A lender punishes defaulters (by denying them credit in the next period) and rewards those who repay on time. This mechanism will work only if the lender is a monopolist. If there are other lenders in the picture, a defaulting borrower can always approach one of them. This may discourage lending and, therefore, hurt the interests of the borrowers (See McIntosh and Wydick 2005). Ahlin and Jiang (2008) also explored the implications of multiple lenders.

Has Microcredit Led to Sustainable Poverty Reduction?

As we have noted, there has been a massive expansion of microcredit over the recent decades in almost all countries of the world. However, despite the great enthusiasm generated over microcredit, whether it has led to sustained poverty reduction is doubtful. Note, first, that there are *two* questions here: (1) Has microcredit reduced poverty? (2) Assuming the answer to the first question is positive, is the reduction sustainable?

Unfortunately, the answer to the first question is ambiguous. The literature, therefore, has not even begun to consider the second question seriously. The results of the assessment of the poverty alleviation effect of microcredit have been modest at best. While Khandker (1998) and Pitt and Khandker (1998) found significant effect of the Grameen Bank in Bangladesh, Morduch (1998) disputed the claim on the basis of the same data! Moreover, Khandker (2005) used what he considered to be *better* data and revised his earlier position on the question, now reporting that the effect had *not* been very strong. Coleman (1999) carried out a field experiment in Thailand in which microcredit was given to some villages but withheld from some others temporarily. The impacts on those who were chosen to receive credit first were, again, rather modest.

A possible reason for the limited impact of microcredit on production and income generation may lie in the fact that the poor are often compelled by circumstances to divert the loans to finance current consumption. Both Rahman (1999) and the Goldin Institute (2007) conclude that, in contrast to the purposes stated on chapter, most of the microloans are, in practice, used for coping with consumption needs. The Indian experience has not been too different. Some studies estimate that between one fifth and one-third of total microcredit in India goes to support productive activities. The rest (i.e. the major part) is used up for other purposes. [See Bateman (2010, p.29).]

Collins et al. (2009) document how microcredit has helped the poor in various countries to *cope with poverty* by smoothing the time pattern of consumption. While this can be counted as a net benefit for the poor, it must be noted that this is a far cry from the originally declared objective of the microcredit movement. The idea was to lift people out of poverty rather than merely to help them to cope with it.

Randomised Control

We have seen above that randomised designs have been used to explore the impacts of particular microcredit mechanisms (such as group lending, frequent repayments etc.). Until recently, however, there was no large-scale randomised *evaluation* of what happens to *poverty* and other economic and social variables if microcredit is introduced in a new market. An important beginning has been made in Banerjee et al. (2010). Half of 104 slums in Hyderabad, India were randomly chosen for opening of an MFI branch while the remainder were not. 15–18 months later, the impacts on the various indicators were observed. However, again the observed impacts were either zero or only modestly positive. There was no impact on such social indicators as health, education or women's decision-making powers. So far as the crucial economic variable viz. the average value of monthly expenditure per capita is concerned, there was, again, no significant effect of access to microcredit. There was considerable variation between households. On the *positive* side, expenditure on durable goods increased in the treated areas. The number of new enterprises has also increased by about one-third.

Longevity of Enterprises

As we said above, the question of sustainability of the impact of microcredit is yet to be taken up for systematic study. One aspect of this question relates to the *longevity* of the new enterprises that are set up with the help of microcredit. The little empirical evidence that we have points towards high rates of failure of these enterprises [See, for instance, Bowen et al. (2009) and Bateman (2010, pp. 74–77)]. While there are inter-country differences, it seems generally to be the case that a microenterprise formed on the basis of a microloan has only a 50 % probability of survival after 1 year. For survival after 3 years the probability is drastically lower.

Before leaving this Section, it should be noted that the question of interest is here *not* whether poverty has decreased over, say, the past three decades. (For India, there is unambiguous evidence that it has decreased, whatever may be the particular poverty index used.) Rather, it is whether it has decreased significantly and sustainably *as a result of the expansion of microcredit over this period.* Available data do not yield a convincingly affirmative answer to this question. As is well-known, some economists are of the opinion that the poverty reduction that has actually taken place has been due, rather, to a faster rate of economic growth.

Lines of Further Research

In this short Section we suggest some lines along which further research on microcredit can take place. A number of suggestions are available in other surveys (for instance, Fischer and Ghatak 2011). We do not repeat them here. However, we offer a few others. Needless to say, the list is not exhaustive.

So far as *theoretical* research is concerned, a remaining task is to include in our models some of the important features of reality which have so far been neglected. One such feature is the presence of informal money-lenders in the picture. Jain and Mansuri (2003) has so far been one of the few attempts to take it into account. Much more remains to be done, both in the analysis of specific microcredit mechanisms and in the evaluation of the impact of microcredit on poverty. In the context of countries like India another such feature of the reality is the presence of the SHG sector. Various aspects of the competition for clients between the SHGs and the NBFC-MFIs and the consequences for repayment rates and the poverty level in the economy seem to be fascinating topics for further research. For simplicity the informal moneylenders may be ignored to start with. But eventually they should enter the analysis.

On the *empirical* side we mention the following points: (i) One obvious area of further study is large-scale randomised control evaluation along the lines of Banerjee et. al. (2010). It is important to know whether or how far the results of that study replicate for other countries or for other regions of India. (ii) It is also important to gather more insights regarding repayment behaviour from experimental games incorporating further features of the ground reality (for instance, those mentioned in the preceding paragraph). (iii) In case of frequent repayments there may be a tendency for borrowers to select projects which yield returns at shorter intervals rather than projects that have longer gestation lags but yields higher rates of return. Whether or how far this has happened or tends to happen is important to know. Again, both laboratory experimental games and field experiments involving randomised control can contribute to our knowledge in this regard.

Conclusion: Limitations of Microcredit and Policy Suggestions

In this Section we conclude the discussion. In doing so we shall speak mainly from the perspective of the Indian economy.

As noted before, the high failure rate of the enterprises formed on the basis of microloans is one of the major factors that prevent microcredit from reducing poverty on a sustainable basis. The high failure rate is, in turn, to be explained by the inability of microcredit to permit the borrowers to ‘scale up’ their operations by ploughing back profits into production. The size of the loans permits the

purchase of only the most primitive instruments. As a result, the scale of production is suboptimal and the average cost of production is too high. Even when the borrowed funds are *not* diverted toward consumption or other uses, the generated income leaves too little investible surplus. So far as the goal of *sustainable* poverty reduction is concerned, this is the crux of the matter. Diversion toward consumption only aggravates the issue (or is itself caused by this more basic problem) [Bateman (2010) mentions both the high failure rates and the scale problem but does not emphasise the causal link between the two]. In this connection we offer the following policy suggestions.

Conventional Bank Credit

One way of solving the scale problem is through a reorientation of the conventional bank credit (rather than through microcredit). In many countries including India it is the small and medium enterprises (SMEs), neither the large nor the very small ('micro' or 'tiny') enterprises, which constitute the most dynamic sector of the economy both in terms of employment generation and in terms of productivity-increasing innovation.

Unfortunately, India suffers from the problem of the 'missing middle'. (See Karnarni 2007). In comparison to many other countries of the world the SME sector in India is insignificant. A complete analysis of the causes behind this state of affairs is beyond the scope of this survey. However, it stands to reason that credit availability may have played a part in it. Data in the *Basic Statistical Returns of Scheduled Commercial Banks* for various years (available on the Reserve Bank of India's website) indicate that while the relative importance of *small industries* in total bank credit in India increased between 1976 and 1990–1991, it has decreased steadily ever since then. It is true that this observation gives only a partial picture (since it refers only to the 'small' rather than to the 'small and medium' sector and since these data include credit to industries only i.e. they exclude services). Yet it indicates the nature of the problem. In particular, it indicates that micro-entrepreneurs may be enabled to graduate into profitable scales of output if 'small' enterprises (which constitute the next size category) have better access to credit.

Joint Enterprises

Some policy suggestions can also be given from within the realm of microcredit. The microcredit movement has demonstrated that group lending can create socially cohesive collections of individuals whose members are willing to support each other in productive enterprise. [Note also that group lending actually creates 'new social capital' of this type rather than just using preexisting social

connections. See, for instance, Feigenberg et al. (2010).] It can be argued that this is the single most important positive development associated with the microcredit movement. Needless to say, the trend is stronger if there is group lending *with joint liability* (as is the case in India).

We submit that it is a short (and natural) step from here to the setting up of joint enterprises by the group members. This may be a viable solution to the scale problem *without* taking the focus away from microcredit. The precise organisational form of such enterprises (in particular, whether these would be cooperative enterprises) is not the issue here. It may be noted, however, that since the ownership will be limited to the group members, such enterprises will be immune to the problem of 'elite capture' which has been the bane of the cooperative movement and has been one of the major causes of its failure in India.

A similar (though not the same) experiment is currently under way in isolated parts of India. In Kerala, for instance, encouragement is given to the formation of neighbourhood help groups (NHGs) which are then advised to formulate *sustainable* projects which are in conformity with *local development plans*. Formation of cooperatives is encouraged because the soft loans of the cooperative banks are thereby made available. But it is *not* a requirement. What is proposed here is an extension of this experiment by bringing microcredit into the picture. The scale effect is likely to overcompensate for the somewhat higher interest rates on microcredit than those on cooperative loans.

Similar experiments can also be done by members of ROSCAs, especially of those with long histories.

SEWA Bank Type of Experiments

One of the many recommendations in the Report of the Raghuram Rajan Committee on financial sector reforms in India (see Planning Commission 2009) deserves scrutiny in the present context. The Committee suggested the establishment of small banks for the purpose of dealing with small borrowers such as micro-entrepreneurs. These banks are better-placed than large banks headquartered in large cities so far as innovative utilisation of local resources and local information is concerned. The reason that this logic has not so far worked in India (as evidenced by the dismal track records of small banks) is that these banks have not been given the opportunity to set their own operational rules (for instance, recruitment rules or staff salary scales). Thus, costs which can only be borne by large banks have been forcibly imposed on them.

In the present context a specific type of small banks may be accorded special importance. This type is exemplified by the SEWA Bank but need not be confined to banks set up by trade unions (such as SEWA). The SHGs that are formed under the Banks-SHG linkage programme of NABARD often form their federations. These federations seem to be natural candidates for starting local banking

operations on lines similar to those followed by the SEWA Bank. SHGs which are not parts of any federation may also be encouraged to come together for this purpose.

To sum up, microcredit may have benefitted the poor to an extent by increasing credit supply (by mitigating the problems of informational asymmetry in the credit market in various ways) and by helping them to cope with poverty. However, how far the original objective of reducing poverty on a *sustainable* basis has been achieved by this method is open to doubt. Indeed, so far as India is concerned, the fact that those regions that have witnessed the maximum number of farmer suicides in recent years also happen to be the regions which have been almost saturated—"carpet bombed" is a phrase that has been used—with microcredit points toward its failure on this front. Yet the fundamental (and novel) feature of the microcredit movement (viz. harnessing and strengthening the *social cohesion* among the poor, especially among poor *women*) does not seem to have been exploited fully for the purpose of financing the process of development. Some of the possible ways of doing so have been highlighted above. It is quite likely that there are others. Microcredit may yet prove to be a powerful instrument of inclusive development.

References

Ahlin C (2009) Matching for credit: risk and diversification in Thai micro-borrowing groups, Bureau of research and economic analysis of development working paper 251

Ahlin C, Jiang N (2008) Can micro-credit bring development? J Dev Econ 86:1–21

Aniket K (2009) Queuing for credit: increasing the reach of microfinance through sequential group lending, Working paper available on the website www.aniket.co.in

Banerjee AV, Duflo E, Glennerster R, Kinnan C (2010) The miracle of microfinance? Evidence from a randomised evaluation. Centre for Micro-Finance, IFMR, Chennai

Bardhan P (1984) Land, labour and rural poverty: essays in development economics. Columbia University Press, New York

Bateman M (2010) Why doesn't microfinance work?. Zed Books, London

Besley T, Coate S (1995) Group lending, repayment incentives and social collateral. J Dev Econ 46:1–18

Besley T, Coate S, Loury G (1993) The economics of rotating savings and credit associations. Am Econ Rev 83:792–810

Bhattacharya S, Banerjee S, Mukherjee S (2008) Group lending and self-help groups: Joint benefit as an alternative governance mechanism. J Int Trade Econ Dev 17:1–19

Bowen MK, Morara M, Mureithi S (2009) Management of business challenges among small and micro-enterprises in Nairobi, Kenya. KCA J Bus Manage 2:16–31

Cassar A, Crowley L, Wydick B (2007) The effect of social capital on group loan repayments: evidence from field experiments. Econ J 117:F85–F116

Coleman B (1999) The impact of group lending in Northeast Thailand. J Dev Econ 60:105–142

Collins D, Morduch J, Rutherford S, Ruthven O (2009) Portfolios of the poor: how the world's poor live on $2 a day. Princeton University Press, Princeton

de Aghion BA, Morduch J (2005) The economics of microfinance. MIT Press, Cambridge

De Soto H (2000) The mystery of capital: why capitalism triumphs in the west and fails everywhere else. Basic Books, New York

Feigenberg B, Field EM, Pande R (2010) Building social capital through microfinance. NBER working paper no. W16018
Field EM, Pande R (2008) Repayment frequency and default in microfinance: evidence from India. J Eur Econ Assoc 6:501–509
Fischer G, Ghatak M (2009) Repayment frequency and lending contracts with present-biased borrowers, London school of economics working paper no. EOPP/2010/21
Fischer G, Ghatak M (2011) Spanning the chasm: uniting theory and empirics in microfinance research. In: Armendariz B, Labie M (eds) The handbook of microfinance. World Scientific Publishing, London
Ghatak M (1999) Group lending, local information and peer selection. J Dev Econ 60:27–50
Ghatak M (2000) Screening by the company you keep: joint liability lending and the peer selection effect. Econ J 110:601–631
Ghatak M, Guinnane T (1999) The economics of lending with joint liability: theory and practice. J Dev Econ 60:195–228
Gine X, Jakiela P, Karlan D, Morduch J (2010) Microfinance games. Am Econ J Appl Econ 2:60–95
Gine X, Karlan D (2009) Group versus individual liability: long-term evidence from Philippine microcredit lending groups. Yale Economics Department working paper 61
Institute Goldin (2007) Improving microcredit programs: listening to recipients—a summary of the pilot phase. Goldin Insitute, Chicago
Jain S, Mansuri G (2003) A little at a time: the use of regularly scheduled repayments in microfinance programs. J Dev Econ 72:253–279
Karnarni A (2007) Microfinance misses its mark. Stanford Social Innovation Review, Summer
Khandker S (1998) Fighting poverty with microcredit: experience in Bangladesh. Oxford University Press, New York
Khandker S (2005) Microfinance and poverty: evidence using panel data from Bangladesh. World Bank Econ Rev 19:263–286
Klonner S (2003) Rotating savings and credit associations when participants are risk averse. Int Econ Rev 44:979–1005
Klonner S (2008) Private information and altruism in bidding ROSCAs. Econ J 118:775–800
Laffont J-J (2003) Collusion and group lending with adverse selection. J Dev Econ 70:329–348
McIntosh C (2008) Estimating treatment effects from spatial policy experiments: an application to Ugandan microfinance. Rev Econ Stat 90:15–28
McIntosh C, Wydick B (2005) Competition and microfinance. J Dev Econ 78:271–298
Morduch J (1998) Does microcredit really help the poor? New evidence from flagship programs in Bangladesh. Harvard University, mimeo
Pitt M, Khandker S (1998) The impact of group-based credit programs on poor households in Bangladesh: does the gender of participants matter? J Polit Econ 106:958–996
Planning Commission (2009) A hundred small steps: report of the committee on financial sector reforms. Sage, New Delhi
Rahman A (1999) Women and microcredit in rural Bangladesh: an anthropological study of Grameen Bank lending. Westview Press, Boulder
Rai AS, Sjostrom T (2004) Is Grameen lending efficient? Repayment incentives and insurance in village economies. Rev Econ Stud 71:217–234
Rashid M, Townsend R (1994) Targeting credit and insurance: efficiency, mechanism design and program evaluation. World Bank discussion paper no. 47
Reserve Bank of India (2011) Report of the sub-committee of the Central Board of Directors of Reserve Bank of India to study issues and concerns in the MFI Sector
Roy Chowdhury P (2005) Group lending: sequential financing, lender monitoring and joint liability. J Dev Econ 78:415–439
Roy Chowdhury P, Roy J (2009) Public–private partnerships in micro-finance: should NGO involvement be restricted? J Dev Econ 90:200–208
Sadoulet L (1999) Equilibrium risk-matching in group lending. ECARES/ Universite Libre de Bruxelles, mimeo.

Stiglitz JE (1990) Peer monitoring and credit markets. World Bank Econ Rev 4:351–366
Stiglitz JE, Weiss A (1981) Credit rationing in markets with imperfect information. Am Econ Rev 71:393–410
Wydick B (1999) Can social cohesion be harnessed to mitigate market failures? Evidence from group lending in Guatemala. Econ J 109:463–475
Wydick B (2008) Games in economic development. Cambridge University Press, New York

Chapter 18
Employment Guarantee and Natural Vulnerability: A Study of MGNREGA in Indian Sundarbans

Prasenjit Sarkhel

Introduction

There is now a general consensus in the literature that the post reform period in India is characterized by higher GDP growth rate compared to the Hindu rate of growth in the pre-1991 period (Panagariya 2011; Dutta et al. 2012; Kotwal et al. 2011).[1] There are also compelling evidences that high GDP growth rate of the order of 6–8 % on an average has resulted in decline of poverty (Bhalla 2011; Datt and Ravallion 2010)[2] both in the rural and urban areas. Estimates reveal that poverty rates declined from 44.5 % in 1983 to 27.5 % in 2004 though the rate of decline is relatively lower in the post reform period during 1993–2004 (Bardhan 2007). On the other hand, evidence also points to disquieting features of post-reform growth, for instance the widening gap in absolute as well as relative consumption between the rich and the poor (Deaton and Dreze 2002; Sen and Himanshu 2004; Sarkar and Mehta 2010).[3] In particular structural changes associated with economic liberalization had adverse labor market implications in that there are now lesser employment opportunities for unskilled labor force. One reason is that the post reform expansion in the manufacturing sector had not been adequate to accommodate the growing rural labor force,[4] manufacturing accounted

[1] The point of structural break in GDP growth rate has however been debated extensively (Wallack 2003; Rodrik and Subramanian 2004; Balakrishnan and Parameswaran 2007).

[2] The last three decades have witnesses an annual GDP growth rate of 6 % and percapita growth of 4.2 % (Bhalla 2011).

[3] Moreover, inter-country comparison attested the fact that while per-capita Gross National Income in India has shown spectacular rise since one and half decade of market reforms, basic social indicators like life expectancy and mean years of schooling has plummeted in the same period compared to other South–Asian countries and China (Sen and Dreze 2011).

[4] Though population growth as a whole in India has slowed down rural labour force continued to grow at 2.8 % per annum (Binswanger 2012).

P. Sarkhel (✉)
Department of Economics, University of Kalyani, Kalyani, West Bengal 741235, India
e-mail: prasenjitsarkhel@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_18,

for only 17 % of the value added and 12 % of employment even after a decade of reform. In fact employment in the organized sector has actually declined while the use of contract labor has increased during the reform period (Kotwal et al. 2011). Employment growth in manufacturing has largely been concentrated in the informal or unorganized sector (Dougherty 2008). Though the service sector has been dominant and has shown the highest growth in the post reform period[5] yet its employment potential for unskilled workers in general had been low due to its bias for skill intensity. At the same time the share of agriculture in GDP has fallen much faster than the share of agricultural labor in total employment in the post reform period (Binswanger 2012). As a result land-labor ratio has fallen over time indicating accumulation of unskilled labor force in rural areas.[6] The exclusion from the labor market gets further accentuated as social security measures like targeted Public Distribution System (PDS) has largely been dysfunctional in the post reform period. For instance, in 2004–2005 almost two-thirds of the poor households were left out of PDS coverage and 62 % of the Below Poverty Line (BPL) and/or Antyodaya Anna Yojana (AAY) cards were in the hands of nonpoor households indicating huge inefficiency in targeting (Svedberg 2012).[7]

In face of such uneven growth labor market interventions in the form of income transfer programs has often been used a counterbalancing public policy that provides social security to the disadvantaged groups.[8] The Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) initiated almost half a decade after economic liberalization in 2005 is one such major government intervention in the rural labor market. The program committed to deliver 100 days of employment to every rural household willing to do unskilled manual work at a wage rate that is at least as much as the minimum wage laws. By far it is one of the largest workfare programs grossing almost .65 % of the Gross Domestic Product (GDP) at the end of 2009–2010 and the total program expenditure expanding from 1.94 % in 2006–2007 to 3.83 % of the total government expenditure in 2009–2010. In an era of fiscal restraint the implementation of MGNREGA throughout rural India reflects the policy-maker's belief in the potential net benefit from delivering of entitlement on a universal scale. However, realization of such benefits crucially depends on the extent to which the program outcome and outreach are ensured.

For the poor vulnerable and marginalized the decision to participate in MGNREGA hinges crucially on the availability of alternative income

[5] Services share in GDP is about 60 % but 30 % in total employment.

[6] Most of the employment in the rural sector has been generated by nonfarm employment, though the odds of finding nonfarm jobs crucially depend on the extent of human capital. As shown in Eswaran et al. (2009) the movement out of agriculture into nonfarm jobs is predominant among male literates as opposed to women and other disadvantaged classes like Scheduled Caste and Scheduled tribes. In backward areas these tendency would have resulted in increased underemployment and unemployment of unskilled labourer.

[7] Also see Khera (2008, 2011) for estimation of "wrongly included" and "wrongly excluded" households as well as extent of "under purchase" in case study of Rajasthan.

[8] See Subbarrao (2003) for a description of the workfare programs across the globe.

opportunities.[9] Thus, it is expected that the scheme is likely to be most useful in agricultural lean season when the income shock is transient in nature.[10] In fact the claim that MGNREGA had reduced distress migration (Menon 2008; Ministry of Rural Development 2012) tends to establish that the program is being used as an instrument of income stabilization and consumption smoothing by rural households. Contrarily, for systematic shocks like extreme weather events the disruption in livelihood may be of longer duration. Here, the discounted potential earnings from program participation vis-à-vis the damage loss is likely to differ across affected households. Accordingly, with altered participation constraints the extent to which affected households choose MGNREGA over migration possibilities in the aftermath of such aggregate shocks needs careful scrutiny. Increasing frequency of extreme weather events across the globe as well as in India makes this issue all the more pertinent.

This chapter attempts to identify the determinants of MGNERGA program participation in the aftermath of an actual natural shock-Cyclone Aila that plagued coastal and riverine areas of West Bengal in 2009. Apart from immediate loss in lives and damage of physical assets the catastrophic event also resulted in long-run disruption of livelihoods.[11] Saline inundation of agricultural fields rendered them unsuitable for cultivation until next few monsoons; the households had to choose between availing the MGNERGA program and migrating out for alternative employment. Analyzing participation decisions of households in four villages of Indian Sundarbans—one of the worst hit areas—we find that households with higher fixed assets take recourse to MGNREGA while the relatively asset poor and unskilled laborer's migrate. Our findings attempts to highlight the limitation of program outreach for a population that is naturally vulnerable.

We discuss this tradeoff between public program and alternative employment in the following order: "MGNREGA: The Design Issues" highlights the essential features of MGNREGA as a distinctive workfare scheme with social security agenda. The extent of program outreach for the economically and socially vulnerable group is analyzed in "Outreach of MGNREGA", "MGNREGA in Sundarbans" discusses the performance of MGNREGA in Indian Sundarbans in the backdrop of Cyclone Aila. The survey results are presented in "The Survey" followed by the conclusion.

[9] However, even when alternative earning potential is low poor implementation of the program could deter the participation rate.

[10] For instance in the Financial Year 2010–2011 it is estimated that 70 % of person days of work were generated in the lean agricultural season (Ministry of Rural Development 2012).

[11] Due to Aila, 500 km of the embankment were washed away affecting over 6.77 million people and killing 137 people in the North 24 Parganas and South 24 Parganas parts of the Sundarbans in addition to either partially or completely destroying 926,000 semi-permanent houses (Development Research Communication and Service Centre 2009).

MGNREGA: The Design Issues

The design of wage employment program as public provision of social security had its antecedents in India. In fact, in explicit realization of the inadequacy of 'trickle down effect' of economic growth, public intervention in the labor market has been put in place from Third Plan onward starting with Rural Manpower Program in 1964 (Planning Commission 2001). Thereafter, a plethora of wage employment program has been introduced often with definite area of operation like Drought Prone Area Program in the 1970s.[12] There were also specific target group-based programs such as Rural Landless Employment Guarantee Program (RLEGP) in 1983–1989 and Jawahar Gram Samridhi Yojana (JGSY) in 1999–2001 that targeted landless and Below Poverty Line (BPL) households. The design of MGNREGA owes much to these past programs—in particular the theme of creating durable community assets through provision of employment in the lean season had been the common agenda of most of these programs.[13] The distinguishing feature of MGNREGA had been its renewed emphasis on providing unskilled employment opportunities with a 'universal' coverage. In the first phase the program covered 200 most backward districts, but by 2008–2009 it extended to all the rural areas. More importantly, it served to establish, for the first time, the right to work as a legal entitlement (Mehta et al. 2011). Unlike previous workfare programs, MGNREGA provides "durable" legal entitlement as it can be withdrawn only through amendments in parliament while earlier schemes were ceased through mere official orders.

Rural households get included in MGNREGA under a two-prong procedure: first the Gram Panchayat (GP) conducts a survey and registers rural households willing to provide unskilled manual labor at a predetermined program wage rate. It is mandatory to provide job cards to all registered households. Second, all job card holders must apply for jobs to the GP that must be provided within 15 days of application and failure to which will entitle the job seeker for an unemployment allowance payable by the State Employment Guarantee Council. The projects under MGNREGA fall under eight categories: water conservation, drought proofing, flood control, rural connectivity, land development, renovation of traditional water bodies, irrigation works both minor and major and provision of irrigation facilities in lands of SC and ST households, and any other works notified by the Ministry of Rural Development. The planning and the choice of work are required to be made in open assemblies of local level institutions like *Gram Sabha*

[12] Some other area specific programs were Employment Assurance Scheme in (EAS) during 1993–1999 where the latter attempted to provide employment in 1752 backward blocks in lean agricultural seasons. The immediate precursor of MGNREGA—the National Food for Work Program II in 2000–2002 was also implemented in drought affected rural areas.

[13] In fact, the operational guidelines of MGNREGA permits the convergence of MGNREGA funds with funds from other schemes for the creation of durable assets.

and the *Gram Panchayat*—the latter being the nodal agency for the implementation of the projects (Ministry of Rural Development 2012).

However, despite the demand driven nature of MGNREGA with fixed outlays and notified wage rates, job rationing becomes inevitable. This could pose a constraint for the outreach of the program. The extent of this rationing is captured by estimating the unmet demand for work across states. Using the NSSO employment and unemployment survey for the 66th round, Dutta et al. (2012) estimate the share of the households that were rationed as the product of rationing rate (i.e., proportion of households who wanted work and did not get it) and demand rate (proportion of households that wanted to work in the scheme). The negative correlation between the poverty rate and the proportion of rationed households leads them to conclude that poorer states have larger unmet demand for work. They further reason that this phenomenon may have arisen due to supply side factors like weak administrative capacities that dampened the higher demand for work via poverty. This could lead to serious setbacks in terms of program outreach if the economically and socially backward people are disproportionately rationed out of the scheme. Using national level data, we next probe the issue of whether there has been any discrimination at the entry point of MGNREGA.

Outreach of MGNREGA

In order to evaluate the efficacy of MGNREGA one must take into account the spread as well as intensity of the workfare schemes. In this respect, the proportion of households that were taken under the coverage of the scheme and the average number of employment days provided in a financial year can be used as the measure of spread and intensity, respectively. During the period 2008–2009 to 2011–2012,[14] on an average, the percentage of registered households provided job cards were close but could not reach full coverage marks.[15] One important indicator of program outreach would then be the extent to which job card holders demanded employment. Thus, if job card attainment truly reflects the demand for wage employment then we also expect to see an overwhelming majority of job card holders expressing their intention for work. However, starting from 28 % in 2008–2009 this figure has only been able to touch 50 % rate in 2011–2012 (See Fig. 18.1).

[14] We restrict our analysis from the period 2008–2009 as from this time onwards the coverage of the scheme extended to the entire rural India. Thus, by excluding the period 2006–2007 and 2007–2008 i.e. when the program started in 200 districts and added 130 districts we avoid the problem of comparability.

[15] Among the states Arunachal Pradesh consistently fell short of the full coverage over this period. Goa started low at 5 % coverage rate in 2008–2009 but later caught up and reached more than 90 % coverage in 2011–2012.

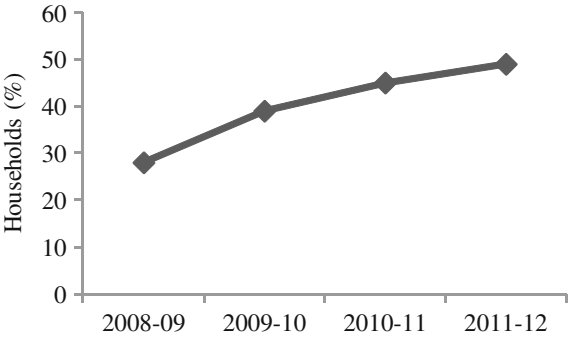


Fig. 18.1 Percentage of job-card holding households demanding employment. *Source* www.nrega.nic.in

It may well be possible that registered households tend to use MGNREGA as insurance and would fall upon the program only when it would provide highest alternative income to its household labor. If that be the case then the participating households are likely to be poor and here the demand for employment must show strong association with measures of poverty. Using state- level figures for households demand for employment and poverty we did not find any consistent statistical association between the head-count index as well as the per-capita Net State Domestic Product (NSDP) with the propensity of program participation (See Table 18.1). In fact, except for the initial period 2008–2009 demand for employment is even negatively related with head-count index.

Our finding echoes well with Dutta et al. (2012), as they did not find any strong correlation with participation rates and poverty.[16] However, micro studies have also come up with opposite facts: for instance Dreze and Khera (2011) in their survey of six states in 2008 noticed that more than 8 % of the sampled participants live in derelict houses while more than 72 % are devoid of electricity at their place of residence. Similarly, another study in five districts of Uttar Pradesh also reports that BPL households constitute 85 % of the program participants in the sample.[17] We also use other economic variables like various measures of occupational pattern to understand the linkage between livelihoods and outreach of the MGNREGA. Table 18.2 reports the correlation between occupational pattern and

[16] Of-course institutional shortfalls in terms of staff inadequacy could also lead to lower project implementation. For instance, two key administrative posts in MGNREGA involve *Gram Rozgar Sahayak (GRS)* and *Program Officer (PO)*. The GRS helps in labour budgeting and project implementation at the village level while the PO enables consolidation of GP level plans into block plans and supervises the implementation at the village level. On an average over the country as whole the actual staff number for GRS and PO fell short of the requirement by 20 and 11 % with 38 % of the states (8 out of 18) reporting above average shortfalls. The correlation coefficient between these measures of shortfalls and percentage of employment demanded is weakly positive (0.16 and 0.14, respectively) and does not seem to indicate a significant relationship with employment provided.

[17] Indian Institute of Management-Lucknow (IIM-L), 'Quick Appraisal of Five Districts of Uttar Pradesh', Report submitted to the Ministry of Rural Development/UNDP, Lucknow: Indian Institute of Management, 2009.

Table 18.1 Demand for employment in MGNREGA and poverty

Poverty measures	Households demand for employment (%)			
	EMP_0809	EMP_0910	EMP_1011	EMP_1112
Per-capita NSDP 2009–10 (constant price)	−0.38**	−0.21	−0.09	0.02
Head-count index	0.26	−0.01	0.02	−0.25

Note EMP_t: Percentage of job card holder households demanding employment at time t, MGNREGA data from www.nrega.nic.in, Head count index from Tendulkar Poverty Estimates from http://planningcommission.nic.in/eg_poverty.htm ** denotes significance level at less than 5 %

Table 18.2 Occupational pattern and MGNREGA participation

Household participation	Occupation					
	Agricultural worker	Secondary worker	Tertiary worker	Casual worker	Self-employed	Salaried worker
HH demanding Job Cards	0.23	−0.002	−0.41**	0.20	−0.09	−0.36
HH got NREGA work	0.18	0.05	−0.37*	0.21	−0.11	−0.31
HH did not seek NREGA work	−0.25	−0.003	0.45**	−0.13	0.02	0.55**
HH sought NREGA work but did not get	0.12	−0.14	−0.06	0.11	−0.04	−0.41**

Source Key Indicators of Employment and Unemployment in India, 2009-10, NSSO 66th Round (2011), Minsitry of Rural Development (2012)
Note * denotes significance level at less than 10 %, ** less than 5 %

various measures of MGNREGA participation across Indian states from NSSO 66th round employment and unemployment survey (2009–2010).

It appears that households in the service sector are not interested in MGNREGA. This may indicate that self-targeting in MGNREGA may exclude people with higher income earning potential from capturing the benefits of the program. However, there are no statistically significant evidences that workers in relatively poor performing sector like agriculture demand MGNREGA employment.

Next, we also report the degree of association between main and marginal workers across Indian states and program participation (See Table 18.3). The proportion of cultivators both, in case of main and marginal workers seems to be positively related with the demand for work though the correlations are significant only for the initial two years. Agricultural laborer does not display any association with demand for employment.

Given the aggregate nature of the data one should not in haste dismiss the outreach of MGNREGA in the rural labor market. Rather the impact varies across states presumably depending on the administrative efficacy of its implementation. For instance, the Ministry of Rural Development (2012) refers to the NSSO survey of MGNREGA in three states of Andhra Pradesh, Madhya Pradesh, and Rajasthan that reports that agricultural labor constitutes the majority among the program

Table 18.3 Main and marginal worker and MGNREGA participation

Household participation	Occupation					
	Main/total	Cult/main	ALab/main	Marg/total	Cult/marg	ALab/marg
EMP_0809	−0.31**	0.58**	−0.23	0.31	0.41**	−0.17
EMP_0910	−0.01	0.47**	−0.15	0.01	0.36**	−0.03
EMP_1011	0.06	0.21	−0.05	−0.06	0.08	0.02
EMP_1112	0.01	0.16	−0.39**	−0.01	0.19	−0.34**

Source Census 2001 for main and marginal workers across Indian States, www.nrega.nic.in for the employment demand figures

Note ** denotes significance level at less than 5 %. Total: total workers, Main: main workers, Marg: marginal workers, Cult: Cutivators, ALab: agricultural laborer, EMP_t: Percentage of Job card holder households demanding employment at time t

participants. In fact Engler and Ravi (2009) find in their longitudinal survey of Medak district of Andhra Pradesh that people often regard NREGA as an alternative to agricultural labour in addition to supplementary income source in the agricultural lean season.

In terms of social vulnerability, the outreach of the program is more apparent. Over the years the participation of backward and disadvantaged class in terms of SC and ST households have more or less remained constant at 40–50 %. Though, movements within groups suggest that participation of ST has been falling slightly over the years that had been partially made up by the SC participation. In fact, for both SC and ST the participation rate exceeds their share in the total population at the state level (Ministry of Rural Development 2012).

The program outreach can be also be gauged in terms of average person days generated in MGNREGA. As such, the average person days in the program remained well below the program objective of 100 days of employment. Starting from 43 days of average employment days in 2006–2007 it reached its highest of 54 days in 2009–2010 again coming down to 47 days in 2010–2011 (Ministry of Rural Development 2012). If the poor had lower program participation in terms of lesser number of guaranteed work days then it must be accounted for in the outcome indicator. To that end, we estimate the ratio of average number of person days (AM) and the harmonic mean (HM) of person days generated for the period 2008–2009 to 2010–2011. The harmonic mean gives higher weight to lower person days and hence the ratio of AM/HM is likely to indicate the disparities in program provisioning across countries. Our estimates suggest that the ratio remained fairly constant during these periods with a value close to 2 indicating more or less constant pattern of employment generation with no marked increase in disparity.

However, 2009 marked an abnormal year in that ten states of India were declared as drought affected. These included Assam, Bihar, Himachal Pradesh, Jharkhand, Karnataka, Madhya Pradesh, Maharashtra, Manipur, Nagaland, Rajasthan, and Uttar Pradesh. As can be observed from Figure 18.1 this period also shows the highest increase in the demand for employment—an increase of over 11 %—compared to later years. In fact, for all the drought affected states except

Manipur and Bihar the demand for employment has been well above the national average of 38 %. However, it needs to be seen whether this increase in demand emanated mostly out of affected and vulnerable population. Micro studies have come up with mixed evidences regarding the effectiveness of MGNREGA as a social security measure in the face of weather shocks. For instance, Johnson (2009) reports increases in wage income from MGNREGA in rain deficit areas. On the other hand, Somanathan and Somanathan (2009) finds that in the aftermath of the Kosi flood in Bihar in 2008 although more than two-third of sample households were aware of MGNREGA there is negligible participation in the program in the post flood period. The following section describes the background of the inter-linkage between natural vulnerability and MGNREGA in the context of India Sundarbans that was devastated by Cyclone Aila in May, 2009.

MGNREGA in Sundarbans

The Indian Sundarbans, an archipelago of 102 islands, is a socio-economically backward area located in the southern part of the state of West Bengal. Spread over 19 administrative blocks in two districts North 24 Parganas (six blocks) and South 24 Parganas (13 blocks),[18] the Sundarbans region has a relatively higher concentration of SC/ST households (36.5 %) than the state average (25.6 %). Between two census years 1991–2001 the population in Indian Sundarbans increased by 17 % but this was accompanied with an eightfold increase in marginal workers though the proportion of main workers remained fairly constant. The majority of the population is engaged in agriculture (60.32 %) but only 25 % of those engaged in agriculture are cultivators themselves and so the region has a higher preponderance of farm wage laborers. Thus, of the total land holding, small and marginal farmers account for 85.22 % that is quite high compared to the state average (43.48 %).[19] In this background of asset poor and predominantly marginal population, the adoption of MGNREGA is likely to have larger acceptance. The underprivileged social and economic setting of the inhabitants also gives us an opportunity to verify the extent to which disadvantaged class chooses MGNREGA as a social security vis-à-vis alternative earning opportunities.

At a first glance, contrary to the expectation, it appears that Sundarbans have generated lower days of employment compared to the state as well as the district in which they reside (Table 18.4). Within 2 years of the inception of MGNREGA the

[18] The six blocks in the North 24-Pgs are Haroa, Minakhan, Hingalgunj, SandeshkhaliI, Sandeshkhali II and Hasnabad, the other 13 blocks in the South 24 parganas are Basanti,Canning I & II, Namkhan, Kakdweep, Sagar, Patharpratima, Mathurapur I & II, Jaynagar I &II and Kulatali.

[19] Farmer owning land between 0.13–0.27 ha (1–2 bigha) are considered marginal farmers while those having land between 0.27–0.67(2–5 bigha) are considered small (Centre for Science and Environment 2012).

Table 18.4 Average person days in MGNREGA (2008–2011)

Regions	2008–2009	2009–2010	2010–2011
Sundarbans	12	22.60	27.50
N-24 Pgs	42	28.31	32.21
S-24 Pgs	19	17.16	20.33
West Bengal	26	46	31.07

Source www.nrega.nic.in

Indian Sundarbans were torn apart by Cyclone Aila in May 2009 that dilapidated almost 500 km of river embankments that provided crucial defence from tidal inundation. Much of the MGNREGA activities were devoted toward building up of flood control infrastructure and roads in the post-Aila period. Thus, the stepping up of the employment generation in the post 2009 period must be seen after controlling for the reconstruction motive.

However, employment provision has been largely inadequate, particularly in the post-Aila years. In fact, the year when Aila struck and restoration and reconstruction was in full force only half of the total people with job cards demanded jobs under MGNREGA (CSE 2012).

Whether this phenomenon is a due to supply side effects of project nonimplementation or is it really the case that there has been a lack of demand for the scheme in these areas is an area that needs careful examination. The following section reports the results of the household survey in Sundarbans to assess the latter. The survey covered 399 households in three villages of Khulna Gram Panchayat in Sandeshkhali II block of North 24 Parganas. Note that by limiting our observation in a single Gram Panchayat we tend to obviate the need to control for supply side impacts—our objective is more directly related to assess the factors that influences the demand for MGNREGA at the household level.

The Survey

Khulna Gram Panchayat in North 24 parganas part of Sundarbans is a habitat that is dominated by backward classes. As of 2001 Census almost 70 % of the population comprises Scheduled Castes and 17 % of the population belongs to the scheduled tribe. The total labor force comprises 66 % of main workers among which cultivators (29 %) and agricultural laborers (38 %) constitute the majority. Although agriculture remains the primary occupation, there is little or no surface or underground irrigation—90 % of the area available for cultivation is un-irrigated. During the Aila more than 1 km of the embankments in the area gave way (Department of Irrigation and Waterways) and saline inundation posed a huge threat to agriculture. In this situation inaction of MGNREGA as an alternative earning opportunity is likely to be welcomed. Using the household list of the three villages (Khulna, Hatgacha and Sitalia) 399 households were chosen randomly for

structured interview during the month of July–September 2011. The questionnaire were organized to enquire about the household's assets in terms of landholding and livestocks, demographic information like family size, caste, and religion, the extent of their social network like involvement in local organizations and political activities; finally they were asked about the number of days the household members found employment in the MGNREGA scheme during 2008–2010.

Data Description

In keeping with population characteristics the sample constitutes large concentration of Scheduled caste households (52 %) followed by Scheduled Tribes (22 %) and Other Backward Classes (19 %); the rest belong to general classes and Muslims. Table 18.5 reports the characteristics of the sampled household.

Observe that the sample consists of landless households as well. In fact though almost all of them possess homestead land around 18 % of the sampled household's does not possess any land. These are households that are either involved in petty or small businesses. In fact, within these households, 5 % of households own neither live stock nor agricultural land. Apart from petty and small businesses households major source of income includes cultivation in own land or rearing fishes in ponds. Notice that on average households agricultural land holding varies between 1–2 bigha so that they are mostly marginal farmers and small farmers. Similar pattern also holds for the area of the fish ponds.

Table 18.5 Household characteristics

Household characteristics	Mean (S.D)	Min	Max
Household size	4.52 (1.55)	2	12
Homestead land (bigha)	0.41 (0.29)	0	2.5
Agricultural land (bigha)	1.05 (1.27)	0	6
Pond land (bigha)	0.17 (.37)	0	5
Total land holding (bigha)	1.64 (1.53)	0	8.5
Livestock (Standard livestock unit)	1.20 (1.42)	0	6.68
BPL score (2005)	29.23 (3.05)	18	33

Note Conversion factors used to calculate Standard Livestock Unit are: bull = buffalo = 1, cow = 0.7, goat = sheep = 0.1, pig = 0.4, poultry = duck = 0.02 (Akter et al. 2008). Figures in parenthesis show standard deviation. 1 bigha = 0.1338 hectare

For the sampled households as a whole the average number of days worked in MGNREGA per member of the households significantly increased in the post-Aila period compared to 2008.[20] The average per person employment days of the sampled households shows a steady increase from 11 days in 2008, 18 days in 2009, and 20 days in 2010 with the median increasing from 9.5 days in 2008 to 15 days in 2010.

Despite this aggregate increase in the per-capita employment there is significant variation when we observe the distribution of employment days across different occupational groups (Table 18.6). Along with the number of days in employment we also report the maximum number of days a member of the household has to stay out of the village in the post-Aila period for supplementing the household income. In fact there has been a sharp increase in the percentage of households that has more members working outside the village in the post-Aila period. However, this could happen because of the wage gap rather than job rationing—but it is also the case that total allocation for MGNREGA in Khulna declined in absolute terms after the year following Aila (See Fig. 18.2).

Although post-Aila employment increased across all categories of household it is clear that farm households have worked a significantly larger number of days compared to other households. Though farm households that supplement their income with other activities accessed the program till 2009, by the next year the majority of them seemed to prefer the exit options and have members who out-migrated for at a least quarter of a year. As expected unskilled manual workers have higher days of employment in the program but at the same time had members in the household who would stay away half a year for outside earning. Thus, we hypothesize that households without any fixed assets are more prone to migration while those having land demands MGNREGA as a supplementary income.

The fact that a segment of the farm households containing members working in the nonagricultural sector has a poor program participation as well as higher rate of migration compared to farm households needs to be explained. It may be noted that per-capita operational land in case of the former is significantly lesser than that of the latter.[21] Given productivity this implies lesser availability of farm proceeds for each of the member households and hence there may be incentive to migrate. This could also indicate that there are other mechanisms at work whereby the selection of the needy within the program is conditioned. In addition, identical average days of employment between the farm households and households in the unskilled manual sector might also hint the presence of "*program capture*".[22] The fact that the latter also has substantially higher rate of migration indicates that

[20] The response of the household was validated by checking the job card and also through local informants.

[21] The per-capita operational land holding for farm households (0.73 bigha) exceeds that of second category of farm households (0.53 bigha) and the difference is significant at less than 5 %.

[22] Thus the nonpoor may exercise his economic power and social connections for gaining preferential treatment in welfare services thereby "capturing" the benefit of a program that is intended for the poor.

Table 18.6 Distribution of employment days under MGNREGA

Major occupation	Per person employment 2008		Per person employment 2009		Per person employment 2010		Maximum number of days migrated	
	Mean (S.D)	Median	Mean (S.D)	Median	Mean (S.D)	Median	Mean (S.D)	Median
Farm owners (fishery pond agricultural land)	5.94 (6.88)	4.75	15.01 (12.91)	13	21.15 (23.98)	16.25	29.77 (81.59)	0
Farm owners +manual work	5.39 (6.33)	4	13.39 (10.09)	13.75	10.92 (13.63)	4.66	115.37 (118.03)	120
Agricultural laborer	8.15 (7.66)	7	12.83 (13.72)	9.75	15.92 (19.97)	12	28.12 (75.38)	0
Private jobs +petty business	3.03 (5.18)	0	6.28 (9.99)	0	10.90 (12.54)	7	22.89 (66.65)	0
Unskilled manual work	8.38 (9.45)	5	16.27 (14.21)	14.5	20.75 (18.98)	19	115.60 (113.41)	180

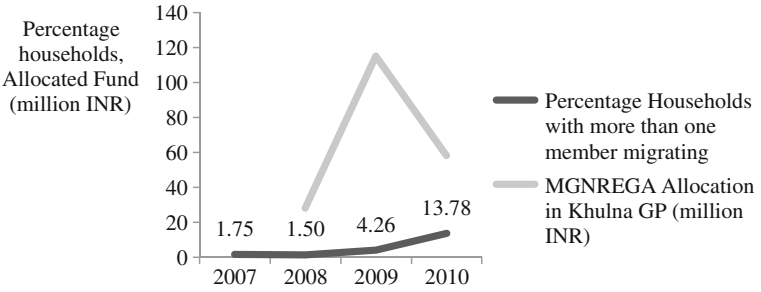


Fig. 18.2 MGNREGA allocation and migration

there are likely to be large unmet demand for work. Thus, apart from asset indicators, we need to focus upon the extent of household's social network that could earn him higher days of employment in the program. Here, we consider two such indicators: whether the household is a relative of the Panchayat authorities and whether the household is an active member of the political party. By active member we specifically asked the household whether he attends the political campaign of the political party in general and during elections. The impact of these two factors on household employment is reported in Table 18.7.

However, it is difficult to reach any conclusion regarding employment in MGNREGA and household's relation with Panchayats: those having some relations in Panchayats got fewer days of employment than nonrelatives but the former also has the higher assets in terms of land holdings and livestock so the need to resort for a wage employment program might be comparatively lesser in their case. In contrast, the result with association of political parties is much straightforward: asset poor households might use their political network to garner higher days of

Table 18.7 Impact of social network on MGNREGA employment

Landholding, livestock and employment (Mean)	Extent of social network			
	Close/Distant relation of panchayat		Member of political party	
	Not related	Related	Nonmember	Active member
Land Holding (bigha)	1.47	1.79	1.71	0.95
	(1.30)	(1.73)	(1.55)	(1.13)
Operational Land (bigha)*	1.01	1.43	1.28	0.68
	(1.16)	(1.62)	(1.44)	(1.04)
Livestock (Standard Live stock Unit)	0.88	1.54	1.26	0.67
	(0.08)	(0.11)	(0.07)	(0.12)
Person days per Household 2008	8.40	3.57	4.25	6.19
	(9.03)	(0.50)	(8.31)	(7.66)
Person days per Household 2009	15.14	8.84	7.09	12.63
	(13.42)	(11.75)	(9.79)	(13.17)
Person days per Household 2010	19.08	12.39	10.60	16.44
	(18.88)	(16.02)	(16.99)	(17.91)

Note * Operational land includes all farmlands and/or fish ponds except homestead lands. Figures in parentheses show standard deviation. 1 bigha = 0.1338 hectare

employment. To understand the causal relationship with economic and social factors, we next estimate the per member employment as a function of the economic and political assets.

Estimation Results

As detailed in the earlier section the average number of days worked under the scheme has been increasing over the three-year period 2008–2010. Evidently, this also implied that the percentage of nonparticipants also declined over the years.[23] This means that the indicator of employment, i.e., per member employment days has a number of zero observations. Here a Tobit model needs to be estimated. However, estimating Tobit regression for each year separately involves the assumption that there is no contemporaneous correlation in the error terms. Thus, in addition we also estimate a pooled Tobit regression where we control for the Aila and Post-Aila periods by defining a period dummy. We have three time periods, and we define two dummies: D_09 that takes a value 1 if observation pertains to the year of Aila, i.e., 2009 and 0 otherwise, and D_10 with one for

[23] From 47 % in 2008 the percentage of nonparticipants declined steadily to 31 % in 2009 and finally to 20 % in 2010.

observations relating to 2010 and 0 otherwise. The regression results are reported in Table 18.8.

The regression results clearly shows that prior to Aila there is but a weak indication of program capture as livestock holdings positively affected the employment days in MGNREGA. In the post Aila period employment is positively related with agricultural holdings: for per bigha increase in agricultural holdings per person employment increased by more than one day and the marginal effect of agricultural landholdings increased even more in the next year (i.e. 2010). At the same time higher endowment of operational holdings in the form of ponds discouraged employment in MGNREGA. Furthermore, in about two years after Aila even literate people are seeking unskilled manual jobs.

The increasing 'capture' of the farm household in the MGNREGA has interesting dynamics behind it. In the pre-Aila period most of the projects under the MGNREGA scheme were directed toward water conservation and irrigation activities that are supposed to increase farm productivity. However, saline inundation after Aila caused huge economic losses to agriculture in particular and at the same time the projects had to be redirected toward reconstruction and rebuilding of embankments. As Table 18.9 shows, more and more person days in the MGNREGA were directed toward flood control and rural connectivity projects, shifting the emphasis from farm-related projects in the post-Aila period. Thus, agricultural households had the option of out-migrating or they could wait back for the monsoon to wash away the salt and regain their field. It seems that the farm households had opted for the latter and crowded out accommodation for the manual workers from the scheme. Our findings finds support in a study by GFK Mode on behalf of the NREGA cell in North 24 Parganas where villagers report that people with relatively higher land endowments have stayed back in the post-Aila period to access the benefits of MGNREGA(Gfk-Mode 2010).

It may also be noted in passing that people with near-relatives in the Panchayat and active members in the political party also have a higher amount of agricultural landholding.[24] Thus, inclusion of farm households might have taken place due to the intensity of their social network as well. Interestingly, the nature of migration is mostly interstate (mostly in South India) rather than typical rural–urban migration to the nearest urban center Kolkata. Thus, among the sampled households almost 65 % of the household has members who migrated in the Southern States of Tamil Nadu and Karnataka and 14 % households have members that moved even out of the country for earning. The interstate migration is mostly seen in case of manual and unskilled workers.[25]

[24] The average agricultural land for those having near relatives in the Panchayat is 1.27 bighas while for those without any relatives the figure is 0.88 bighas. The difference is significant at less than 1 %. Similarly the difference in landholdings across households having politically active members and those that haven't is positive and significant at less than 5 %.

[25] Incidentally these are the groups that have below sample average land-holdings and higher than average adult members. Thus per-capita landholdings are abysmally small for these groups to eke out their living.

Table 18.8 Tobit estimates

Variables	PPD_0809		PPD_0910		PPD_1011		PPD_0811	
	Coeff. (std.err)	Marg. effects	Coeff. (std.err)	Marg. effects	Coeff. (std.err)	Marg. effects	Coeff. (std.err)	Marg. effects
D_09(1 = 2009,0 otherwise)	–	–	–	–	–	–	9.63***	4.42
D_10(1 = 2010,0 otherwise)	–	–	–	–	–	–	15.67***	2.75
Agricultural Land (bigha)	0.17 (0.68)	0.06	2.43*** (0.87)	1.13	2.55** (1.02)	1.48	1.55*** (0.54)	0.71
Education dummy (1 = Household members above school secondary,0 otherwise)	0.83 (1.76)	0.30	2.47 (2.25)	1.15	5.45*** (2.56)	3.17	3.74** (1.37)	1.71
Pond Land (bigha)	−1.11 (2.21)	−0.40	−5.45** (2.96)	−2.54	−6.29** (3.30)	−3.66	−4.29** (1.76)	−1.96
Member of Political Party (1 = active member,0 otherwise)	3.01 (2.75)	1.10	7.04** (3.47)	3.28	11.69*** (3.83)	6.81	7.11*** (2.11)	3.25
Relation with Panchayat (1 = household related with panchayat,0 otherwise)	−9.6*** (1.8)	−3.50	−9.45*** (2.37)	−4.41	−3.70 (2.66)	−2.15	−8.59*** (1.39)	−3.93
Livestock (SLU)	1.53** (0.62)	0.55	1.90** (0.80)	0.89	1.01 (0.93)	0.59	2.08** (0.49)	0.95
ST dummy (1 = if household belongs to ST 0 otherwise)	5.53*** (1.80)	2.02	2.21 (2.31)	1.03	9.09*** (2.70)	5.29	1.01 (1.45)	0.46
OBC and Muslim dummy (1-household belongs to OBC or is Muslim 0, otherwise)	−0.012 (2.38)	−.004	−8.08** (3.09)	−3.77	−11.32*** (3.31)	−6.59	−2.14 (1.48)	−0.98
Constant	−1.28 (3.13)		0.85 (3.97)		2.53 (4.42)		−4.29** (1.76)	
	N = 373, LR Chi2(8) = 62.28*** Log likelihood = −857.57		N = 368, LR Chi2(8) = 87.65*** Log likelihood = −1122.48		N = 369, LR Chi2(8) = 69.75*** Log likelihood = −1356.56		N = 1110, LR Chi2(9) = 228*** Log likelihood = −3441.27	

Note PPD_t: Persondays per household for period t, t = 2008–2009, 2009–2010, 2010–2011, Marginal effects are conditional on being uncensored ** denotes significance level at less than 5 %, *** less than 1 %. 1 bigha = 0.1338 hectare

Table 18.9 Distribution of the types of projects during 2008–2011

Types of projects worked upon	2008–2009 (% HH)	2009–2010 (%HH)	2010–2011 (% HH)
Conservation/Renovation of water bodies	53.77	6.08	2.58
Irrigation works	26.23	12.17	0.32
Flood control (embankments)	11.06	73.00	82.26
Rural connectivity	8.54	8.37	14.84

Concluding Observations

The inadequate expansion of the manufacturing sector in the post reform period together with dominance of skill intensive service sector led to lower employment opportunities for the unskilled labor. Though the share of agriculture in GDP displayed a falling trend in the past decade, yet claims on agricultural product increased as land-labor ratio continued to fall. In this situation MGNREGA was introduced as an instrument of public policy to ensure social security and redistribution of income in the rural labor market. Committing huge public resource for MGNREGA, so far as the available evidence goes, have been partly successful to reach the poor but has not really taken the poorest under its aegis. We locate this limitation of the program outreach by identifying the tradeoff between MGNREGA participation and outmigration in the presence of systematic income shock—Cyclone Aila in Indian Sundarbans. Our results suggest that in the event of a weather shock the asset poor households may choose to opt away from the schemes while those with relatively higher endowment of fixed asset may use their social network to gain access to the program. In other words, with binding subsistence constraint, the poor may discount the potential program earning at a higher rate than the relatively wealthier household. From the point of view of the planner, reallocating MGNREGA funds from projects like water conservation and irrigation works that improves rural livelihood toward community infrastructure like flood control and rural connectivity seems to be the rational choice in the post-disaster scenario. However, this in turn may create incentive for landed households to wait back and expend effort for alternative employment in the program. With fixed budget outlay this could also crowd out the poorest of the poor from the intended social security of MGNREGA. Our analysis thus drives home the point that to be effective as a social security in face of complex weather shocks MGNREGA program design must account for the differences in time preference of its potential beneficiaries.

Acknowledgments The household survey reported in this chapter is funded by University Grants Commission. The chapter has benefitted immensely from discussion with Sarmila Banerjee. For insightful comments on earlier drafts we are grateful to participants in the Workshop on Livelihood, Employment and Vulnerability organized by Department of Economics, Calcutta University on January 5, 2012.

References

Akter S, Farrington J, Deshingkar P, Freeman A (2008) Livestock, vulnerability, and poverty dynamics in India. Findings from the ODI livelihood options project panel survey. Discussion Paper No. 10. Targeting and Innovation. International Livestock Research Institute (ILRI), Nairobi, Kenya, p 43

Balakrishnan P, Parameswaran M (2007) Understanding economic growth in India: a prerequisite. Econ Polit Weekly 42(27–28):2915–2922

Bardhan P (2007) Poverty and inequality in China and India: elsuive link with globalization. Econ Polit Weekly 42(38):3849–3852

Bhalla SS (2011) Inclusion and growth in India: some facts, some conclusions. Asia Research Centre Working paper, 39

Binswanger MH (2012) India 1960–2010: Structural non farm sector, and the prospects of agriculture. Department of Agricultural and Resource Economics University of California, Berkeley

Centre for Science and Environment (2012) Living with changing climate impact, vulnerability and adaptation challenges in Indian Sundarbans. Centre for science and environment, New Delhi

Datt G, Ravallion M (2010) Shining for the Poor too ? Econ Polit Weekly 7:55–61

Deaton A, Dreze J (2002) Poverty and inequality in India—a re-examination. Econ Polit Weekly 37:3729–3746

Development Research Communication and Service Centre (2009) Report of consultation for regeneration of agriculture and livelihood systems in Aila affected areas, Kolkata http://www.drcsc.org/resources/agriculture%20and%20livelihood%20after%20AILA.pdf

Dougherty S (2008) Labour regulation and employment dynamics at the State level in India. OECD Economics Department Working Papers, No. 624

Dreze J, Khera R (2011) The battle for employment guarantee. Oxford University Press, New Delhi

Dutta P, Murgai R, Ravallion M, van de Walle D (2012) Does India's employment guarantee scheme guarantee employment? Econ Polit Weekly XI VII(16)

Eswaran M, Kotwal A, Ramaswami B, Wadhwa W (2009) Sectoral labour flows and agricultural wages in India, 1982–2004: has growth trickled down? Econ Polit Weekly 44(2):46–55

Engler M, Ravi S (2009) Workfare as an effective way to fight poverty: the case of India's NREGS. Social science research network: http://ssrn.com/paper=1336837, Accessed on 14 June 2012

Gfk-Mode Pvt Ltd. (2010) NREGA scheme on embankment. Report prepared for NREGA Cell, North 24 Parganas, Barasat, http://north24parganas.gov.in/mgnrega/pdf/Embankment.pdf

Johnson D (2009) Can workfare serve as a substitute for weather insurance? The case of NREGA in Andhra Pradesh Working Paper 32, Chennai: Institute of Financial Management and Research (IFMC)

Khera R (2008) Access to the targeted public distribution system: a case study in Rajasthan. Econ Polit Weekly 43(44)

Khera R (2011) India's public distribution system: utilization and impact. J Dev Stud 47(3):1–23

Kotwal A, Ramaswami B, Wadhwa W (2011) Economic liberalization and Indian economic growth: what's the evidence? Discussion Paper 11–13. Indian Statistical Institute, Delhi, Planning Unit

Mehta A, Shepherd K, Shashanka A, Shah B, Kumar A (2011) India chronic poverty report: towards solutions and new compacts in a dynamic context. Indian Institute of Public Administration, New Delhi

Menon SV (2008) Right to information Act and NREGA: reflections on Rajasthan. Paper 7351. Munich Personal RePEc Archive.Ministry of Rural Development 2012, Munich

Ministry of Rural Development (2012) MGNREGA Sameekhsha an anthology of research studies on the Mahatma Gandhi National Rural Employment Guarantee Act, 2005, 2006–2012, Orient BlackSwan

Panagariya A (2011) Avoiding lopsided spatial transformation. In: Ezaz G (ed) Reshaping tomorrow: is South Asia ready for the big leap. World Bank, OUP, New Delhi

Planning Commission (2001) Government of India Report of the Working Group on Rural Poverty Alleviation Programs for the Tenth Five Year Plan (2002–2007)

Rodrik D, Subramanian A (2004) Why India can grow at 7 percent a year or more: projections and reflections. Econ Polit Weekly 39(16):1519–1526

Sarkar S, Mehta BS (2010) Income inequality in India: pre- and post-reform periods. Econ Polit Weekly 37:45–55

Sen A, Himanshu (2004) Poverty and inequality in India—II: widening disparities during the 1990s. Econ Polit Weekly 39:4361–4375

Sen A, Dreze J (2011) Putting growth in its place: it has to be but a means to development, not an end in itself. Outlook, 14 Nov 2011

Somanathan E, Somanathan R (2009) Climate change: challenges facing India's poor. Econ Polit Weekly 44(31):51−58

Subbarrao K (2003) Systemic shocks and social protection: role and effectiveness of public works programs. Social protection Discussion Paper Series. World Bank, Washington D.C

Svedberg P (2012) Reforming or replacing the public distribution system with cash transfers? Econ Polit Weekly 7:53–62

Wallack J (2003) Structural breaks in Indian Macroeconomic data. Econ Polit Weekly 38(41):4312–4315

Chapter 19
Inequality, Public Service Provision, and Exclusion of the Poor

Sukanta Bhattacharya

Introduction

The rhetoric of inclusive growth has no doubt caught the imaginations of development planners and policy-makers almost everywhere in the world. Recent policy and research documents from various institutions, ministries, individual researchers are abound with mention of inclusive as the most important ingredient for development. In India, the country which has observed sustained high growth of GDP for almost a decade now, every recent policy document has a reference to the target of making the growth pro-poor and inclusive. This continuing fascination about inclusive growth on the part of the researchers and policy-makers indicates that the economics community now recognizes the fact that the market-driven growth in some developing countries is not enough for development. In fact, there are now several evidences from individual countries which show that growth elasticity of poverty reduction decreases with high growth. Moreover, that higher income inequality reduces the poverty impact of growth is also well-documented. So inequality enhancing high growth not only may lead to adverse poverty impacts, but also has the potential to reduce the poverty impact of future growth.

Obviously, one way to counterbalance this adverse poverty impact of the top-end high growth is through redistributive mechanisms of the governments. The governments often use progressive tax structure for financing the provision of public services meant for all to achieve this end. R. H. Tawney (1964), back in 1931, discussed the growth and significance of public provision for education, health, and social services, and noted that "the standard of living of the great mass of the nation depends, not merely on the remuneration which they are paid for their labor, but on the social income which they receive as citizens" (1964, 133). But, the problem is that, in this era of globalization, the governments of the developing economies are

S. Bhattacharya (✉)
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: sukanta_70@yahoo.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_19,

increasingly shying away from the domain of provisioning of public services and thereby unwittingly losing out on a very important instrument for redistribution.

Many academicians, as well as development practitioners, forcefully argue that a policy geared toward private provisioning of the public services is good not only for the wealthy who can obtain better services by paying for it, but also for the poor because it reduces the pressure on the existing public facilities thereby ensuring better quality for those who continue to use them. This argument, at best, ignores several crucial forces that are very important in determining the ultimate outcome of such a policy. Growth, when it increases the incomes of the rich disproportionately, induces social segregation. For the public services that are excludable, the rich, getting dissatisfied with the quality of public services provided by the government, organize their own clubs to make provision of superior quality services accessible only to the club members or preach for private provisioning. Each club is a kind of social closure which is formed by the rich to exclude the less fortunate individuals from some common enterprises or social communities. If such clubs are formed to provide different public utilities like electric supply, water supply, supply of health services or education facilities, then eventually the club members will seize to participate in the public domain and this withdrawal will lead to a general deterioration in the standard of the common pool services available to the mass from public sources. Since the richer section of the society has more options and, therefore, stronger voices, so the non-participation of the affluent in the public sphere would adversely affect the accountability of the system. The poorer section does not have many alternatives to fall back and thus the level of welfare of the poor people is adversely affected. This is clearly a demand-side impact.

In this context, Graham (1998) noted that in many Latin American countries, the second stage of economic reforms aims to change the institutional framework ranging from public finance to social security, health care, and education. Provision of these services no longer belongs exclusive to the public domain. However, the differential impacts of the choice of institutions for provision of social services on different segments of the society have hardly been studied. Blank (2000) and Ireland (1990) considered models in which each consumer chooses to consume either social or private goods and a vertical market segmentation results from differential demand for quality. In the context of this choice, the role of competition and the use of incentives have further been discussed in Besley and Ghatak (2003).

On the supply-side, once the rich withdraw from the domain of publicly provided services, even a benevolent government would readjust the tax structure that is meant primarily for raising revenues to provide for these services. One can clearly see that in high growth developing countries, the governments generally tend to lower the tax rates as income growth takes place. The rich, being more organized, can effectively lobby for lowering of the tax rates once they start buying public services from private facilities. Once this tax restructuring takes place, this may also have an adverse impact on the quality of public services available for the poor because now there is a trade-off between less crowding of the public facilities and less revenue for running those same facilities. Of course, the net impact is

ambiguous and the trade-off can be resolved either way. The claim that less crowding of public facilities will definitely make the poor better-off is over-ambitious at best and not grounded on any solid theoretical justification.

The existing literature on exclusion mostly blames the backwardness of institutions for the phenomenon and prescribes the institutional reforms to correct it. In this paper, we argue that even the modern market-based institutions that evolve with increasing inequality may cause exclusion. What is more worrying is the fact that a benevolent government's optimal policy response may not be enough to tackle this problem. It calls for thinking along a different line to determine the policy course that the government should adopt in a growing economy. This is a new area and extensive future research is required to properly understand the problem and possible responses from the policy-makers.

The differential impacts of the choice of institutions for provision of social services on different segments of the society have hardly been studied. A recent paper by Gulati and Ray (2011) have shown the possibility of exclusion from social provision for the poorer section of the society in presence of acute income inequality. Empirically, the effect of inequality on the overall quality of public good has been studied by Motiram and Nugent (2007). Osberg et al. (2004) discussed the relation among inequality, public goods, and social mobility in a survey.

In this chapter, we propose a model of public service provision that attempts to address some of these issues. We also make a preliminary attempt to link it to the debate around the inclusive versus non-inclusive growth that presently permeates even into the world of mainstream academic economics. The paper is organized in the following manner. In section "The Model", the basic model is set up. Section "The Nash Game" tells the possible impacts of different types of income growth. Section "Concluding Comments" concludes.

The Model

We consider an economy consisting of two types of people—poor with income Y_1 and rich with income Y_2 where $Y_2 > Y_1$. The total number of poor people in the economy is N_1 while the number of rich people is N_2. Furthermore it is assumed that the poor and rich form a constant fraction of population. So, $N_1 = (1 - \delta)N, N_2 = \delta N$ where $0 < \delta < 1$. The rich people's income is a multiple β of the income of poor, i.e., $Y_2 = \beta Y_1$ where $\beta > 1$. The preference of a typical individual over private good and public service is

$$U = x^{\alpha} g^{1-\alpha} \tag{19.1}$$

Here g represents the quality of the public service. Higher the value of g, better is the quality of public service that is provided. Better quality of a public service would mean a public school with a better infrastructure, a better health care facility with access to all modern equipments and properly trained staff. In this model, the

decision about the quality of public service is purely supply determined. The choice variable at the hands of the government is tax rate and given the tax rate, g is determined by the amount of tax revenue.

We assume that the cost of providing public service of a higher quality is also higher. The cost function for providing public service of quality g to m beneficiaries is given by

$$C = F + mg$$

where F is the fixed cost of provision of public services. We also assume that the marginal cost of increasing quality increases with the number of beneficiaries. The tax revenue from a proportional tax at rate t is given by $t(N_1Y_1 + N_2Y_2)$. Hence given t, g can be determined from $t(N_1Y_1 + N_2Y_2) = F + (N_1 + N_2)g$ under pure public provision. For analytical simplicity, throughout the paper we will use a particular characterization of F. We assume $F = \gamma(NY_1)$ where $\gamma < 1$ is the parameter indicating the level of fixed cost.

The individual budget constraint is given by $x = (1 - t)y$, where t is the tax rate chosen by the government and y is the income of a representative individual. Government chooses t to maximize social welfare $W = N_1V_1 + N_2V_2$ where V_i is the ith group's indirect utility. The resulting choice is

$$\begin{aligned} t^* &= (1-\alpha) + \frac{\alpha}{1-\delta+\delta\beta} \cdot \frac{F}{NY_1} \\ &= 1-\alpha + \frac{\alpha\gamma}{1-\delta+\delta\beta} \end{aligned} \tag{19.2}$$

Accordingly, the optimal quality of public service is given by,

$$\begin{aligned} g^* &= (1-\alpha)\left\{[(1-\delta)+\delta\beta]Y_1 - \frac{F}{N}\right\} \\ &= (1-\alpha)\{(1-\delta)+\delta\beta-\gamma\}Y_1 \end{aligned} \tag{19.3}$$

The Nash Game

This economy consists of two distinct groups with different average incomes. The rich may not be satisfied with the quality of public service that the government provides and may desire a better quality. Suppose now that rich form a club among themselves for providing their desirable quality of public service on its own. The club chooses a tax t_c on top of the government tax t.

The club's desirable quality of public service is again supply determined and is given by $g_2 = \frac{t_c N_2 Y_2}{N_2} - \frac{F}{N_2} = \left(t_c\beta - \frac{\gamma}{\delta}\right)Y_1$. In presence of the club, there is now a difference in the quality of public service that the government provides since the

government provides only for the poor. It is now given by $g = \frac{t(N_1Y_1+N_2Y_2)}{N_1} - \frac{F}{N_1} = \frac{t[(1-\delta)+\delta\beta]-\gamma}{(1-\delta)}Y_1$.

The government and the club play a simultaneous Nash game to choose the taxes. Club takes t as given and chooses t_c to maximize

$$V_2 = ((1-t-t_c)Y_2)^{\alpha}\left(\left(t_c\beta - \frac{\gamma}{\delta}\right)Y_1\right)^{1-\alpha} \tag{19.4}$$

The best response for the club is given by

$$\hat{t}_c = (1-\alpha)(1-t) + \frac{\alpha\gamma}{\delta\beta} \tag{19.5}$$

Government on the other hand takes t_c as given and chooses t to maximize

$$W = N_1[((1-t)Y_1)^{\alpha}(g)^{1-\alpha}] + N_2[((1-t-t_c)Y_2)^{\alpha}(g_2)^{1-\alpha}] \tag{19.6}$$

where

$$g = \frac{t[(1-\delta) + \delta\beta] - \gamma}{(1-\delta)}Y_1$$

and

$$g_2 = \left(t_c\beta - \frac{\gamma}{\delta}\right)Y_1$$

The first-order condition for the government's optimization problem is

$$N_1Y_1^{\alpha}\left[(1-\alpha)(1-t)^{\alpha}g^{-\alpha}.\frac{\delta g}{\delta t} - \alpha(1-t)^{\alpha-1}g^{1-\alpha}\right]$$

$$-N_2Y_2^{\alpha}\alpha(1-t-t_c)^{\alpha-1}(g_2)^{1-\alpha} = 0$$

Plugging in $\hat{t}_c = (1-\alpha)(1-t) + \frac{\alpha\gamma}{\delta\beta}$, and after some manipulations we obtain

$$\frac{1-\alpha-t+\frac{\alpha\gamma}{1-\delta+\delta\beta}}{(1-t)^{1-\alpha}\left(t-\frac{\gamma}{1-\delta+\delta\beta}\right)^{\alpha}} = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\frac{\frac{\delta\beta}{1-\delta+\delta\beta}}{\left(\frac{1-\delta}{1-\delta+\delta\beta}\right)^{\alpha}} \tag{19.7}$$

The above equation solves the equilibrium tax rate for the government under a mixed system of provision of public services. Note that the right-hand-side of (19.7) is independent of t. The left-hand side is a decreasing function of t for[1] $t \in \left[\frac{\gamma}{1-\delta+\delta\beta}, 1-\alpha+\frac{\alpha\gamma}{1-\delta+\delta\beta}\right]$. Moreover, within the said interval the left-hand side goes from $+\infty$ to 0. Thus existence of an equilibrium t is guaranteed. We denote the Nash equilibrium value of t as $\hat{t}$. Notice that the right-hand side of (19.7) of is

[1] For all $\beta > 1$, $\frac{\gamma}{1-\delta+\delta\beta} < 1-\alpha+\frac{\alpha\gamma}{1-\delta+\delta\beta}$ can be easily verified.

increasing in β while the left-hand-side is falling. Thus, the equilibrium tax rate, $\hat{t}$, is falling in β.

Incentive for Club Formation

A club will be formed only when the indirect utility of the rich with the club is in place is greater than the indirect utility of the rich in absence of the club.

Definition 1 $\bar{t}$ is defined as the maximum government tax rate under which the rich have the incentive for club formation.

Lemma 1 $\bar{t}(\beta,\delta,\gamma) = 1 - \left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha} - \frac{\gamma}{\delta\beta}\left\{1 - \delta\left(\frac{\beta}{1-\delta+\delta\beta}\right)^{\alpha}\right\}$

Proof Mathematically, the condition for club to exist, can be expressed as

$$((1-\hat{t}-\hat{t}_c)Y_2)^{\alpha}(\hat{g}_2)^{1-\alpha} \geq ((1-t^*)Y_2)^{\alpha}(g^*)^{1-\alpha} \tag{19.8}$$

Putting the values of $\hat{t}_c$, $\hat{g}_2$ and t^*, g^* in (19.8) and after canceling terms, we obtain

$$\begin{aligned} \left(1-\hat{t}-\tfrac{\gamma}{\delta\beta}\right)\beta^{1-\alpha} &\geq \tfrac{1-\delta+\delta\beta-\gamma}{(1-\delta+\delta\beta)^{\alpha}} \\ \Leftrightarrow \quad \hat{t} &\leq 1 - \left(\tfrac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha} - \tfrac{\gamma}{\delta\beta}\left\{1-\delta\left(\tfrac{\beta}{1-\delta+\delta\beta}\right)^{\alpha}\right\} \\ &= \bar{t}(\beta,\delta,\gamma) \end{aligned}$$

where the last equality follows from definition of $\bar{t}$.

$\bar{t}$ is the highest level of the tax rate under which a club will be formed. In other words if the tax rate chosen by the government is less than or equal to $\bar{t}$ then the rich people will agree to charge a tax t_c on top of it and form a club. Note that $\frac{\delta\bar{t}}{\delta\beta} > 0$ implying that $\bar{t}$ increases with β. Moreover, $\lim_{\beta\to 1}\bar{t} < 0$ which indicates that for low levels of inequality the formation of the club is not incentive compatible. However, as $\beta \to \infty$, $\bar{t} \to 1-\delta^{1-\alpha} > 0$. Hence, there exists a critical $\beta > 1$ at which $\bar{t} = 0$. We denote this critical β by β_0.

A club is formed for the provision of public service for the rich whenever $\hat{t} \leq \bar{t}$. We now present our first proposition which shows that given sufficient inequality the rich withdraws from the public system and form their own club for provision of public service.

Proposition 1 Given any $\delta, \gamma \in (0,1)$, there exists a critical threshold $\beta_c > \beta_0$ such that for all $\beta \geq \beta_c$ the rich have the incentive to form a club for provision of the public service.

Proof Notice that as $\hat{t}$ decreases with β continuously and $\lim_{\beta\to\infty}\hat{t}=0$. On the other hand $\bar{t}$ continuously rises with β and $\lim_{\beta\to\infty}\bar{t}=1-\delta^{1-\alpha}>0$. Hence, there exists some $\beta_c\in(\beta_0,\infty)$ such that for all $\beta\geq\beta_c$, $\hat{t}\leq\bar{t}$.

The result seems very intuitive. As discussed in the introduction, the emergence of a club catering to one section of the society to provide a public service of their desirable quality often occurs when the average income of this group is significantly above the average income of the rest.

The Quality of the Public Service for the Poor

When the rich form their own club for public service provision, the quality of the public service meant for the poor falls. The government maximizes the aggregate utility and the tax collected by the government does not benefit the rich once they provide for their public service through their own club. As a result, the government reduces the tax rate once the club is formed. This has a negative impact on the quality of the public service provided by the government for the poor. On the other hand, since the tax proceeds for the government are spent to provide public service only for the poor and not the rich, there is a positive effect on the quality as well. As it turns out, whenever the rich have incentive to form the club, the first effect dominates the second effect and the poor receive a lower quality. This is spelled out in the next proposition.

Proposition 2 Given any $\delta\in(0,1)$, $\beta\geq\beta_c$ implies lower quality of public service for the poor under the mixed system relative to that under pure public provision.

Proof With the club in place, the quality of the public service received by the poor is $\hat{g}$ where

$$\hat{g}=\frac{\hat{t}[(1-\delta)+\delta\beta]-\gamma}{(1-\delta)}Y_1$$

Under pure public provision, the said quality is g^* where

$$g^*=(1-\alpha)\{(1-\delta)+\delta\beta-\gamma\}Y_1$$

Therefore,

$$\begin{array}{lccc} & \hat{g} & \leq & g^* \\ \Leftrightarrow & \frac{\hat{t}[(1-\delta)+\delta\beta]-\gamma}{(1-\delta)} & \leq & (1-\alpha)\{(1-\delta)+\delta\beta-\gamma\} \\ \Leftrightarrow & \hat{t} & \leq & (1-\alpha)(1-\delta)+\left(\frac{1-(1-\alpha)(1-\delta)}{(1-\delta)+\delta\beta}\right)\gamma \end{array}$$

Henceforth, we denote $(1-\alpha)(1-\delta)+\left(\frac{1-(1-\alpha)(1-\delta)}{(1-\delta)+\delta\beta}\right)\gamma$ as $\tilde{t}$. Thus the quality of public service received by the poor falls if and only if $\hat{t}<\tilde{t}$. We show that $\hat{t}\leq\bar{t}\Rightarrow\hat{t}<\tilde{t}$, i.e., if there is enough incentive for club formation for the rich, then the quality of the public service received by the poor falls.

First suppose that $\alpha\beta\leq 1$. In this case, $\frac{1-\delta+\delta\beta}{\beta}=\frac{1-\delta}{\beta}+\delta\geq\alpha(1-\delta)+\delta$ since $\alpha\leq\frac{1}{\beta}$. Hence,

$$\begin{array}{rccccc} & \frac{1-\delta+\delta\beta}{\beta} & \geq & 1-(1-\alpha)(1-\delta) & & \\ \Rightarrow & \left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha} & \geq & (1-(1-\alpha)(1-\delta))^{1-\alpha} & > & 1-(1-\alpha)(1-\delta) \end{array}$$

since $1-(1-\alpha)(1-\delta)<1$. Thus,

$$1-\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha}<(1-\alpha)(1-\delta)$$

Since

$$\begin{aligned}\bar{t}&=1-\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha}-\frac{\gamma}{\delta\beta}\left\{1-\delta\left(\frac{\beta}{1-\delta+\delta\beta}\right)^{\alpha}\right\}\\&<1-\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha}\end{aligned}$$

for any $\beta>1$,

$$\begin{aligned}&\hat{t}\leq\bar{t}\\&\quad\Rightarrow\hat{t}<1-\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha}\\&\quad<(1-\alpha)(1-\delta)\\&\quad<(1-\alpha)(1-\delta)+\left(\frac{1-(1-\alpha)(1-\delta)}{(1-\delta)+\delta\beta}\right)\gamma\\&\quad=\tilde{t}\end{aligned}$$

Thus, $\hat{t}\leq\bar{t}\Rightarrow\hat{t}<\tilde{t}$.

Now consider the case $\alpha\beta>1$. Suppose $\hat{t}\leq\bar{t}$, but $\hat{t}>\tilde{t}$. Then,

$$\begin{aligned}&L(\hat{t})<L(\tilde{t})\\&\quad=(1-\alpha)^{1-\alpha}\frac{\delta}{(1-\delta)^{\alpha}}\cdot\frac{1}{(1-(1-\alpha)(1-\delta))^{1-\alpha}}\end{aligned}$$

since $L(t)$ is decreasing in t. But $\hat{t}$ solves 7 and hence

$$L(\hat{t}) = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\frac{\delta\beta}{(1-\delta)^{\alpha}}\cdot\frac{1}{(1-\delta+\delta\beta)^{1-\alpha}}$$

$$= (1-\alpha)^{1-\alpha}\frac{\delta}{(1-\delta)^{\alpha}}\cdot\frac{(\alpha\beta)^{\alpha}}{\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha}}$$

Plugging this in the last equation and canceling terms, we find that

$$\frac{(\alpha\beta)^{\alpha}}{\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha}} \leq \frac{1}{(1-(1-\alpha)(1-\delta))^{1-\alpha}}$$

or,

$$\frac{\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha}}{1-(1-\alpha)(1-\delta)} \geq \left[\frac{\alpha\beta}{1-(1-\alpha)(1-\delta)}\right]^{\alpha} > 1$$

where the last inequality follows from $\alpha\beta > 1$ and $1-(1-\alpha)(1-\delta) < 1$. Hence, $1-\left(\frac{1-\delta+\delta\beta}{\beta}\right)^{1-\alpha} < (1-\alpha)(1-\delta)$ which in turn implies $\bar{t} < (1-\alpha)(1-\delta)$. This is a contradictory to $\bar{t} \geq \hat{t} > \tilde{t}$. Therefore, $\hat{t} \leq \bar{t} \Rightarrow \hat{t} < \tilde{t}$.

The last proposition states the main result of the paper. Once the rich stop participating in the public system of provision, the poor who are compelled to consume the service provided by the public system receive lower quality instead of less crowding at the public facilities. In presence of a club, the government knows that the public service it is providing is not serving one section of the population. Hence a lesser benefit of public service provision is now internalized. As a result, government would have the incentive to produce a lower quality of public service. For any value of $\alpha < 1$, it can be shown that the poor's utility also falls under the mixed system of public service provision relative to pure public provision once income inequality is sufficiently high.

Pattern of Growth and Quality of Public Service

In this section, we consider two different types of income growth. One type of income growth enhances β keeping δ constant, i.e., due to the growth, the income gap between the rich and the poor widens. This type of growth is called non-inclusive.[2] The other type of income growth increases δ keeping β constant, i.e.,

[2] It may be clarified here that non-inclusive growth in our model implies increasing inequality between the incomes of the rich and the poor. It does not necessarily imply that the income of the poor remains stagnant or decreases over the process of growth. In the scenario characterized by

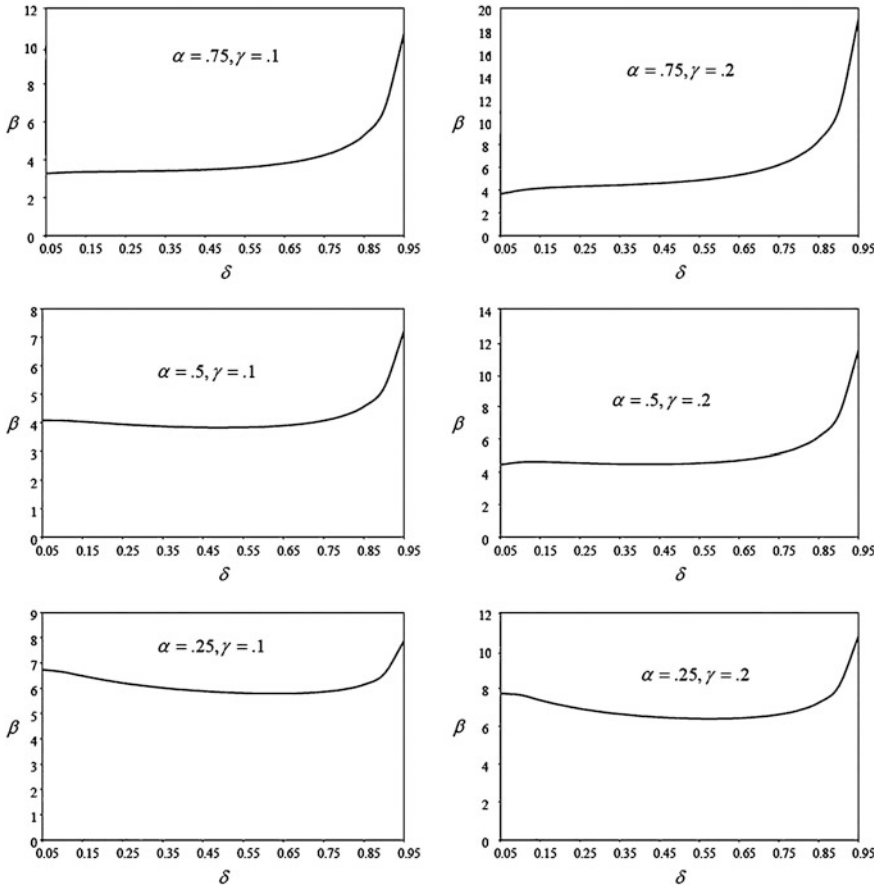


Fig. 19.1 Simulation pictures:Critical β-δ frontier for different parameter values. *Note* The rich have incentive for club formation for every β-δ configuration above the schedule

this type of growth makes some of the poor rich while maintaining the same income gap between the rich and the poor. This is called inclusive growth in the sense that the benefit of growth accrues to the people belonging to the poorer section of the population.

We have already shown in the last section that whenever the rich provide for themselves, the quality of the public service that the government provides for the rest of the population falls. The rich form their own club whenever $\hat{t}(\beta, \delta, \gamma) \leq \bar{t}(\beta, \delta, \gamma)$ for any given value of preference parameter α in $(0, 1)$. Using simulation, we try to trace out the critical (δ, β) frontier for different values of γ

(Footnote 2 continued)
non-inclusive growth, the level of poverty as measured by the percentage of people below poverty line remains same in our model.

and α above which the club will be formed. Except β, the other parameters are fractions in $(0, 1)$ interval, while $\beta > 1$ is our measure of income inequality. The simulation results are reported in Fig. 19.1.

The critical (δ, β)-frontiers are U-shaped in all the cases. If we start at any $(\delta - \beta)$ combination above the frontier, the rich form their own club for public service provision. Now if δ starts increasing, there may be an incentive for switching back to pure public provision beyond a certain level. Notice that the rich want to form their club for provisioning of the public service when they are dissatisfied with the quality of the public service that the government provides. The government supplies an average quality by collecting taxes from the rich and the poor. As the income of the rich rises, the demanded quality also rises. The shortfall of the average quality from the demanded quality increases further and beyond a certain level, the rich have enough incentive to provide for themselves even after paying the government-designated tax if they foresee that the government will adjust the tax rate downwardly once they walk out of the public system.

As the number of poor decreases sufficiently, the average quality under pure public provision is still lower than the rich enjoy under the mixed system, but the shortfall gets lower and lower as the number of rich increases while by switching back to pure public provision the rich can save the fixed cost that they incur for their own provision. For sufficiently high values of δ, the trade-off is resolved in favor of switching back. Once the rich switch back to pure public provision, the remaining poor are also better-off. This is not the case if as a result of growth β continues to rise. The rich walk out of the public system once the frontier is crossed and never switch back to pure public provision however large β becomes.

This helps us in relating the nature of income growth to its impact on the quality of service that the poor receive. If only the income of the rich increases (characterized in terms of increasing β in our model), the quality of service that the poor receive falls once the rich start providing for themselves. The effect is permanent in the sense that there is never any switching back by the rich and the quality of the service received by the poor remains lower than that under pure public provision. On the other hand, if the income growth is inclusive in nature, i.e., as a result of the income growth some people from the group of poor become rich (characterized in terms of increasing δ in our model), the utility of the remaining poor may fall if the club is formed, but the impact is transitory. As income growth continues and δ rises beyond a critical level (at a given β), the rich would switch back to the pure public provision and the remaining poor will become better off.

Concluding Comments

In this chapter, we propose a model where the marginal utility of quality of public services rises with income. As a result, given the same proportionate tax rate on income, the rich's demand for better quality public services is higher. We show that if the government chooses the tax rate by maximizing a utilitarian social

welfare function, the rich would have incentive to form their own club for self-provision of public services despite being compelled to pay the government taxes if their income rises significantly above the poor. We also show that once the rich form the club, the quality of public services for the poor is lower than that under the pure public provision. This may be one of the reasons behind the popular discontent among the people, particularly the poor, about the quality of public services in many developing countries in recent times despite high per capita income growth.

We also argue that if income growth continues to expand the divide between the incomes of the rich and the poor, then the poor's utility falls at some stage after the rich arranges for their own provision of public services. Moreover, this effect is permanent and the poor's utility will go on falling as the income growth of this nature continues. On the other hand, if the income growth is inclusive in the sense of making some of the poor rich, the poor's utility may fall once the club is formed by the rich, but this decline is transitory. Once enough poor people become rich as a result of the income growth, the rich would switch back to pure public provision and from then on the remaining poor's utility would start increasing with income growth.

There are several interesting extensions of the model that may be considered for future research. Instead of a proportionate tax on the rich and the poor at the same rate, the implications of a progressive tax structure may be examined. One may also consider several income classes which we suspect would lead to multiple club-formation and supply of several quality levels for the public services. It is also possible to bring in demand-side effects on the quality of public services. Typically, in a democratic setup the government may choose the tax rates with an eye to reelection possibilities. In that case, the median voter's utility function will receive higher weight in the Government's objective function. On the other side, the rich generally have better lobbying power with the Government and thus can influence the policy-making process in their favor more than the poor. All these potentially would affect the equilibrium and the ultimate outcome in a richer model.

References

Besley T, Ghatak M (2003) Incentives, choice, and accountability in the provision of public services. Oxf Rev Econ Policy 19:235–249

Blank RM (2000) When can public policy makers rely on private markets? the effective provision of social services. Econ J 110:C34–C49

Graham C (1998) Private markets for public goods—raising the stakes in economic reform. Brookings Institution Press, Washington

Gulati N, Ray T (2011) Inequality, neighbourhoods and welfare of the poor. Discussion Paper 11–07, Indian Statistical Institute, Delhi

Ireland NJ (1990) The mix of social and private provision of goods and services. J Public Econ 43:201–219

Motiram S, Nugent JB (2007) Economic and political inequality and the quality of public goods. Int J Dev Issues 6:142–167

Osberg L, Smeeding T, Schwabish J (2004) Income distribution and public social expenditure: theories, effects and evidence. In: Neckerman K (ed) Social inequality. Russell Sage Foundation, NY, pp 821–859

Tawney RH (1964) Equality. London: Unwin Books, first published in 1931, 4th edition

Chapter 20
Stakeholder Attitudes and Conservation of Natural Resources: Exploring Alternative Approaches

Biswajit Ray and Rabindranath Bhattacharya

Introduction

Individuals make decisions embedded in a social context and their attitudes affect their decisions. Based on this argument, there is a long history of empirical research in social sciences (beyond economics), that elicits subjective testimony on feelings, beliefs, values, expectations, plans, attitudes, and behavior. This body of empirics was excluded from the neoclassical economic analysis on the assumption that individual preferences remain unchanged, despite the fact that economic theorizing often includes reference to attitudes, beliefs and the like. An important example is the data on stakeholder attitudes toward environment (Infield 1988). Its importance has been documented recently (Agrawal 2006) and pro-environmental attitudes studies have started to catch the attention of economists.

Following Ajzen and Fishbein (1977), attitude is a human psychological tendency, which people express by evaluating a particular object with favor or disfavor, or, for example, in our problem at hand, like or dislike of community-based forest management. Attitude consists of beliefs, which are associations people establish between the attitude object and various attributes. For example, in the phrase, 'a reserve forest is part of a country's wealth,' 'reserve forest' is the attitude object, 'country's wealth' is an attribute, and 'is part' is a relational term.

However, attitudes cannot be observed directly and have to be inferred from overt responses. The techniques of attitude measurement can be broadly organized into direct self-report methods (e.g., interviews and questionnaires) and implicit

B. Ray (✉)
Department of Economics, University of Calcutta, Kolkata, West Bengal 700050, India
e-mail: biswajit_roy1002@yahoo.com

R. Bhattacharya
UGC DRS-1, Department of Economics, University of Calcutta and Center for Studies in Social Sciences Calcutta (CSSSC), Kolkata, India
e-mail: rnbeco@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_20, © Springer India 2013

measurement techniques (e.g., observation, priming and response competition measures). Studies measuring environmental attitudes generally use direct self-report methods, and much less frequently implicit techniques.

In an attempt to organize the measures of environmental attitudes, Dunlap and Jones (2002) have proposed a 4-fold typology of measures based on environmental issues (e.g., water pollution, population growth) and expression of concern (i.e., beliefs, attitudes, intentions, and behaviors related to environmental issues). Despite the possibility of large number of such measures, only three have been widely used. These are the Ecology Scale (Maloney and Ward 1973), the Environmental Concern Scale (Weigel and Weigel 1978), and the New Environmental Paradigm (NEP) Scale (Dunlap and Van Liere 1978; Dunlap et al. 2000). These three scales examine multiple phenomena or expressions of concern, such as beliefs, attitudes, intentions, and behaviors. They also examine concerns about various environmental issues, such as pollution and natural resources degradation.

Given these measures of attitudes toward environment, one pertinent query in resource and environmental economics is whether environmental attitudes are actually reflected in environmental behaviors. While Shogren and Taylor (2008) believe that economics is all about "behavior," considerable debate has surfaced in the social psychology literature regarding the attitude–behavior consistency (Ajzen and Fishbein 1977). In most cases, these two attributes are not consistent and consistency of attitudes and behaviors are conditional (Ajzen and Fishbein 1977). One reason might lie in the structure of stakeholder attitudes (Milfont and Duckitt 2010). In the context of public goods provision like co-management of natural resources, these reasons may vary from the very prevalence of free-riding behaviors of the stakeholders to their social attributes like their social status and identity. The evidence that women are better rule-makers than men (Agarwal 2009) and that organizations with fair leaders are more successful than those with the authoritarians (Sinha and Suar 2005) may testify that identity and status matter in our understanding of the attitude–behavior debate, at least in the context of community-based natural resource management.

Against this backdrop, the importance of understanding stakeholder attitudes in economic setting has been growing. When it comes to the question of survival of community organizations, it appears to be more pertinent. Without pro-environmental attitudes in collective behaviors, no organizations can sustain in the long-run (Ostrom 1990). Negative attitudes entail high transaction costs, leading to the decline of an organization. Failure of the Joint Forest Management in some parts of India resonates this argument. On the other hand, the 1000-years-old Spanish Irrigation system is a classic example of how stakeholder attitudes contribute to the sustainability of organizations.

Interdisciplinary studies including social psychology and behavioral economics attempt to analyze the mentalities of the stakeholders. Often they challenge the existing ideas of rational choice and lay the foundations of new theories and methods of understanding human attitudes and behaviors. It is in this context that this article assesses two popular methods of measuring stakeholder attitudes and behaviors in community resource management. The first approach is referred to as

the scaling procedure, and the second one is the procedure of behavioral social sciences, based mainly on experimental economics paradigm.

To induce subjective data theoretically more robust, this article invokes two interaction theories, namely the *expectation states theory* and *social identity theory*, and compares them with the game theory, the foundation of behavioral economics. In this way, it complements the existing literature. To understand stakeholder subjectivity we consider women and social status as the two most important attributes of rural forest dependent households and highlight a case study of seven community forests in the state of West Bengal, India to illustrate how the two approaches work in practice. It may be noted that our effort is to understand the methodological issues involved and not the analytical dimensions of community forestry.

The structure of the article is as follows. The next section deals with how scales are employed in social sciences, mainly emphasizing the subject-centered scales pertinent to our natural resource conservation context. Since scales measure stated preferences of stakeholders, knowing actual preferences of the people becomes meaningful. Hence, we describe the experimental method, which captures the actual attitudes or behaviors in section three. Theoretical predictions are made in the fourth section to make the discourse complete. "Understanding Stakeholder Attitudes and Behaviors: A Field Study" presents a case study of community forestry in section five to help us understand how these methods can be employed. Finally, we conclude.

Scale-Based Approach

Scales are unique features of any perception studies. A scale is a collection of words/phrases/statements called items, the responses to which are scored and combined to yield a scale score. Scale scores can be categorized according to level of measurement. At the nominal level of measurement, scale scores are used to assign the classification categories to which the objects of measurement are grouped. At the ordinal level, scale scores rank order the measured objects along the classificatory dimension. At the interval level, scale scores reflect the relative distances between and among measured objects. At the ratio level, scale scores indicate the absolute distance of any measured object from a true-zero point on the scale.

Scales are classified, according to the source of scale score variation, as *stimulus centered*, *subject centered*, or *response scales* (Netemeyer et al. 2003). For the purpose of this chapter, the term scale will be limited to those instruments that are constructed by any researcher to obtain quantitative data on psychological constructs (e.g., attitudes) for which appropriate standardized instruments are not available. Since these variables are latent in nature, to quantify them, we construct different types of scales depending upon our objectives of study. Typically, scales provide the participant with a limited number of response options. Hence, they are

in close-ended formats. For measuring stakeholder attitudes in the context of natural resource management, the scale construction may involve the following stages:

Scale Design

Designing a scale requires, first, some theoretical background, so that we can consider (or construct) a well-defined psychological variable. Such constructs may be, for instance, co-management or leadership quality. Definition of the variable depends on the larger theory that propels the research. Definition includes distinctions (what the variable is and what it is not), dependencies (how the variable is a function of more basic or previously defined terms), and relations (how the variable is related to other variables). How the variable is to be measured depends on a number of considerations, such as how best to represent the variable, who the respondents will be, the context and conditions under which the measure is to be administered, and the research design of the study, especially the analyses planned. In short, the theory of the scale should give the researcher directions as to scale content (the writing of items/statements) as well as scale format (the type of scale to construct).

Scale Content

A useful preliminary to item writing is to conduct open-ended interviews with representative subjects from the target respondent population, which can elicit a wide range of statements about the variable in question. The interviewee's own words can then be used in writing the items. Such use can provide a degree of authenticity that in turn can contribute to the scale's validity. It will also mean that readability of the scale will be less of a problem. Further, it is good to check on the readability level of the scale to make sure that it is appropriate to the level of the respondent population.

The initial pool of items can be written to be homogeneous/unidimensional or heterogeneous/multidimensional in content. The scale design should indicate what degree of content heterogeneity is desired, based on the theory of the scale.

Scale Format

Scale format consists of item(s) and the response choices. Items may consist of full sentences, phrases, or even single words, describing some attribute of an object (e.g., "The leader of the Joint forest Management Committee appears trustworthy").

Alternatively, they may be the state of the object ("The leader is active"), or some event involving the object ("The leader is reflecting the members' forest resource needs"), having varying degrees of specificity or generality.

Response options in structured verbal scales vary in their underlying measurement dimension (e.g., agree-disagree, like-dislike, important-unimportant). They also vary in response format. Rating response formats differ in the number of scale points (choices) given to the respondents (2-, 3-, or 5- point scales are the most common), and in the way scale points are anchored. Anchors can be words (yes–no, true–false), phrases (strongly agree-strongly disagree), or more extended prose as in behaviorally anchored scales (e.g., Campbell et al. 1973). Rating scales may be anchored at each scale point or only at selected scale points (e.g., at the ends and the middle of the scale). Response choices may be unweighted or weighted using multiple weights. Rating response formats may be one-sided (zero to positive or to negative values) or two-sided (with both positive and negative sides of the continuum).

In general, researchers choose the simpler format. When rating response formats are used, more scale points are better than fewer, because once the data are obtained, one can always combine scale points to reduce their number, but one cannot increase that number after the survey is done. Further, more scale points can generate more variability in response, a desirable scale characteristic if the response is reliable. Inordinate use of the middlemost scale point can be avoided by eliminating that scale point, that is, by using an even number of scale points. This has the further advantage of ensuring that the underlying dimension will be linear or can be made linear.

Scale Development

Scale development consists of collecting data with the use of a preliminary form and analyzing the data in order to select items for the final study. A small N pilot study before the main data collection can be used to check out how easily the scale instructions are followed, how well the scale format functions, how long the scale takes to complete, and especially how appropriate the scale items are for the target respondent population. As a rule, the sample should be representative of the target respondent population. There can be exceptions, however; for example, in developing stimulus-centered scales, one could use a sample that is more homogeneous than samples from the target population.

Scaling methods lie at the heart of any scale construction. They are used to select items. They include stimulus centered, subject centered, or response scales. Given the scope of the chapter, we confine our discussion to the subject-centered scaling methods only. Interested readers may consult Netemeyer et al. (2003) for understanding scaling procedures in social sciences.

Subject-Centered Scale Methods

Subject-centered scales are probably the most widely used scales in attitudinal research (Dawis 1986). The classic method for developing subject-centered scales is the Likert method (Likert 1932).

The Likert Method

Likert (1932) employed the psychometric methods to scaling non-ability test items. The Likert procedure can be described as follows:

1. A number of items/statements are written to represent the content domain like co-management. If each statement has five response options, the scale is referred to as five-point scale. It is most frequent in use. Scoring weights from 1 to 5 are assigned to the five-point scale. Direction of scoring (whether 1 or 5 is high) is immaterial provided it is consistent for all items.
2. The items are administered to a large group of respondents (at least 100). Each respondent's item rating choices are scored and the item scores summed to constitute the respondent's total score.
3. Items are selected according to their ability to discriminate between high and low scorers on total score. Likert used a group-difference procedure (difference in item means between high-scoring and low-scoring groups, e.g., uppermost 25 % and lowermost 25 %). One could also use an item-total score correlation procedure, as is currently done in many studies (Liu et al. 2010). Maximizing item—total-score correlation will also maximize the scale reliability.
4. The best discriminating items are then selected to constitute the scale, and the scale score is obtained by summing the item scores for the selected items. At this point, scale scores can be treated as normative scores (i.e., transformed to standardized scores, used to determine percentile equivalents for specific populations, etc.).

Thus to implement the Likert method we require only (a) computing total score, (b) computing item—total-score correlations, and (c) computing scale reliability for the final set of items. Incidentally, reliability should be computed for every research use of Likert scales, not just at scale development, because reliability is a function not only of the scale but also of the respondent sample. Unfortunately, not all scales that are purported to be Likert-type scales are constructed according to the Likert procedure. If, in such scales, the correlation of the items with total scale is not high, then the interpretation of the scale score is problematic.

In this context, we can use factor analysis to select representative items for Likert-type scales (for details see Dawis 1986).

Scale Evaluation

Scales are evaluated primarily based on two criteria—reliability (the proportion of scale score variance that is not error variance) and validity (the proportion of scale score variance that accurately represents the construct or the proportion of criterion variance that is predicted by the scale). These two criteria are complex concepts. In practice, most of the environmental management perception studies examine reliability through Cronbach's alpha (Cronbach 1951).

Reliability and Validity

Different kinds of reliability estimates may be required for different kinds of scales. Because reliability is a function of sample as well as of instrument, it should be evaluated on a sample from the intended target population.

Reliability of a scale is based on true score model of classical test theory (see Netemeyer et al. 2003). If x, T and e be the observed attitude score, the true score and the measurement error respectively, $X = T + e \ldots(1)$.

Assuming that Cov $(T, e) = 0$, Var (x) = Var (T) + Var (e). Reliability is then defined as ρ_{tx} = Var (T)/Var (x) …(2).

Because true score and error are unobservable, we infer about reliability from the observed score reflecting observed attitudes. Broadly, three general types of reliability are computed: (a) test–retest reliability, (b) alternative form of reliability, and (c) internal consistency reliability.

(a) Test–retest reliability is examined by using a single statement over two times. The respondents remain same over time. If the correlation coefficient between the score in time 1 and the scores in time 2 is high (>0.80), respondents' attitudes are said to be reliable. However, the time gap depends on the construct. In case of attitudinal studies, generally 2 weeks are treated as suitable (Robinson et al. 1991). It should be set, so that perfect recall fades away and no major changes (like introduction of development policy) take place in the study environment between the two periods since this policy may affect the respondents' attitudes.
(b) Alternative form reliability is the same procedure as the test–retest with the exception that in this method two different yet related statements are used in the two periods.
(c) Internal consistency is most popular of the three, because the test–retest and alternative form reliability measures are subjects to constraints of time, cost and availability of the respondents at different occasions. There are two methods of measuring internal consistency, namely split-half reliability and coefficient alpha or Cronbach's alpha.

Cronbach's alpha (α) is the most frequently used measure of internal consistency/reliability. It measures the interrelatedness among a set of statements/items

designed to measure attitudes toward a single construct (like conservation). Suppose that the conservation attitudes scale consists of four statements. Computing α is a means of partitioning total score variance into true score (variance across individuals along a single statement) and error component variances. $\alpha = 1$—error variance. The diagonal elements of co-variance matrix represent the variability of a statement across a given set of respondents. $\alpha = 1$—(unique variances of scores/total score variance).

If the statement scores are standardized, then standardized α is computed using the formula: $\frac{kr}{1+(k-1)r}$, where k is the number of statements and r is the average correlations among the statements in the scale. Finally, if we use statements that are dichotomized like the yes/no type of statements, then we use Kuder–Richardson 20 (KR-20) formula to compute coefficient α. $\alpha = \frac{k}{k-1}\left(1 - \frac{\sum pq}{\sigma_x^2}\right)$, where p represents each statement's mean and $q = 1 - p$. Statistical software like SAS, SPSS, BMDP, Stata, or Wessa can be used to compute alpha.

When a scale represents a unique construct like attitude towards conservation, it is referred to as unidimensional. However, if a scale contains statements reflecting more than one dimension, then we need to ensure unidimensionality first, and then compute alpha because alpha does not measure dimensionality of a scale. Problem arises when we sum scores of different scales like co-management scale, scale of attitudes toward organization etc., to arrive at a composite score to represent a unique dimension like community-based conservation. However, this may be feasible if all the statements are unidimensional; otherwise, this may lead to interpretational ambiguities.

It may be noted that as the number of statements increases, alpha tends to increase. However, there is no appropriate scale length. Redundant statements may lead to "attenuation paradox", whereby increasing alpha beyond a certain point does not increase internal consistency. Further, it increases as inter-item correlation increases, and it decreases as a function of multidimensionality. Because of its interdisciplinary uses, a scale with alpha of at least 0.70 and inter-item correlation of 0.30 is useful.

It may be noted that there is a trade-off between reliability and validity. However, scholars prefer reliability to validity (e.g., Mehta and Heinen 2001; Liu et al. 2010). Hence, we skip discussing validity here (See Netemeyer et al. 2003 for details).

Other Issues

There are also other considerations (see Schmitt 1996). For example, the character of the score distribution generated by the scale—in part, is a function of the respondent sample. Most users would prefer a scale that ordinarily produces a reasonably normally distributed set of scores. Scales, like ability tests, can be so

constructed as to produce the shape of score distribution that is desired, by selecting the appropriate items. Another concern is that the scale should produce sufficient score variation to be useful. An old rule of thumb is that the coefficient of variation should be between 5 and 15 %.

In scale construction, there can be no single "best" method. A hybrid approach, tailored to the situation based on purpose, context, and limitations of the researcher, might be better than the other standard approaches discussed here. Finally, a big question: Is a scale necessary? Are there other, less intrusive ways of getting the same information or measurements? To address this we move to the next section.

Experimental Approach

Experimental approach tests many behavioral hypotheses at the core of individual decision making and sometimes replaces the standard assumptions made about the decision makers as well. They highlight some of the lessons that can be learned from conducting behavioral experiments in developing communities. We focus on a topic that has always been at the heart of community organizational studies—pro-social behavior: cooperation in social dilemmas, trust and reciprocity, fairness and altruism.[1]

Pro-social behaviors account for the rise and decline of organizations (Bouma et al. 2008). In social dilemmas, without state support, many communities depend on local norms of trust and reciprocity to substitute for formal institutions and complete otherwise incomplete contracts and rules of conduct to provide public goods. Further, high levels of trust and reciprocity make communities more productive, because they allow people to reciprocally behave in pro-social manner. For example, in community forestry organizations, stakeholders often adjudicate disputes locally through an understanding of the norms of fairness and altruism (Ostrom 1990). Given this importance, we describe experiments related with social preferences below and note the key findings and lessons learned therefrom.

Cooperative Behaviors in Social Dilemmas

Research on the cooperative behaviors of individuals has used three different experiments: the prisoner's dilemma game (PDG), the voluntary contribution mechanism game (VCMG), and the common pool resource game (CPRG). Each game has a particular social dilemma setting for the participants in which one strategy leads to the social optimum, while the dominant/Nash strategy produces a

[1] Chakravarty et al. (2011) provide an excellent review of experimental economics.

socially inefficient outcome. PDG, typically a symmetric two-person game, deals with two strategies: cooperation and defection, the latter strictly dominating the former. In PDG, players confront a normal form matrix with the payoffs measured in the local currency and, have to choose simultaneously 'cooperation' or 'defection'.

VCMG allows players to contribute to a public good, which has the incentive structure of an n-person prisoners' dilemma. This game is repeated for a number of rounds. At the beginning of each round, participants are given an endowment of tokens that they can place in a self-account that only benefits her and a public account that benefits everyone in the group in any combination. The amount contributed to the public account is a measure of the cooperativeness of the participant.

In CPRG, players cooperate by sustainably extracting from a community resource accessible to all players. The resource has rivalry in use—one player's extraction reduces the resources available for others to extract. Since this game usually involves a non-linear public bad, instead of the payoff function participants are presented with a table, which relates their extraction choice to the aggregate choice of the rest of his group members. Players use the table to identify their best response to any extraction taken by the rest of the group.

Evidence shows that only a minority free ride as a first impulse (Cardenas and Carpenter 2008). Approximately one-third of players cooperate in the PDG, contributions of half the endowment to public account are common in the VCMG and extracting only three-quarters of the Nash level appears to be the norm in the CPRG.

Determinants of Cooperation

Several factors affect cooperation depending on the variation in play. The behavioral difference may be age driven (Carpenter et al. 2004b). Allowing participants to socially sanction free riders also enhances cooperation even with some variation in the sanctioning technology (Carpenter et al. 2004a). Similarly, allowing participants to discuss the game between rounds has a similar effect both in CPRG (Cardenas et al. 2002) and in VCMG (Carpenter et al. 2004b). Furthermore, societies where formal institutions are missing or weak develop and rely on prosocial norms and preferences to build informal enforcement mechanisms that make cooperation in these communities sustainable (Ostrom 1990).

Group composition is another important predictor of cooperation. For example, the gender composition of groups affects the level of pro-social behaviors (see Croson and Gneezy 2009 for an overview) and so does the composition of the poor people in the group (Cardenas 2003).

Trust and Reciprocity

The most popular framework of trust and reciprocity experiments is the Berg et al. (1995) investment or trust game (TG). Two players are given a fixed amount of money as a show-up fee. The sender as first-mover can send as much of her stake money (or nothing) to an anonymous recipient as she wishes. The amount sent is tripled before it reaches the receiver to make sending money socially efficient. The receiver then sends back as much as she wishes (or nothing).

No investment is the unique subgame perfect Nash equilibrium. Realizing that the receiver has no incentive to return anything the sender invests nothing in the partnership. Despite this theoretical prediction, Berg et al. (1995) find that, on an average, senders send 50 % of their endowment and trustees return 30 % of what they receive. Extreme cases do exist (e.g., Danielson and Holm 2007). In general, average play is nowhere near the prediction based on egoistic preferences.

There is also a positive relationship between trust and reciprocity. To explain this, we need to have a prior understanding of the norms of reciprocity. When a participant simply repays an investment as if it were an interest-free loan, it is referred to as *balanced reciprocity* as found in the slums of Nairobi (Greig and Bohnet 2005). In contrast, most of the data from developed countries support the norm of *conditional reciprocity* in which the two players consider the relationship more as a partnership in which both players earn profit (Danielson and Holm 2007). Trust and reciprocity are positively correlated in the second case and it turns out that conditional reciprocity is socially more efficient.

It is instructive to note a few methodological issues involved. First, although trustees' behaviors in the TG measures reciprocity; senders' behaviors might reflect both altruism and trust. Senders may see their transfer as a donation without any expectation of a return. To examine this potential altruism, experimenters have measured pure trust following Cox (2004) and running dictator games (DGs) in addition to TGs. In the dictator game (DG), the dictator as a first-mover makes a transfer (even zero transfer is possible) to a recipient, who accepts the amount sent to her. As the game is anonymous, transfers refer to a measure of altruism. Assuming that trust and altruism are additively separable, pure trust is then just the difference between the transfer in the TG and that in the DG. The amount of pure trust is between 21 (Carter and Castillo 2003) and 36 % (Ashraf et al. 2006).

Among the experimenters, a growing consensus is that the standard trust question such as 'Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?' is a better measure of trustworthiness than trust (Johansson-Stenman et al. 2004). This result is a test of validity of the survey trust questions. However, the validity of the measures elicited in the TG is not unambiguous (Karlan 2005).

Finally, since the social capital literature now focuses more on trust and trustworthiness, interest in running the trust games in developing communities has increased. This enhances analyzing the links between the behavioural measures, the survey-based measures like the attitude measures through scaling and

economic outcomes. At the community organizational level, importance of such studies is quite significant. Bouma et al. (2008) and Mukherjee (2011) are two instances of such studies.

Fairness and Altruism

Fairness and altruism refer to distributional norms. In the simplest case, altruism motivates how generously one person must treat another when the second person has no power to control the outcome. These norms govern many philanthropic acts. Things are a little more complicated, however, when the second person can retaliate against perceived injustices. To differentiate the norms that dictate behaviour in these situations, we use the term, 'fairness'. Two simple games measure norms of fairness and altruism. One is the DG. The second is the ultimatum game (UG). In the ultimatum game, two players split a pie between themselves. The proposer offers a share to the responder who accepts or rejects the offer. Accepted offers are implemented; and rejections result in both players receiving nothing. Because responder could do better by accepting lower offers, rejecting should be an empty threat. Knowing this, proposers need not offer more than some small amount.

In the UG, however, most behaviors deviate from the subgame perfect equilibrium (of very small offers) in systematic ways. In each case, the mean allocation to the second person is substantially greater than zero and, in the UG, low offers are routinely rejected which suggests that fairness norms are enforced. Henrich et al. (2001) find that the local payoffs to cooperation and the degree of market integration explain 68 % of the variation in UG offers. Furthermore, societies where teamwork is essential for production exhibit strong sharing norms.

The study by Henrich et al. (2006), which is a replication of the Henrich et al. (2001) study, supports the idea that distributional norms are local phenomena and may vary with local economic conditions and that distributive norms are supported by costly punishment. Generally, those who are fairer are more likely to insist on fairness by others (Cardenas and Carpenter 2008). There is, however, some variation in UG and DG play across communities.

Interestingly, deservingness is a strong predictor of altruism (Durlauf 2002). Implementing trust and dictator games along with a survey Castillo and Carter (2003) show how to exploit the variation in behavior and socioeconomic characteristics at both the individual and group level to identify whether peer effect influences choices.

Finally, literature shows that framing the DG and UG games in a socially embedded manner may be more informative than the standard ones. In the comparative DG, based on a set of relational social distance circles emanating from the individual, proposers are asked to propose an allocation to a family member (the innermost circle), a person in your village (the next circle), and a stranger from a different village (the outermost circle).

Challenges in the Field Experiments

Experimenters confront several challenges in the field. Here we delineate a few of these relevant for our topic.

In the field experiments, participants are non-students individuals, who may not be literate, and accustomed to thinking abstractly and obeying instructions unlike the students in the lab. Hence, making participants understand the game is the most important challenge. Using simple instructions[2] that rely more on pictures, diagrams, examples, and also reading out, and acting out greatly improve the understanding of the participants. In addition, paper and pen experiments are actually preferred to computerized experiments in the field.

Furthermore, participants' belief about the action taken by the experimenter for the purpose of the experiment may influence their decision making (Henrich 2000). Therefore, the experiment must seem more interesting and straightforward to the non-student participants, so that they do not form negative perceptions of the experiment and the experimenter. Accordingly, credibility of both the protocol and the experimenter has to be enhanced by spending considerable time in the studied communities or, by teaming up with local people more credible to the participants.

In order to incentivize participants one should also be sensitive to their needs. In communities that are not well integrated into markets, cash payments are not particularly appealing, and therefore less desirable than payments in durables (e.g., see Pender 1996). Regardless of the incentive type, paying on average one to two days wage for a half-day session seems to create the necessary salience for participants in the field.

In the field experiment, recruitment of participants is much more likely to happen verbally, and therefore peer effects might lead to non-representative sampling. When friends or relatives show up for a particular session and the experimental design is highly sensitive to social ties, one may prefer to assign them to different sessions.

In this context, the employment of appropriate method is equally important: some prefer *strategy method* (Henrich et al. 2006) while others prefer *game method* (Bouma et al. 2008). Consider the following *strategy method* used in a TG in which the investor has a stake size of 100 INR. Trustees are presented with all possible 11 decisions of investment to be made by the sender: (0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100), which when multiplied by three appear before the receiver as (0, 30, 60, 90, 120, 150, 180, 210, 240, 270, and 300). Responders are asked to respond to this set of sending strategies of the sender. However, in game method nothing of this sort is shown to the responder. By strategy method, experimenter can collect more information. However, the potential problem that may arise with the *strategy method* is that many people, especially the rural participants in developing communities do not usually think in terms of strategies.

[2] Most of the lab-based experiments are conducted via z-Tree software (Fischbacher 2007), where pay-offs are constructed accordingly.

Therefore, it is hard for them to think of what they would do in every possible situation, while it is easy to think about what they will do in the situation in which they find themselves.

A final concern is preventing cross talk. Cross talk arises when participants, who have already participated in an experiment talks about the experiment to other participants who have not yet participated. It is a major challenge. Experimenters should keep record of the exact day and time of each session in order to keep track of possible cross talk effects.

One of the most important lessons from this review of experiments might be that pro-social predispositions are heterogeneous and theory needs to account for this fact (Cardenas and Carpenter 2008). The reason might lie in the differences in stakeholder attitudes and behaviors. Hence, to give this treatise completeness with respect to community-based natural resource conservation, below we make theoretical predictions on people's prosocial behaviors.

Theoretical Prediction on Pro-Social Behaviors

Game theory predicts very general trends for people's social preferences like cooperation in public goods setting, but cannot differentiate among different possible levels of cooperation within and between groups in such setting. One reason might be that game theory says little or nothing about where people's expectations that help them reach some levels of cooperation come from, how and why they persist, or when and why we might expect them to arise (Kreps 1990). Accordingly, concept of subgame perfect Nash outcome does not hold in experiments (Cardenas and Carpenter 2008). Consider the one shot trust game of Berg et al. (1995) where game theory predictions on pro-social behaviors are:

1. Individuals/groups stick to their Nash strategy, "invest nothing" that reflects their egoistic behaviors.
2. An individual's pro-social behavior does not depend on her social attributes–whether she is a woman or a man or whether she bears a low or high status. She will always invest nothing.
3. Pro-social behavior of a group does not depend on the composition of the group.

One implication of these predictions is that neither actor's gender nor gender composition of the playing group affects social preferences. However, evidence shows (Table 20.1) that an actor's gender may influence pro-social behaviors in numerous ways.[3] Social scientists argue that beyond economic calculus, social psychological factors like actors' identity and performance expectations may also

[3] Croson and Gneezy (2009) provide a systematic review on gender differences in social preferences.

Table 20.1 Social attributes and pro-social behavior: experimental evidence

Study	Study details	Behavior
Public goods game		*Cooperation*
Sell and Wilson (1991)	Group size = 4, 92 subjects, repeated	Male > Female
Seguino et al. (1996)	Group size = 5–52, 139 subjects, one shot	Female > Male
Sell (1997)	Group size = 4, 244 subjects, repeated	Male > Female
Eckel et al. (2010)	Group size = 4, 80 subjects, 40 rounds in a game of Voluntary Contribution Mechanism—20 with and 20 without punishment, status based on a trivia quiz	High Status people punish less and low status people follow them systematically.
Ultimatum Game		*Rejection*
Eckel and Grossman (2001)	96 subjects, each plays four rounds as proposer and four rounds as respondent	Male = Female (Male as proposer and female as respondents) Male = Female (Female as proposer and male as respondents)
Solnick (2001)	89 proposer/respondent pairings	Male = Female (Male as proposer and female as respondents) Male = Female (Female as proposer and male as respondents)
Dictator Game		*Altruism/Donation*
Bolton and Katok (1995)	$10 as a pie, no offer more than $5	Male = Female
Eckel and Grossman (1998)	Double-blind game with a $10 pie, 140 pairs, no restriction on offer	Female > Male
Trust game		*Trust and Reciprocity*
Ashraf et al. (2006)	Continuous game, Strategy method	Male = Female (trust) Male = Female (reciprocity)
Buchan et al. (2008)	Continuous game, interaction of gender by first name	Male > Female (trust) Female > Male (Reciprocity)
Bellemare and Kröger (2003)	Continuous game, Dutch players	Female > Male (trust) Male > Female (reciprocity)

Note Male = Female implies no significant difference. Male > Female implies more pro-social behaviors of male than female

impinge on their social preferences (Sell 1997; Chen and Li 2009; Eckel et al. 2010). Therefore, integration of complimentary theoretical perspectives with game theory may predict pro-social behaviors better.

Two interaction theories appear to be relevant in our context: *expectation states theory and social identity theory*. We consider these two theories, because they are based on social status and identity. Gender is very much a component of status as well as identity. Therefore, whichever theory is invoked, the group's gender composition may make a difference in the level of cooperation.

Expectation States Theory

People often expect an individual or a group to perform better than others based on certain characteristics and as a result defer more to them. These are referred to as performance expectations. The greater the performance expectation of one actor compared to another, the more likely that actor will be given chances to perform in the group and enjoy high status. Thus, relative performance expectations create relative social standing or status of different actors and maintain a hierarchy of participation, evaluation, and influence among the actors. This is referred to as the "*power and prestige structure*" that constitutes the "status structure" of the group (Correll and Ridgeway 2003).

Based on this status structure in any social setting high-status individuals/groups can influence decisions of the low-status more than decisions of the high status, while low-status people is more likely to influence neither of them. The socially significant attributes of individuals based on which people anticipate the quality of individuals' future task performances are called *status characteristics*. These characteristics are either societally specific (computer expertise) or diffuse (e.g., gender and ethnicity). Specific status characteristics carry cultural expectations for competence at a limited, well-defined range of tasks, and consequently relevant in a limited range of settings. Diffuse status characteristics, which evoke societally defined stereotypes (Sell 1997), on the other hand, carry very general expectations for competence, in addition to specific expectations for greater or lesser competence at particular tasks. They affect performance expectations across a wide range of settings like provision of public goods (Sell et al. 1993).

Expectation states theory (EST) developed by Berger and Fisek (1974) conveys that this relative social status of individuals may influence their pro-social behaviors (see Correll and Ridgeway 2003 for development of the theory). EST argues that women generally are less likely to influence men; and that society also holds, in general, negative expectation about women's task performance. Thus, the theory considers women as possessing the low status compared to men (for discussions see Meeker and Weitzel-O'Neill 1977; Wood and Karten 1986). These observations are confirmed by economic evidence: women are more asset-constrained (Agarwal 2001) and both females and female-headed households are more poor than men (Quisumbing et al. 2001) leading to their low societal status. EST says that when high and low status actors interact in a social setting, relative status hierarchy is activated between them. As a result, the former can influence the latter and the latter defer the former by obeying them. This further encourages

high-status actors to maintain their status by cooperating more toward collective orientation. On the other hand, when high-status actors interact with other actors of the same status, attempt to influence does not work and it leads to a low cooperation. However, low-status people cooperate more when they are in a low-status group. The reason may be that low-status actors are empowered in the absence of the high-status actors and that none of them influences others' decisions. Thus, EST predicts that

1. *High-status people (e.g., Men, rich, and high-caste people) will cooperate less in the group of all high-status actors than in the group containing low-status actors (e.g., women, poor, and low-caste people).*
2. *Low-status actors will cooperate more in groups of all low-status actors than in groups containing high-status actors.*

Is EST applicable in community-based natural resource management (CBNRM)? The theory assumes that the success of the task is important for all group members, whereas public goods settings like community forestry are characterized by mixed motives including free riding. Sell (1997) posits that so long as people consider *influence attempts* and form different *performance expectations* of the stakeholders as found in community-based conservation, the scope of EST can be stretched by including additional motives.

Social Identity Theory

Social identity is commonly defined as a person's sense of self, derived from perceived membership in social groups. The simplest measure of social identity of an individual is group membership. When we belong to a group, we are likely to derive our sense of identity, at least in part, from that group. While standard economic analysis focuses on individual-level incentives in decision making, emphasis on group identity is now being given in mainstream economics too (Akerlof and Kranton 2000; Eckel and Grossman 2005; Basu 2006; Davis 2006; Chen and Li 2009).

The theory developed by Tajfel (1972) and Tajfel and Turner (1986) analyzes intergroup relations based on the distinction between ingroup and outgroup. Ingroups are the groups we identify with, and outgroups are the ones we do not identify with. As group identity affects individual behavior (Chen and Li 2009), many experiments assess whether and how people interact with ingroup and outgroup members differently. Most of these experiments confirm Tajfel's finding that group membership creates ingroup enhancement at the expense of the outgroup. With this is linked the concept of "common fate" or collective fate (Brewer and Kramer 1986; Sell 1997). Brewer and Kramer (1986) found that those, who are linked by a common fate involving a common determination of payoffs, do, indeed, cooperate more than those who are not so linked. Similarly, Orbell et al. (1988) argue that when the group identity is stronger more stakeholders tend to

depend on common fate and think group outcome as own outcome and create greater cooperation in social dilemmas.

Most literature suggests that gender is an important social identity, and this is more the case for women than for men precisely because women are traditionally of lower status compared to men (See especially, Lorenzi-Cioldi 1991; Pichevin and Hurtic 1996). Thus, social identity and status are interrelated. Social identity is less strong for high status actors, because they depend more on their individual attributes and capabilities than on collective fate (Sell 1997). People of low status believe that the status category determines outcomes, unlike the high-status people. When these persons interact with other lower status individuals, however, group identity would be activated and cooperation should be high.

Social identity prediction 1: Low-status actors (e.g., women, poor and low-caste people) will cooperate more in groups of all low-status actors than when in groups containing high status actors (e.g., Men, rich, and high-caste people).

The two theories are distinguished by predictions about the behavior of high-status individuals like men. Based on the idea that status is important for tying together people who share a common fate, low-status actors should be affected by the group's gender composition. High-status actors will not be affected by the status composition of the group because they depend less on common fate.

Social identity prediction 2: High-status actor's cooperation will not differ according to group's gender and/or status composition.[4]

Understanding Stakeholder Attitudes and Behaviors: A Field Study

In this section, we delineate how stakeholder attitudes and behaviors[5] can be measured with the help of the methods discussed above and examine the theoretical predictions on their behaviors in the context of community forestry in India. For this, we use the data of a field survey and a series of trust experiment conducted among 543 forest dependent households in two phases (2006–2007 and 2009–2010) in seven community forest areas in West Midnapore and Jalpaiguri districts of West Bengal, India. The attitudinal surveys were administered in 2006–2007 while the experiments were conducted in 2009–2010. This break is due

[4] It is not necessary that only these two interaction theories be invoked. Rather, any theory that appears to be relevant should be invoked in order to understand human behaviors. For an example, theory of heterogeneity conveys that *similar people are more likely to behave similarly*. In other words, all-men and all-women groups, for example, will be more pro-social than the mixed group, because women are homogeneous in their social preferences and so are men, while in the mixed groups, gender effect works.

[5] Although stakeholder refers to government, NGOs and the local people, in this study we consider only the local people as stakeholder. Hence, the term 'local people' and 'stakeholders' have been used interchangeably in this chapter.

to political turbulence in the study areas. Since no respondents in the household survey and in economic experiments participated more than once, our data are not panel data. The details of the data collection are given in Ray and Bhattacharya (2011). We use the data of a sub-sample of 196 participants, who participated in a trust game. The game setup is as in Berg et al. (1995). We describe this experimental protocol below.

Examining Attitudes–Behaviors Consistency: An Experiment

Are respondents' conservation behaviors consistent with their attitudes? If so, who are more pro-social: men or women? Since authors find that conservation friendly decision making, and hence behaviors are affected substantially by bargaining between men and women, relative strength of women does matter (Agarwal 2010). Hence, we pose here a query: Does women's proportional strength matter in revealed social preferences (like trust, reciprocity and cooperation) of stakeholders?[6] In other words, do respondents do what they state in the context of community forestry? If their revealed behaviors do not significantly correspond to their stated conservation-related attitudes, we can identify the free-riding tendency of stakeholders, although in a naïve manner.[7] To examine these issues and give completeness to this study, we have used the experimental setup of Berg et al. (1995) in several sessions. It is the most popular way of measuring trust via economic experiment.

In this game, senders are given a certain amount of money where they decide about how much of it to send to receivers and how much to keep. Any positive amount sent is tripled (to make sending socially efficient) before it reaches receivers, who then decide how much of the tripled amount of money received is to be sent back. Fractions sent and fractions returned are treated as measures of individual player's trust and reciprocity respectively.

We extended the above experiment as follows.[8] Our participants played in groups rather than as individuals. Each group comprised two/three members with varying proportions of females (0, 33, 50, 67 and 100 %, respectively) and played

[6] We extend the existing studies (Bratton and Ray 2002; Agarwal 2010) by examining the relationship between women's proportional strength and social preferences. Although gender does matter in social preferences, we also extend the existing literature by examining the threshold effect of gender on prosocial behaviors.

[7] One interesting aspect of our study is that it tries to capture (but not quantify) free-riding behaviors of resource users in a trust game when free-riding tendency is generally detected in public goods experiments.

[8] Instructions of the experiment are available upon request from the authors.

as both senders and trustees with prior knowledge about the roles.[9] Based on ex-ante matching, sender groups anonymously paired with the responder groups played the game once. The game was double blind. Neither the experimenter nor the players knew the identity of the groups they were paired with. In community forestry, some social groups like women and low caste groups are often discriminated. They are given less importance, denied from decision-making power etc. Hence, we let each group know about the gender composition of its paired group to examine whether the group sends/returns less to its counterpart when it knew that the group it had been paired with had more female members. The stake size was 50 INR (1USD = 48 INR) per member, half day's wage in the study site. To measure cooperation each group was given 15 min to make decisions on the amounts to be sent or returned. Players were told that the amount sent and the amount received would be shared equally.

Results

Stakeholder Attitudes Towards Co-management and Organization

Table 20.2 describes attitudes of the respondents toward joint forest management as a management paradigm as a whole and stakeholder attitudes toward members of the local forest management organization and forest department officers. Interestingly, around 70 % of the local people 'strongly agree' or 'agree' with the statements on co-management. A slight decline is noticed in their attitudes toward their fellow members (56–63 %). Thus, more than 50 % of the local people appear to hold positive attitudes toward community-based conservation since out of a maximum score of 15 in each scale they have scored 11.47 and 10.47, respectively.

Coefficient alpha displays the importance of each statement in the two scales. While the attitudes data of the local people toward the two scales are reliable (as alpha values exceed 0.71 in each case), all the statements do not have equal impact on the people as revealed in the alpha values corresponding to each item/statement. If a statement is deleted from the co-management scale, because of the formula of alpha its value will decline. For instance, the most important aspect of co-management scale that makes people's attitudes consistent is '*closeness of people and government'*. It is because without considering the statement people's attitude toward co-management becomes significantly less reliable (alpha decreases from 0.78 to 0.61). Similarly, the respondents treat '*mutual trust in local development'* between the local people and organization as the very important dimension of organization.

[9] The objective of 'playing both roles' is to double the experimental data within a given time. Many authors have done this before (see Bouma et al. 2008 in the context of watershed management).

Table 20.2 People's attitudes toward Community-based forest management in West Bengal and reliability analysis

Statements	SA	A	NR	D	SD	Mean ± SD	Cronbach's alpha if item deleted	Corrected item-total Correlation (r_{i-t})
Co-management Attitudes						11.47 ± 3.04	0.78	
1. Joint Management of local forest (JFM) both by the Forest Department and the people of this village is better than forest management either by village community alone or state alone, and also better than keeping it as an open access	36.73	32.65	10.71	10.71	9.18	3.77 ± 1.30	0.74	0.58
2. I like JFM Program as it integrates me with our village forests	35.20	37.76	13.27	9.18	4.59	3.89 ± 1.14	0.73	0.58
3. Only, a management system such as JFM can bring government and Forest Department closer to people	35.20	36.22	9.18	13.27	6.12	3.81 ± 1.22	0.61	0.69
Attitudes towards organizations						10.47 ± 3.85	0.84	
1. FPC and Forest Department officers are managing local forest well	33.16	30.61	5.10	23.47	7.65	3.62 ± 1.37	0.82	0.65
2. Efforts of local FPC members and the Forest Department to minimize human-elephant and other resource-related conflicts are praiseworthy	32.65	24.49	7.14	18.37	17.35	3.39 ± 1.53	0.78	0.70
3. Members of the local forestry organizations and Forest Department personnel trust each other very much, which help in local community economic development	33.16	28.06	4.08	16.33	18.37	3.45 ± 1.53	0.71	0.76

Liu et al. (2010) use similar type of interpretation of alpha in the context of protected area-community conflict in China.

It is important to note that in the case of attitudes toward organization, respondents' average score is lower and the alpha is greater than in the co-management scale. It may be because attitudes toward co-management are general worldview while attitudes toward organization are specific to the study areas. Hence, when some respondents (30–35 % as in Table 20.2) hold negative attitudes toward local JFMC, the reason might range from inequity in access to forest benefits and distribution of conservation responsibilities to asymmetric power relation (Vedeld 2000). However, the overall attitudes of the local people appear to be pro-environment.

Since different social groups may hold different perceptions and attitudes toward co-management, the attitudinal profiles of our respondents can be analyzed more rigorously if we consider the social attributes of the people. Two such important social characteristics are gender and social status of the respondents. We consider these attributes because they can vary from general to site-specific ones.

Regarding the gender environment relation, women are closer to nature and men to culture (Agarwal 1992).Yet, women are the vulnerable stakeholder in co-management: they depend most on natural resources, but do not have any decision-making power (Agarwal 2001). On the other hand, in selecting the community forestry leader and assigning important decision-making options, power relation does matter in community resource management (Vedeld 2000). High-status people may promote conservation to reap significant private benefits (Adhikari 2005), or they may prevent it from taking place when this goes against their interests (Shanmugaratnam et al. 1992).

In our study areas, women appear to be more positive both toward community-based resource management (4.11 vs. 3.57, $p < 0.01$) and organization (3.70 vs. 3.30, $p < 0.05$) than male. Coefficient alpha reveals that they are also more consistent than men are.

Many scholars (e.g., Eckel et al. 2010) have demonstrated the importance of status differences in decision making (See Heffetz and Frank 2011 for an excellent review of status studies in economics). In this article, we consider only the *social status* of a household. It is determined through a participatory technique. Three elements emerged as indicators of an individual's status. If a participant or any of her family members either: (1) holds mortgaged land, (2) works in government sectors like schools and/or (3) has useful socio-political connections, the participant is assigned to have high social status. Social status of resource users is diffused in the sense that it is not confined to our study setting only.

Table 20.3 surprisingly demonstrates that respondents with high social status are more co-management friendly and have favorable attitudes toward the local JFMC members and forest officers. However, their attitudes are not uniformly consistent. Alpha values of high-status people differ substantially between the two constructs. Because economic persons are often involved in status seeking behaviors (Heffetz and Frank 2011), involvement of high-status people in community forestry issues may raise or at least maintain their social status.

Table 20.3 Stakeholder Attitudes differentiated by gender and status

Attitudinal measures	Average attitude scores				Cronbach's alpha			
	Gender		Status		Gender		Status	
	Women	Men	Low	High	Women	Men	Low	High
Co-management attitudes	4.11	3.57	3.69	4.06	0.82	0.73	0.76	0.69
Attitudes towards organizations	3.70	3.30	3.36	3.86	0.91	0.77	0.82	0.88

Another reason might be that high-status people derive more benefits from conservation than the low-status people do. Hence, they support the local JFMC.

Pro-Social Behaviors of the Local People

The trust experiment shows (Table 20.4) that the most and the least trusting are the all-women and all-men groups, respectively. In the mixed group, men contribute

Table 20.4 Revealed pro-social behaviors of participants

Gender composition of the playing groups	Pro-social behaviors			
	Trust (Fraction sent)	Reciprocity (Fraction returned)	Cooperation (in min)	Index of pro-social behaviors
All-men group	0.44	0.35	12.05	0.42
Mixed gendered group (Overall)	0.47	0.48	10	0.50
Men in mixed group	0.49	0.49	–	0.19
Women in mixed group	0.45	0.46	–	0.19
Mixed group with male dominance	0.55	0.57	7.33	0.68
Mixed group with female dominance	0.40	0.46	9.56	0.49
All-women group	0.58	0.49	10.50	0.54
Status composition of the playing group				
All high-status group	0.48	0.37	12.20	0.42
Mixed status group (Overall)	0.46	0.40	10.84	0.47
High status	0.45	0.39	–	0.18
Low status	0.48	0.42	–	0.18
Mixed group with high-status dominance	0.28	0.45	10.20	0.43
Mixed group with low-status dominance	0.48	0.44	10.69	0.49
All low-status group	0.48	0.47	10.70	0.50

Note Overall, fraction sent = 0.48, fraction returned = 0.45, time taken for cooperation = 10.79 min and index of pro-social behavior = 0.51 per group

more than the contributions of the women of the mixed groups and the all-men group. Women in the mixed groups contribute less than the all-women groups. Furthermore, of all the mixed groups, mixed group with male dominance (i.e., mixed group with one-third women) trust their partners most.

In case of reciprocity or trustworthiness, all-men groups are least trustworthy, while mixed groups and all-women groups are, on average, equally trustworthy. Interestingly, 33 % women's groups are most trustworthy: they return 57 % of their received amount of money.

In this experiment, we measure cooperation as the time taken by a group to make its sending/returning decision. The more a group takes time to send/return money to its counterpart, the less the members of the group are cooperating with themselves. Thus, all-men groups appear as the least cooperative groups, taking 12 out of 15 min on average. Overall, investment is profitable since fraction returned exceeds one-third (0.45).

We measure the index of pro-social behaviors as the average of the index of fraction sent plus index of fraction returned plus (1-index of time taken for decision-making). Values of all the indexes used are calculated by the formula: {(actual-minimum)/(maximum-minimum)} and then averaged across households to get the average index values for a group or for a JFMC. Table 20.4 shows that the mixed groups with male dominance (i.e., 33 % female groups) perform best: They are the most trusting, trustworthy and cooperative groups. Hence, they are most pro-social (index of pro-social behaviors = 0.68).

We note that players' status affects social preferences. Low-status groups trust others as much as the high-status groups do. However, they reciprocate and cooperate more than the high-status groups as reflected in the difference in the values of index of pro-social behaviors of the two groups. Hence, composition of the playing group affects the level of social capital such as trust. Cardenas (2003) reinforces our claim. He finds in a CPR game that groups composed of mostly poor people cooperate more and conserve common property resource better than groups, which consist of poor people and more affluent local property owners. However, more research is needed to understand what matters more in social preferences-gender or status differences.

Thus, the game theory prediction that rational people always adopt Nash strategy (here 'No investment') does not hold in this study. In fact, Nash strategy has rarely been found to hold in field experiment (see Cardenas and Carpenter 2008). In contrast, in this study the predictions of the two interaction theories hold strongly.

Examining Attitude–Behavior Consistency of Participants

The Pearson's correlation coefficients in Table 20.5 reveal that attitudes of the respondents toward co-management and the local organizations are negatively associated with time taken for collective decision making and positively with fraction sent and fraction returned, and with the index of pro-social behavior as a whole. Thus, behaviors of the participants appear to be consistent with their

Table 20.5 Examining Attitude–behavior consistency between the playing groups

Attitudinal measures	Behavioral measures			
	Trust	Reciprocity	Cooperation (time taken for decision-making)	Index of pro-social behaviors
Co-management attitudes	0.51	0.44	−0.38	0.59
Attitudes towards Organizations	0.62	0.41	−0.47	0.69
Aggregate attitudes	0.64	0.47	−0.48	0.72

Note 1. Figures represents Pearson's correlation coefficients
2. Attitudinal measures are significantly correlated with behavioral measures at 1 % level. $N = 76$ playing groups

attitudes revealed in their responses to the attitude scales: those, who state that they are more co-management friendly or have positive attitudes toward the JFMC members and the Forest Department officers, send and return more money to their partner groups. They also take less time to arrive at a collective decision. However, relatively low values of all the correlation coefficients indicate the presence of free riding to some extent in every JFMC.

Forest Attribute, Transaction Costs, and Performance of Organizations

Does a more pro-social organization incur low cost and endure long? With the help of Table 20.6 and Fig. 20.1, we try to find a plausible answer to this question in the context of the surveyed JFMCs. At the outset, we deal with the measurement of some variables used in this study. First, sacrifice for conservation captures the voluntary environmental collective action for conservation. It is measured as annual labor days spent by a household in cooperating with other JFMC members for local forest management (1 Labor Day = 8 h in the study areas). The index of this variable has been constructed in line with the other indexes used in this chapter. Second, transaction costs in the co-management context refer to costs of enforcing common property right to forest and hence we measure it as the costs of monitoring, sanctioning, information collection and sharing, and conflict resolution (Ray and Bhattacharya 2011). Fourth, forest benefits consist mainly of non-timber forest products (NTFPs) valued at market price.[10] Fifth, we measure inequalities in transaction costs and forest benefits by Gini coefficient. Sixth, for

[10] There is also a timber benefit. Timbers are auctioned to the local related industries, and a 15 % (in BTR) to 25 % (WMP) share of sale proceeds are given to the concerned JFMC. Since this sum is generally used in local infrastructure development like road construction etc., these timber benefits are community benefits in nature unlike the individual household's forest benefits like fuelwood, fodder etc. Since timbers are not felled annually, and felling differs across JFMCs depending upon the forest health, it is difficult to add timber benefits to NTFPs to arrive at the

Table 20.6 Linking pro-social behaviors with stability of organizations

JFM organizations	Index of pro-social behaviors	Index of sacrifice for conservation	Stability of organization	Annual average variable transaction costs days	Inequality in distribution of transaction costs (Gini)	Inequality in distribution of forest benefits (Gini)	Trend in forest condition
Salboni	0.72	0.8	Stable	17	0.16	0.08	Improving
Mahuldanga	0.52	0.62	Stable	20	0.08	0.08	Stable
Chharadhan	0.64	0.66	Stable	21	0.12	0.10	Improving
Panialguri	0.62	0.32	Stable	23	0.13	0.25	Improving
Kalkut-Cheko	0.49	0.25	Quasi-stable	23	0.35	0.20	Improving
Banschati	0.29	0.31	Weak	30	0.28	0.30	Declining
Poro-Basti	0.40	0.2	Quasi-stable	35	0.36	0.23	Stable

Source Calculation based upon Authors' village Survey

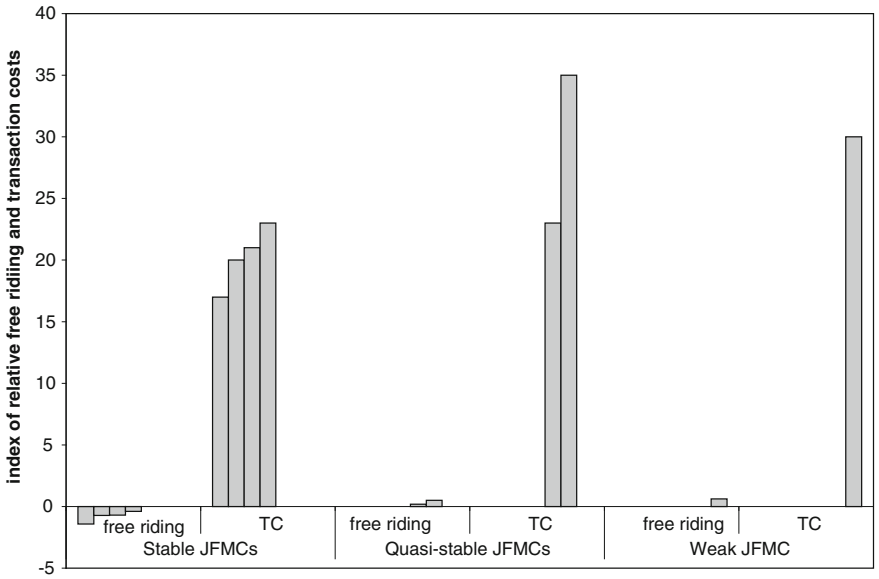


Fig. 20.1 Stability of JFMCs, transaction costs of co-management and the free riding behaviors of participants

the state of crown density, the relative abundance of forest products, disappearance of tree species and changes in forest area we rely on the narratives of the locals, the foresters' evaluations and authentication by our forestry experts. Like Varughese and Ostrom (2001), we define a trend in forest condition as improving/stable/declining if vegetation (tree species) and forest cover thus assessed were

(Footnote 10 continued)
aggregate forest benefits for a JFMC. Hence, we consider only the NTFP benefits in this chapter for each respondent and hence each JFMC.

expanding/unchanged/degrading. Our comparison periods include the time when the local JFMCs started managing the local forests, and 2007, at the time of our field survey. Finally, following Ostrom (1992) we define a JFMC as *stable* if it can make sustained claims over forest resources and take any offences to external court or administrative bodies or even lodges cases against the offender. A JFMC is *quasi stable* if it can make the offender surrender to the Forest officials, while *weak* organizations do not take such actions.

Table 20.6 exhibits that all the JFMCs with lower or moderate transaction costs have stable or improving forests. By contrast, Banschati is the only JFMC with declining forest. Therefore, it appears that forest trends in the study area are inversely associated with transaction costs. Chakrabarti et al. (2001) also note such phenomena in 12 IFRI (International Forestry Resources and Institutions) sites of northern West Bengal, India.

Similarly, stable organizations have more pro-social members and they voluntarily contribute for conservation. These organizations also bear low transaction costs. Moreover, we observe that in stable JFMCs every member bears almost equal transaction costs. Stable organizations have also stable or improving forests.

Following Ostrom (1992), we consider distribution of forest benefits as one of the indicators of performances of the JFMCs. Organizations with more equal distribution appear to have low transaction costs. However, even though, a relatively more unequal benefits distribution is observed in case of Panialguri JFMC, it has already gained reputation in the area for its mutual trust and high collective action (Ray and Bhattacharya 2011). Therefore, Panialguri JFMC may endure long in the region. Banschati, on the other hand, is a declining forest area where distribution of forest resource benefits is markedly unequal and transaction costs are higher than the other forest villages. During our survey, we found members of that organization often involved in illegal infractions and timber felling, leading to the unsustainable resource harvest. Deteriorating cooperation coupled with abysmal poverty has instilled in the local residents a sense of treating the local forest as a *defacto* open access resource. Ray and Bhattacharya (2011) find that costs of conflict resolution in the Banschati JFMC are substantial (30 %). High transaction costs might have discouraged local people to cooperate. As a result, it has become a weak organization.

In short, no definite relationship (Pearson's correlation coefficient = 0.49, $p > 0.10$) is found between transaction costs and average forest benefits in the study sites. However, equalities in the benefits and cost matter in reducing transaction costs (Poteete and Ostrom 2004, pp. 454). See Appendix 1 for the correlations between the variables.

Free Riding and Stability of JFMCs

In order to examine the effect of free riding of the JFMC members on organizational stability and transaction costs, we construct an index of relative free riding for each JFMC as follows:

$I_{\text{free riding}} = \left(\frac{I_{\text{ForBen}} - I_{\text{SacCon}}}{I_{\text{ForBen}}}\right)$, where I_{ForBen} and I_{SacCon} refer to average values of the index of forest benefits and sacrifice for collective action respectively for the participants of each JFMC. A very low or negative value of $I_{\text{free riding}}$ implies that the members of the JFMC may be treated as operating as a collective. Figure 20.1 shows that stable organizations are collectives while members of the quasi stable and weak JFMCs tend to free ride. Furthermore, JFMCs with greater tendency to free ride like Poro-Basti and Banschati have incurred higher transaction costs. One implication of this observation together with the findings in Table 20.5 may be that the local organizations are not optimal in the sense of complete absence of free-rider problem and that the local forest department must be more instrumental in motivating other non-member users of JFMC villages to become members.

Conclusions

CBNRM has become a leading paradigm for conservation with development. It is equally important for sustainable rural livelihoods. Nonetheless, the experiences of different countries in CBNRM are mixed. Reasons vary across culture and context (Ostrom 1990). The most important ones are the organization-people conflicts, lack of social capital, and hence negative perceptions and attitudes of the resource user communities. The obvious outcomes are asymmetric power relations among the local people and consequently, social inequalities in the form of vulnerabilities of women (Agarwal 2001) and the poor as resource users (Adhikari 2005). This marginalization of weaker sections in CBNRM might affect environmentalities of the locals (Agrawal 2006). Positive environmentalities reduces transaction costs of working together for enforcing common property rights (Ray and Bhattacharya 2011) and makes people pro-environment. Here lies the importance of understanding attitudes and behaviors of the resource using communities for achieving conservation with development. When it comes to the question of survival of resource management organizations, proper understanding of stakeholder attitudes is important, because environmental co-management policy is also a social policy. Since we make decisions as social beings, we need an integrated interdisciplinary approach to understanding the people.

With this aim in mind, we reviewed the methods of measuring human attitudes and behaviors and cited some important cautions while using these methods. We find that game theory makes incorrect behavioral predictions. Since individuals/groups make behavioral decisions embedded in social context, the implication is that without considering social factors like social identity and social status, economic theories cannot properly understand human behaviors. We invoke two interaction theories from social psychology to make refinements in our theoretical predictions on agents' behaviors and our review theoretically robust.

Finally, we cite a case study of forest co-management from West Bengal, India to show how these methods can be used in practice. Here we are concerned more

with methodologies than with analyses. Since empowerment of the locals in community decision making is often identity based (women or men) and social standing (high or low social status), we have considered gender and social status as the two key stakeholder attributes to understand the attitude-behavior (in)consistency of the surveyed people. Interestingly, we see that people, who hold positive attitudes toward co-management and the local forest management organization, are consistent in expressing their mentalities. While women are more pro-social in all-women groups and men in mixed groups, we observe that status of the participants also affects social preferences. On the other hand, forestry organizations, which bear low transaction costs, are stable. People of these organizations also free ride less. In this sense, transaction costs can also indirectly measure attitudes of stakeholders. In other words, organizations with high transaction costs decline, because members hold negative attitudes and are less pro-social.

In short, institutions cannot sustain without the positive mentalities of stakeholders. To unravel this we need to invoke theories and complement them with the methods that explain human behavioral differences in ways that indicate some practical policy implications.

Acknowledgments We acknowledge our debt to the Departmental Research Support (DRS)-I of the Department of Economics, University of Calcutta for funding the field experiments and surveys. We thank Sarmila Banerjee for her continued encouragement. We also thank Puspendu Maity, Jayanta Chakraborty and their team members, foresters, and villagers of the study sites for their kind cooperation during the surveys. Usual disclaimer applies.

Appendix

Appendix 1: Pearson Correlation Matrix on the Attributes of JFMCs

	Index of prosocial behaviors	Stability of organization	Transaction cost (TC)	Inequality in TC	Inequality in Benefits	Forest trends	Free riding
Index of prosocial behaviors	1						
Stability of organization	0.90^{***}	1					
Transaction cost (TC)	-0.82^{**}	-0.73^{**}	1				
Inequality in TC	-0.65^{*}	-0.77^{**}	0.73^{**}	1			
Inequality in benefits	-0.71^{**}	-0.76^{**}	0.75^{**}	0.61^{*}	1		
Forest trends	0.86^{***}	-0.73^{**}	-0.63^{*}	-0.28	-0.47	1	
Free riding	-0.91^{***}	-0.86^{***}	0.89^{***}	0.78^{**}	0.85^{***}	-0.60^{*}	1

***, **, and * indicate that the correlations are significant at the 0.01, 0.05 and 0.10 levels respectively (one-tailed)

References

Adhikari B (2005) Property rights and collective action: understanding the distributive aspects of common property resource management. Environ Dev Econ 10(1):7–31
Agarwal B (1992) The gender and environment debate: lessons from India. Fem Stud 18:119–158
Agarwal B (2001) Participatory exclusions, community forestry and gender: An analysis and conceptual framework. World Dev 29(10):1623–1648
Agarwal B (2009) Rule making in community forestry institutions: the difference women make. Ecol Econ 68:2296–2308
Agarwal B (2010) Does women's proportional strength affect their participation? governing local forests in South Asia. World Dev 38(1):98–112
Agrawal A (2006) Environmentality: technologies of government and the making of subjects. Oxford University Press, New Delhi
Ajzen I, Fishbein M (1977) Attitude-behavior relations: a theoretical analysis and review of empirical research. Psychol Bull 8:8–918
Akerlof GA, Kranton RE (2000) Economics and identity. Quart J Econ 15(3):715–753
Ashraf N, Bohnet I, Piankov N (2006) Decomposing trust. Experimental Economics 9(3):193–208
Basu K (2006) Identity, trust and altruism: sociological clues to economic development. Cornell University Center for Analytic Economics Working Paper 06–05. http://ssrn.com/abstract=956080
Bellemare C, Kröger S (2003) On representative trust. Tilburg University CentER Discussion Paper 2003-47
Berg J, Dickaut J, McCabe K (1995) Trust, reciprocity and social history. Games Econ Behav 10(1):122–142
Berger J, Fisek MH (1974) A generalization of the theory of status characteristics and expectation states. In: Berger J, Conner TL, Fisek HH (eds) Expectation states theory: a theoretical research program. Winthrop, Cambridge, pp 163–205
Bolton GE, Katok E (1995) An experimental test for gender differences in beneficent behavior. Econ Lett 48:287–292
Bouma J, Bulte E, van Soest D (2008) Trust and cooperation: social capital and community resource management. J Environ Econ Manag 56(2):155–166
Bratton KA, Ray LP (2002) Descriptive representation, policy outcomes, and municipal day-care coverage in Norway. Am J Polit Sci 46(2):428–437
Brewer M, Kramer RM (1986) Choice behavior in Social dilemmas: effects of social identity, group size and decision framing. J Pers Soc Psychol 50:543–549
Buchan NR, Croson R, Solnick SJ (2008) Trust and gender: an examination of behavior and beliefs in the investment game. J Econ Behav Organ 68(3–4):466–476
Campbell JP, Dunnette MP, Arvey RP, Hellervik LV (1973) The development and evaluation of behaviorally based rating scales. J Appl Psychol 57:15–22
Cardenas JC (2003) Real wealth and experimental cooperation: evidence from field experiments. J Dev Econ 70(2):263–289
Cardenas JC, Carpenter J (2008) Behavioural development economics: Lessons from field labs in the developing world. J Dev Stud 44(3):311–338
Cardenas JC, Stranlund J, Willis C (2002) Economic inequality and burden-sharing in the provision of local environmental quality. Ecol Econ 40(3):1–17
Carpenter J, Daniere A, Takahashi L (2004a) Social capital and trust in Southeast Asian cities. Urban Stud 41(4):853–874
Carpenter J, Daniere A, Takahashi L (2004b) Cooperation, trust, and social capital in Southeast Asian urban slums. J Econ Behav Organ 55(4):533–551
Carter M, Castillo M (2003) An experimental approach to social capital in South Africa. Working paper, Department of Agricultural Economics, University of Wisconsin

Castillo M, Carter M (2003) Identifying social effects in economic field experiments. Working paper, Department of Agricultural and Applied Economics, University of Wisconsin
Chakrabarti M, Khaling S, Bhattacharya J, Ghosh SR, Sarkar A, (2001) Functioning of joint forest management in the forests of North Bengal: observations from 12 IFRI sites. Paper presented at a dissemination Seminar, December 2001. St.Joseph's College, Darjeeling, India
Chakravarty S, Friedman D, Gupta G, Hatekar N, Mitra S, Sundar S (2011) Experimental economics: a survey. Econ Polit Wkly 56(35):39–78
Chen Y, Li SX (2009) Group identity and social preferences. Am Econ Rev 99(1):431–457
Correll SJ, Ridgeway CL (2003) Expectation states theory. In: Delamater J (ed) Handbook of social psychology. Kluwer Academic/Plenum Publishers, New York, pp 29–51
Cox JC (2004) How to identify trust and reciprocity. Games Econ Behav 46(2):260–281
Croson R, Gneezy U (2009) Gender differences in preferences. J Econ Lit 47(2):448–472
Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. Psychometrika 16:297–334
Danielson A, Holm H (2007) Do you trust your brethren? Eliciting trust attitudes and trust behavior in a Tanzanian congregation. J Econ Behav Organ 62(2):255–271
Davis JB (2006) Social identity strategies in recent economics. J Econ Methodol 13(3):371–390
Dawis RV (1986) Scale construction. J Couns Psychol 34(4):481–489
Dunlap RE, Jones RE (2002) Environmental concern: conceptual and measurement issues. In: Dunlap RE, Michelson W (eds) Handbook of environmental sociology. Greenwood Press, Westport, pp 482–524
Dunlap RE, Van Liere KD (1978) The new environmental paradigm. J Environ Educt 9:10–19
Dunlap RE, Van Liere KD, Mertig A, Jones RE (2000) Measuring endorsement of the new ecological paradigm: a revised NEP scale. J Soc Issues 56:425–442
Durlauf SN (2002) On the empirics of social capital. Econ J 112(483):459–479
Eckel CC, Grossman PJ (1998) Are women less selfish than men: evidence from dictator experiments. Econ J 108:726–735
Eckel CC, Grossman PJ (2001) Chivalry and solidarity in ultimatum games. Econ Inq 39(2):171–188
Eckel CC, Grossman PJ (2005) Managing diversity by creating team identity. J Econ Behav Organ 58:371–392
Eckel CC, Fatas E, Wilson R (2010) Cooperation and status in organizations. J Public Econ Theory 12(4):737–762
Fischbacher Urs (2007) z-Tree: Zurich toolbox for ready-made economic experiments. Exp Econ 10(2):171–178
Greig F, Bohnet I (2005) Is there reciprocity in a reciprocal-exchange economy? Evidence from a slum in Nairobi, Kenya. Working paper, Kennedy School of Government
Heffetz O, Frank RH (2011) Preferences for status: evidence and economic implications. In: Benhabib J, Bisin A, Jackson M (eds) Handbook of social economics, vol 1A. North-Holland, The Netherlands, pp 69–91
Henrich J (2000) Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga Indians of the Peruvian Amazon. Am Econ Rev 90:973–979
Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H, McElreath R (2001) In search of homo economics: behavioral experiments in 15 small-scale societies. Am Econ Rev 91(2):73–78
Henrich J, McElreath R, Barr A, Ensminger J, Barrett C, Bolyanatz A, Cardenas JC, Gurven M, Gwako E, Henrich N, Lesorogol C, Marlowe F, Tracer D, Ziker J (2006) Costly punishment across human societies. Science 312:1767–1770
Johansson-Stenman O, Mahmud M, Martinsson P (2004) Trust, trust games and stated trust: evidence from rural Bangladesh. Working paper, Department of Economics, Goteborg University
Infield M (1988) Attitudes of a rural community towards conservation and a local conservation area in Natal S Africa. Biol Conserv 45:21–46

Karlan D (2005) Using experimental economics to measure social capital and predict financial decisions. Am Econ Rev 95(5):1688–1699
Kreps DM (1990) A course in microeconomic theory. Princeton University Press, Princeton
Likert R (1932) A technique for the measurement of attitudes. Arch Psychol 140:5–55
Liu J, Ouyang Z, Miao H (2010) Environmental attitudes of stakeholders and their perceptions regarding protected area-community conflicts: a case study in China. J Environ Manag 91:2254–2262
Lorenzi-Cioldi F (1991) Self-stereotyping and self-enhancement in gender groups. Eur J Soc Psychol 21:403–417
Maloney MP, Ward MP (1973) Ecology: let's hear it from the people. an objective scale for measurement of ecological attitudes and knowledge. Am Psychol 28:583–586
Meeker BF, Weitzel-O'Neill PA (1977) Sex roles and interpersonal behavior in task-oriented groups. Am Sociol Rev 42:91–105
Mehta JN, Heinen JT (2001) Does community-based conservation shape favourable attitudes among locals? An empirical study from Nepal. Environ Manag 28:165–177
Milfont TL, Duckitt J (2010) The environmental attitudes inventory: a valid and reliable measure to assess the structure of environmental attitudes. J Environ Psychol 30:80–94
Mukherjee P (2011) Attitudes, institutions and cooperation: does gender matter in joint forest management in West Bengal, India? M.Phil Dissertation, Rabindra Bharati University, Kolkata, India
Netemeyer RG, Bearden WO, Sharma S (2003) Scaling procedures: issues and application. Sage Publications, London
Orbell JM, Van de Kragt AJC, Dawes RM (1988) Explaining discussion-induced cooperation. J Pers Soc Psychol 54:811–819
Ostrom E (1990) Governing the commons. Cambridge University Press, Cambridge
Ostrom E (1992) The rudiments of a theory of the origins, survival, and performance of common-property institutions. In: Bromley DW, Feeny D, McKean MA, Peters P, Gilles JL, Oakerson RJ, Runge CF, Thomson JT (eds) Making the commons work: theory, practice, and policy. ICS Press, San Francisco, pp 293–318
Pender JL (1996) Discount rates and credit markets: theory and evidence from rural India. J Dev Econ 50(2):257–296
Pichevin M, Hurtic M (1996) Describing men, describing women: sex membership salience and numerical distinctiveness. Eur J Soc Psychol 26:513–522
Poteete AR, Ostrom E (2004) Heterogeneity, group size and collective action: the role of institutions in forest management. Dev Change 35(3):168–181
Quisumbing AR, Haddad L, Christine P (2001) Are women overrepresented among the poor? An analysis of poverty in 10 developing countries. J Dev Econ 66(1):225–269
Ray B, Bhattacharya RN (2011) Transaction costs, collective action and survival of heterogeneous co-management institutions: case study of forest management organizations in West Bengal, India. J Dev Stud 47(2):253–273
Robinson JP, Shaver RP, Writesman LS (1991) Criteria for scale selection and evaluation. In: Robinson JP, Shaver PR, Writesman LS (eds) Measures of personality and social psychological attitudes. Academic Press, San Diego, pp 1–15
Schmitt N (1996) Uses and abuses of coefficient alpha. Psychol Assess 8(4):350–353
Seguino S, Stevens T, Lutz MA (1996) Gender and cooperative behavior: economic man rides alone. Fem Econ 2(1):1–21
Sell J, Wilson RK (1991) Levels of information and public goods. Soc F 70:107–124.
Sell J, Griffith WI, Wilson RK (1993) Are women more cooperative than men in social dilemmas? Soc Psychol Q 56:211–222
Sell J (1997) Gender, strategies, and contributions to public goods. Soc Psychol Q 60(3):252–265
Shanmugaratnam N, Vedeld T, Mossige A, Bovin M (1992) Resource management and pastoral institution building in the West African Sahel. World Bank Discussion Papers, No. 175, The World Bank, Washington DC

Shogren JF, Taylor LO (2008) On behavioral environmental economics. Rev Environ Econ Policy 2(1):26–44
Sinha H, Suar D (2005) Leadership and people's participation in community forestry. Int J Rural Manag 1(1):125–143
Solnick SJ (2001) Gender differences in the ultimatum game. Econ Inq 39(2):189–200
Tajfel H (1972) Social categorization. In: Moscovici S (ed), Introduction a' la psychologie sociale. 1, Paris, Larousse, pp 272–302
Tajfel H, Turner JC (1986) The social identity theory of intergroup behavior. In: Austin W, Worchel S (eds) The social psychology of intergroup relations monterey. Brooks, Cole, pp 7–24
Varughese G, Ostrom E (2001) The contested role of heterogeneity in collective action: some evidence from community forestry in Nepal. World Dev 29(5):747–765
Vedeld T (2000) Village politics: heterogeneity, leadership and collective action. J Dev Stud 36(5):105–134
Weigel R, Weigel J (1978) Environmental concern: the development of a measure. Environ Behav 10:3–15
Wood W, Karten SJ (1986) Sex differences in interaction style as a product of perceived sex differences in competence. J Pers Soc Psychol 50:341–347

Chapter 21
Indoor Air Pollution and Incidence of Morbidity: A Study on Urban West Bengal

Sabitri Dutta and Sarmila Banerjee

Introduction

Air pollution causes more than 3 million premature deaths worldwide every year (WHO factsheet, 2011). Studies show that there is a strong association between poor quality of air and the human mortality and morbidity. Among five environmental risks explaining DALY,[1] indoor smoke is the second most important factor and it affects mostly the low and middle income countries. The indoor pollution is more aggressive because the same amount of air pollutant released indoors, produces roughly thousand times more exposure-impact than it would do outdoors (Smith, 2008). Sub Saharan Africa and South East Asian countries, especially Indian sub-continent, are the major bearer of the risk from indoor pollution from solid fuels. The concentration of indoor pollutants, particularly SPM and RPM, is alarming in India. It sometimes ranges between 2,000 $\mu g/m^3$ and 18,000 $\mu g/m^3$ against the safe level of 100 $\mu g/m^3$ recommended by NAAQM standard (Smith 1996).

This chapter mainly deals with the challenges faced in assessing morbidity impact of indoor air pollution in urban areas of developing world which has so far been a relatively neglected terrain in the research agenda. In urban context, the problem of vehicular pollution enjoys maximum attention and the indoor pollution is generally considered as a special problem for the rural households where

[1] DALY stands for Disability Adjusted Life Years and is accepted as a measure of morbidity; this measure combines deaths and illness, allowing for the level of disability (impaired health) resulting from the illness and the number of years of life affected by this disability (if the person survives) or lost completely (if the person dies).

S. Dutta (✉)
DumDum Motijheel Rabindra Mahavidyalaya, 208/B/2 DumDum Road, Kolkata 700074, India
e-mail: sabitri.d@gmail.com

S. Banerjee
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: sarmila_b@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_21, © Springer India 2013

biofuels are the major source of energy. In general, urbanization leads to higher levels of income and places the households in the upper tier of energy ladder where dirty fuels are progressively replaced by cleaner fuels.[2] Several factors that contribute to this trend include limited access to the biomass fuel, inconvenience in the transport and storage, and improvement in the availability of the commercial fuels in urban areas. In spite of this fact, use of these traditional fuels in many cities of the developing world remains high among the low income groups. Out of total energy use by the urban laborers in India, 57 % come from the traditional sources (Reddy 2004). More interestingly, the use of the biomass energy is not insignificant among the urban middle income employees also and even among the executive class. In fact, the annual per capita energy consumption of low income households in urban areas does not differ significantly from that of rural poor. Moreover, the congestion factors are more active in urban environment leading to a magnification of exposure effect.

The extent of indoor pollution is related to the extent of ventilation, which in turn depends on the structure and type of the house, cooking and heating mode, the type of fuel used and many other factors. It implies IAP is mostly dependent on the socio-economic status of the household along with the availability of cleaner fuel, improved stove, etc. This is different from the idea that pollution is a negative externality and exposure to the pollution is beyond the control of the individuals and it has to be sorted out only by the public authorities in local or global levels; rather in case of IAP it is a consequence of household activities and depends on household decisions regarding the covariates and these decisions are ultimately influenced by the socio-economic characteristics of the decision maker like income, education, etc. Households internalize the adverse effects of pollution while making their fuel choice. Apart from fuel type, the exposure depends largely on the living conditions, social location, demographic characteristics, and the level of awareness (Dasgupta et al. 2004, 2006; Ou et al. 2008, various reports of WHO), etc. However, the most important determinant of the level of exposure is the time spent in different microenvironments, where these socio-economic factors play crucial role in creating different microenvironments.

A few studies on the relationship between morbidity and exposure to air pollution are available on the developed countries where a negative correlation is observed between the level of exposure and the level of income, education, and standard of living. However, most of the studies till early 1990s used ambient concentration as an indication of pollution. The effects of outdoor pollution is suggestive of that of indoor pollution but limited in two senses, first, exposure to outdoor air pollution (OAP) can be reduced by changing activities or applying protective devices (viz. mask); second, a brief episode of exposure to OAP can be different from daily exposure to IAP (Duflo et al. 2008). In a significant departure, Brooks and Sethi (1997) computed an index based on toxicity of 150 most

[2] Energy Ladder Hypothesis.

significant air toxics rather than ambient concentrations for the entire US. The study then looked into whether exposure varied across various socio-economic characteristics including income, race, educational attainment, housing tenure, and the propensity of communities to engage in collective action. The study found greater exposure in black communities even after controlling for a variety of variables. Extent of inequality in exposure is the concern of environmental equity.

Recent efforts have also been directed to see the impact of exposure to air pollution in developing and middle-income countries in terms of its effect on human health. In a study of six regions of Sao Paulo, Brazil, Martins et al. (2004) found a negative effect of PM_{10} on both percentage of people with college education and high family income and a positive association with the percentage of people living in slums. Thus, the study concluded that socio-economic deprivation is an effect modifier of the association between the PM_{10} and respiratory deaths. In another study on air pollution and income, Wheeler (2000) found that low-income families were more likely to live in polluted areas. In a study on the effect of air pollution on the child hospitalization for asthma in USA Neidall (2004) showed that the effect of pollution is greater for children of lower socio-economic status. WHO (2010) report on environment and health risk indicated that the social inequalities are a major challenge for environmental health policies and the review confirmed that the people living in adverse socio-economic conditions suffer twice as much from environmental exposures than their wealthier counterpart.

In India, Suksena and Dayal (1997) attempted to assess the total exposure to the air pollution and calculated a dose response function showing relationship between morbidity and mortality due to PM_{10} exposure. In a different study Suksena et al. (2003) estimated pollution exposure for infants of the low income urban localities in Delhi. There are few other studies also available which attempt to estimate the exposure to the indoor pollution and establish a relationship between the morbidity and this exposure. Murty et al. (2003) carried out some primary survey on large groups of households in Delhi and Kolkata, collected information on their socio-economic status as well as health status and assessed pollution exposure in terms of ambient concentration of RPM to find out their willingness to pay for pollution reduction. The most extensive study on exposure index has been carried out by Kathuria and Khan (2007) on Delhi where in the absence of specific time-use data they have modified the ambient concentration level by member specific vulnerability weights (which are dependent on the age-sex composition of the household) to assess the average exposure level for each household. But still a gap remains in this area of research especially for the urban households of India.

This chapter aims to study the nature of relationship between the level of IAP and incidence of morbidity in Urban West Bengal. Initially, the type of fuel used has been taken as a proxy for IAP and then a methodology has been proposed to assess exposure to IAP at the individual as well as household level to study the relationship between morbidity and this exposure. For the first exercise unit level NSSO 60th Round data (2004–2005) on urban West Bengal has been utilized and

to construct the exposure index a primary survey has been carried out on Kolkata[3] during 2008–2009 to gather information on time spent by each individual member of a household in different microenvironment. The rest of the chapter has been organized as follows: Section 2 ("Fuel Use and Morbidity: Data Exploration") and 3("Fuel Use and Morbidity: Data Analysis") try to analyze relationship between the prevalence of morbidity and the type of fuel use by controlling for a number of socio-economic characteristics of the household for urban West Bengal. Section 4("Primary Survey on Kolkata: Design and Frame") describes the design and frame of the primary survey, section 5("Construction of Exposure Index") proposes a methodology to construct a microenvironment-specific exposure index, section 6("Incidence Analysis") carries out the incidence analysis and finally, section 7("Concluding Remarks") concludes the chapter by presenting a sensitivity analysis to indicate the nature and direction of effective policy interventions.

Fuel Use and Morbidity: Data Exploration

By using NSSO 60th Round Household Survey on Morbidity, Health Care and the Condition of the Aged a sample of 1877 observations has been extracted on urban West Bengal. Information was available on the incidence of morbidity (member-wise) experienced over the last 15 days and hospitalization episode over last 1 year period. This section discusses variable construction and data exploration. Data analysis would be reported in the following section.

Variable Construction

Two morbidity variables, M_1 & M_2 have been defined to represent health status of each household as follows:

M_1: Incidence of Morbidity; if there is any morbidity episode reported for at least one member of the household over the reference period then M_1 has been assigned a value 1, otherwise 0. Thus, M_1 is a dummy variable.

M_2: Index of Morbidity; for all households M_2 has been defined as the percentage of morbid members. So, the value of the index is non-negative by definition.

Here, no distinction has been made between once-morbid member vis-à-vis members suffering from multiple episodes of morbidity and no account has been taken of the nature and type of disease. Attempt has been made to explain M_1 & M_2 in

[3] The survey was financially supported by the UGC as a UPE project on Spatial Concentration of Pollution and Morbidity Impact in the City of Kolkata: Efficient Intervention Strategies.

terms of a number of socio-economic variables including the pattern of energy use. The independent variables are:

MPCE: Monthly per capita expenditure of the household; it is expected to have a negative effect on the level of morbidity.

LCI: Living condition Index; on the condition of living information was available on different aspects like (i) house structure (HST), (ii) latrine type (LTP), and (iii) drainage system (DRN). Since these aspects are likely to be correlated, hence, an attempt has been made to reduce dimensions and produce a comprehensive index LCI, by carrying out a Principal Component Analysis (PCA). Kaiser–Meyer–Olkin (KMO) measure equaled 0.659, ensuring sampling adequacy to carry out PCA and the Bartlett's test of sphericity produced a Chi square value 1203.00, which turned out to be significant at less than 1 % level. One Principal Component was extracted with associated eigenvalue equaling 1.939, which taken as a whole explained 65 % of the total variation (see Table 21.1 and Fig. 21.1). LCI is also expected to have a negative effect on the level of morbidity.

The factor scores obtained (using the software SPSS) for this component are used to define the Living Condition Index (LCI) for the households of West Bengal.

LCI = [(Actual factor score-minimum factor score)/(Maximum factor score-minimum factor score)]*100.

FUEL: A dummy variable has been defined with clean fuel (LPG and Electricity) represented by 0 and dirty fuel (Firewood, Biomass, Coal and Kerosene) as 1; though in literature kerosene is considered as a relatively clean fuel, the study of Lahiri and Lahiri (2001) on indoor air pollution of Kolkata has shown that due to low quality of publicly available Kerosene oil and congested living in the city, the blue smoke emitted from the kerosene stove at the time of ignition is highly polluting in nature and has significant deleterious effect on human respiratory system. The effect of FUEL, so defined is expected to have a positive effect on the level of morbidity. In fact, the level of morbidity should depend on the extent of exposure to indoor pollution. To estimate that one needs information on not only the type of fuel used but on the kitchen type, hours spent at home, occupation, and the quality of ambient air. In the absence of all such specific and detailed information we have taken FUEL as the proxy of exposure to indoor air pollution.

EDU: It stands for the level of education of the household. To simplify the matter we have taken level of education of the head of the household and defined an educational achievement index as: $EDU = (\text{Actual}/\text{Maximum}) \times 100$. It is expected to have a negative effect on the level of morbidity.

Table 21.1 Extracted components with factor loadings

Variables		HST	LTP	DRN	Percentage of variations explained
Component	$\lambda 1 = 1.939$	0.400	0.441	0.402	64.632

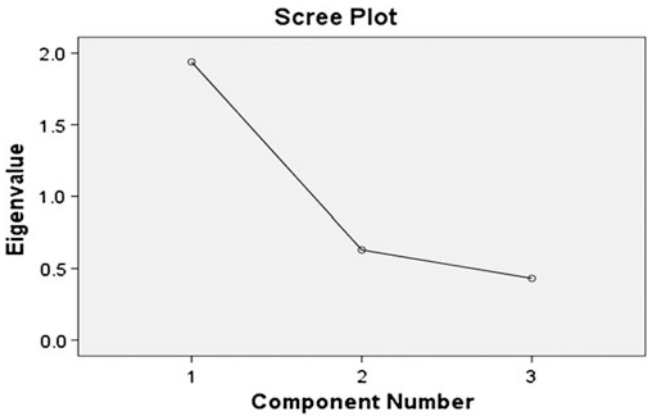


Fig. 21.1 (KMO = 0.659, Chi square = 1.203E3)

Data Exploration

In the next step, the sample distribution of all these variables are scrutinized in terms of their descriptive statistics (Table 21.2).

Since for linear regression to yield good estimates we need all variables to conform normal distribution hence, depending on the descriptive measures we check whether the data are reasonably symmetric and, if so, whether the tails are sufficiently thin or not. If the mean is larger than median then the distributions are positively skewed, the opposite would be the case for median exceeding the mean value. The analysis of the Inter Quartile Range (IQR) helps to decide whether the tails of the distribution are thinner or fatter than a regular normal distribution. On the basis of IQR we estimate a pseudo standard deviation.[4] If this is smaller than the standard deviation then the tails are heavier than normal. Thus, an idea about the required transformation of the data can be formed by comparing mean-based vis-à-vis order-based statistics of the data when one is trying to model an average behavior (Mukherjee et al. 1998). In the data set we have two attributes represented by two dummy variables: M_1 and FUEL along with four quantitative variables: M_2, MPCE, LCI, and EDU. Out of these latter four variables except MPCE the three others are Index-variables with the range of variation confined within [0, 100]. MPCE is the only free variable with the mean value much higher than the median value indicating the presence of positive skewness. Here the value of IQR is very high (716.00) and pseudo standard deviation is smaller than standard deviation indicating fat right tail.

To reduce the fatness of the tail and get back the mean-centricity of the distribution the higher values of the observations need to be moderated by applying

[4] A pseudo standard deviation is (IQR/1.35) which should be equal to the actual standard deviation for a Normal density.

Table 21.2 Descriptive statistics (NSS 60th round)

Descriptive statistics	M_1	M_2	*MPCE*	*LCI*	*FUEL*	*EDU*
Mean	0.49	17.99	1062.45	79.28	0.51	58.02
Median	0.00	0.00	875.00	88.27	1.00	66.67
Standard deviation	0.50	25.58	753.27	18.79	0.50	35.24
IQR	1.00	25.00	716.00	18.12	1.00	66.67
Pseudo standard deviation	0.74	18.53	531.3	13.43	0.74	49.42
Skewness	0.04	1.77	2.78	−1.37	−0.04	−0.39
Kurtosis (normalized)	–2.00	2.80	13.42	1.33	−2.00	−1.10
Sample size	1877	1877	1877	1877	1877	1877
Observations with 0 value	956				921	
Observations with 1 value	921				956	

some nonlinear transformation technique. Generally, for positively skewed data set with fat right tail (i.e., outliers at the higher end) logarithmic transformation works well. For negatively skewed distribution with fat left tail (i.e., outliers at the lower end) power transformation X^p with p equals 2 works better. So, being guided by the χ^2 value of goodness of fit we have replaced *MPCE* by *lnMPCE* (Fig. 21.2) and *LCI* by LCI^2 *(sqLCI)*. For M_2 there is a problem of positive skewness with actual standard deviation exceeding the pseudo standard deviation. However, here the possibility of logarithmic transformation has been ruled out as quite a large number of households have M_2 value equals zero, indicating the absence of any

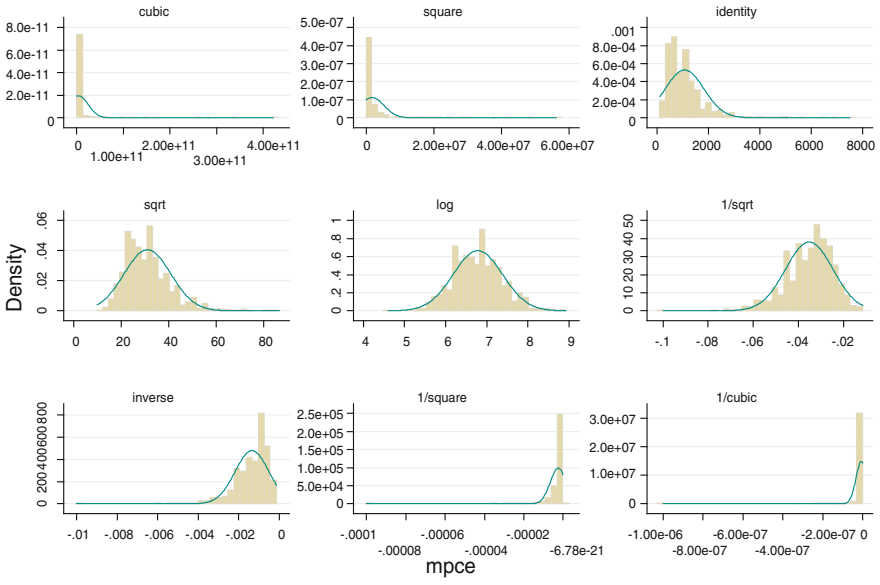


Fig. 21.2 The ladder-graph for MPCE

morbidity episode. So, M_2 has been replaced by $\sqrt{M_2}$ ($sqrtM_2$). For the remaining variables there is no indication of thick tail as kurtosis is negative and pseudo standard deviation is greater than standard deviation. Moreover, the ladder analysis failed to suggest any significant improvement of goodness of fit with any other functional form. Therefore, they have been left in their original (raw) forms.

Fuel Use and Morbidity: Data Analysis

Analysis of Correlation

The pair-wise correlation between all selected variables has been analyzed at the next stage (see Table 21.3).

Two alternative indicators of morbidity ($\boldsymbol{M_1}$ **&** $\boldsymbol{sqrtM_2}$) are highly correlated and the association of the latter one with almost all the explanatory variables is statistically stronger. Contrary to our expectation the variable ***FUEL*** is negatively related to both the morbidity indicators; however, the sign of association with ***lnMPCE*** and ***sqLCI*** conforms our expectation that with the increase in income and living condition households will switch over from dirty to clean fuel. The statistically significant positive linear association between ***EDU*** and $\boldsymbol{sqrtM_2}$ is little counter-intuitive. Here, two plausible explanations may be extended: (a) with the increase in the level of awareness via higher achievement of education perceived illness of the household may go up and (b) the relationship between the reduction in morbidity and higher educational attainment may be effective after some threshold level of achievement.

Since the pair-wise correlation between all explanatory variables is statistically significant the possibility of multi-collinearity needs to be verified. Before proceeding any further we have run an OLS regression of $\boldsymbol{sqrtM_2}$ on all the four regressors and checked the VIF (Variance Inflationary Factor). Since the VIF factors are lying between 1.92 and 1.10 with a mean value of 1.62, i.e., all the VIF values are way below 10, hence, the threat of multi-collinearity is not much compelling and standard regression exercise can be carried out on the data set.

Table 21.3 Pair-wise correlations*

Variables	M_1	$sqrtM_2$	*lnMPCE*	*sqLCI*	*FUEL*	*EDU*
M_1	1.000					
$sqrtM_2$	0.908*	1.000				
lnMPCE	0.076*	0.159*	1.000			
sqLCI	0.089*	0.094*	0.523*	1.000		
FUEL	−0.056*	−0.064*	−0.587*	−0.485*	1.000	
EDU	0.023	0.050*	0.578*	0.449*	−0.564*	1.000

*Significant at 5 % level

Table 21.4 Regression results_NSS_round_60

Independent variables and diagnostic values	Dependent variable M_1		Dependent variable $sqrtM_2$	
	PROBIT		TOBIT	
	Regression 1	Regression 2	Regression 3	Regression 4
lnMPCE	0.126*	−3.74***	1.46***	−16.55***
sqlnMPCE		0.285***		1.317***
LCI		0.0306***		0.158***
sqLCI	0.00004***	−0.0002**	0.0001*	−0.0009**
FUEL	−0.0375	−0.089	0.200	−0.0834
EDU	−0.002*	0.0065**	-0.0098*	0.028*
sqEDU		−0.0001***		−0.0004***
Constant	−0.994**	10.99***	−9.75***	46.10***
LR χ^2(k)	20.39***	52.95***	37.40***	68.98***
No. of observations	1877	1877	1877	1877

***significant at < 1 % level; **at < 5 % level; *at < 10 % level

Regression Analysis

Two sets of regressions are run with two alternative representations of morbidity (***M_1 & $sqrtM_2$***). Since the former one is a dummy variable we have run PROBIT regression and for the latter index we have run TOBIT regression with truncation point at value 0. The results are reported in Table 21.4.

From Table 21.4 a few observations are in order. The value of LR χ^2 (k) is significant at less than 1 % level for all cases making all regressions statistically acceptable. Regressions (1) and (2) have taken the incidence variable (M_1) as the study variable and regressions (3) and (4) have taken the index of morbidity (M_2) as the study variable. In regression (1) three regressors, viz., *lnMPCE, sqLCI,* and *EDU* are significant, but the sign of ***lnMPCE*** and ***sqLCI*** is positive and the most important influence i.e., *FUEL* is statistically insignificant with a negative sign.

As mentioned earlier that one possibility may be the inherent weakness in taking FUEL as the proxy for pollution exposure and it needs to be replaced by a better measure. The other may be the presence of some degree of nonlinearity in the proposed relationship. To be more specific, marginal increase in income, living standard or the level of education may not affect the decision regarding fuel use and morbidity status of the household significantly. However, when these improvements are substantial, the favorable effect on the study variable may be realized. That means, there may exist some threshold levels after which all the proposed explanatory factors would be causally effective. To verify this possibility we have additionally introduced ***(lnMPCE)^2***, ***(EDU)^2***, and ***(LCI)*** as three other regressors. The results are reported under Regression 2.

For both MPCE and EDU, the variables have significant influence on the incidence of morbidity at level and square with opposite but expected signs. As MPCE increases M_1 decreases up to a certain level but after a threshold level of income the effect gets reversed. At a moderate level of income an improvement in

Table 21.5 Descriptive statistics (NSS 60th round): clean and dirty fuels

Descriptive statistics	*LCI*		*EDU*		*MPCE*		*M2*	
	Clean	Dirty	Clean	Dirty	Clean	Dirty	Clean	Dirty
Mean	88.31	70.58	78.37	38.65	1439.00	700.40	19.47	16.56
Standard deviation	9.45	21.28	26.78	31.12	838.20	411.20	25.99	25.12
Skewness	−1.64	−0.73	−1.31	0.15	2.44	5.30	1.59	1.97
Kurtosis (STATA)	7.87	2.64	4.12	1.94	13.23	62.13	5.08	6.64
Total	921	957	921	957	921	957	921	957

expenditure leads to better living standard and, therefore, a better health status. However, at a very high level of income due to change in lifestyle and change in health-perception the reported morbidity may increase with further enhancement of income. Similarly, for EDU the favorable effect of educational attainment on the incidence of morbidity is realized after a certain threshold level. Educational upliftment is necessary but any marginal increment may not alone be sufficient. There has to have a significant leap. Finally, with the introduction of threshold level the living condition index is also behaving in the same way as the level of education. However, here the incremental effect is yet to be statistically significant.

Next we have carried out regression exercises by taking M_2 as the study variable. When M_2 is regressed on *lnMPCE*, *sqLCI*, *FUEL,* and *EDU* for the first time FUEL has the desired sign (though statistically insignificant) and lnMPCE and sqLCI are significant with wrong sign. So, the additional regressors are introduced in regression (4) to verify the presence of threshold effects. The outcome was remarkable. Except for FUEL everywhere the variables are statistically significant with desired signs and the threshold level of ***lnMPCE*** is 6.28 (MPCE is rupees 535.51), of ***LCI*** is 87.78, and of ***EDU*** is 35 (implying a primary level education). Since in our dataset the mean value of EDU was higher than this threshold level so, it showed the expected sign and the mean value of LCI was lower than this threshold and so here the expected sign was not observed initially. Finally, some comments should be made on the constant term. The constant term in regression (4) is significant and positive, indicating the presence of morbidity (to some extent) even in the absence of all these explanatory factors.

So far our analysis has succeeded in finding out the threshold levels which have important bearing on the assessment of causal influence of different exogenous variables on the study variable and, therefore, form the pragmatic basis of policy evaluation. At the end of this analysis it must be admitted that we have failed to establish the expected link between the nature of energy use and the extent of morbidity at the household level.

Before going into any subsequent analysis here we go for a further check for the relationship between the fuel type and the morbidity for the households of urban West Bengal simply by exploring the descriptive statistics of the relevant variables for the state. This is shown in Table 21.5. The table shows that the clean and dirty fuel classification of all other variables show significant differences, except for M_2 (Morbidity index). Minimum difference is found in case of morbidity for all the

statistics. In fact average (mean) morbidity is lower in case of dirty fuel using households than that for clean fuel using households. Apparently this seems to be quite surprising. But this implies that fuel type is not a very good proxy for IAP while explaining household level morbidity in West Bengal.

So, we have to propose a better measure of air pollution exposure which will utilize information on not only the type of fuel used but on the extent of time spent by each member of the household in different microenvironment, the vulnerability of each member reflected in terms of age-sex composition of the household, the concentration of pollution in the ambient air, and all other relevant details. To collect such specific and detailed information a primary survey has been conducted on the city of Kolkata during 2008–2009.

Primary Survey on Kolkata: Design and Frame

Household Survey: The Basic Frame

A detailed household survey has been conducted on 80 wards of the Kolkata Municipal Corporation (KMC) out of 141 wards, spread over East, West, North, and South regions of the city between November 2008 and February 2009.[5] Information was primarily collected on demographic profile, living condition, levels of education, type of occupation, and extent of pollution exposure in terms of time-use pattern as well as environmental perception. An attempt has been made to assess the impact of pollution on local residents in terms of an analysis of the nature of morbidity experienced by them during the recent past (with one year as recall period) and to record the way they correlate these effects with the quality of their living environment. From each of these purposively selected wards four roads were randomly selected and from each road with a random start, following the right hand rule, 7 or 8 household were interviewed. In all, responses were collected from 2,396 households and with average family size of 4.2 that generated a database of nearly 10,000 people.

In accordance with population density we have selected 22 wards from North Kolkata, 24 wards from South Kolkata, 16 wards from East Kolkata, and 18 wards from West Kolkata (Fig. 21.3).

Out of these 80 wards, for 30 wards National Environmental Engineering Research Institute (NEERI) collected air quality data for us as part of another research project titled "Assessment of Damage Cost in Old Tannery Complex of Kolkata" funded by the Department of Environment, Government of West Bengal during November 2008–February 2009. These 30 wards are ward 28, 33, 34, 35,

[5] The survey was conducted as part of a UGC Project funded under UPE Scheme titled 'Spatial Concentration of Pollution and Morbidity Impact in the City of Kolkata: Efficient Intervention Strategies'.

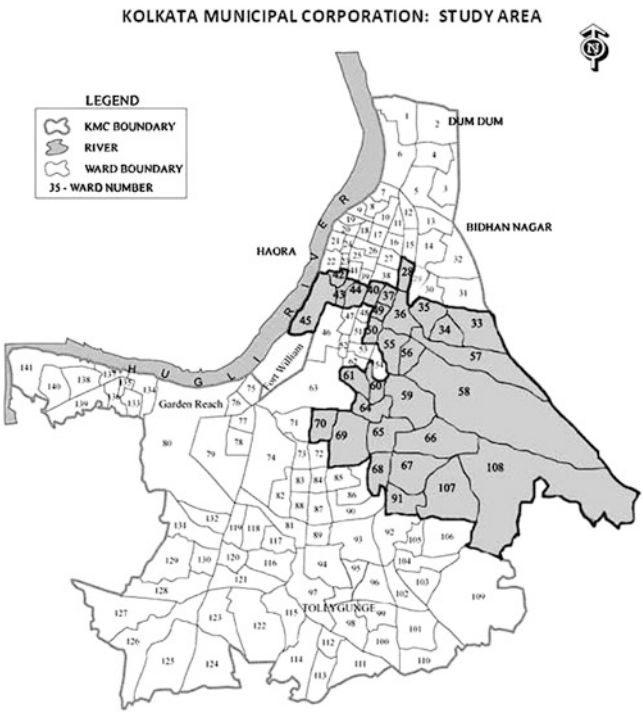


Fig. 21.3 Kolkata Municipal Corporation—study area

36, 37, 40, 42, 43, 44, 45, 49, 50, 55, 56, 57, 58, 59, 60, 61, 64, 65, 66, 67, 68, 69, 70, 91, 107, and 108, which are mostly spread over East–West direction in and around central Kolkata.

So, as an illustration of the methodology that have been proposed to construct a pollution exposure index we are concentrating on these 30 wards for the time being. Table 21.6 reports the excess concentration of RPM in these 30 wards. In the survey it was possible to collect reliable information on demographic profile, living condition, kitchen configuration, fuel type, levels of education, type of occupation, and extent of pollution exposure in terms of time-use pattern as well as environmental perception. No credible information could be collected on income or expenditure of the households because of the reluctance of the respondents to provide this particular information. An attempt has been made to assess the impact of pollution on local residents in terms of an analysis of the nature of morbidity experienced by them during the recent past (with one year as recall period) and to record the way they correlate these effects with the quality of their living environment.

Since the economically weaker section is likely to reside in more congested areas, in the absence of income data we have considered the percentage of slum population in each ward to represent the overall economic standing of the ward.

Table 21.6 Excess RPM concentration in the survey-wards

RPM concentration ($\mu g/m^3$): standard = 100 $\mu g/m^3$

Ward no.	RPM	Excess	Ward no.	RPM	Excess	Ward no.	RPM	Excess
28	347	247	45	141	41	64	204	104
33	359	259	49	261	161	65	437	337
34	272	172	50	238	138	66	494	394
35	418	318	55	448	348	67	254	154
36	121	21	56	403	303	68	219	119
37	259	159	57	252	152	69	379	279
40	153	53	58	247	147	70	169	69
42	266	166	59	258	158	91	356	256
43	145	45	60	174	74	107	119	19
44	125	25	61	182	82	108	262	162

Source NEERI 2009

To form some a priori idea about pollution exposure information was collated from the West Bengal Pollution Control Board (WBPCB) on the share of polluting industries[6] out of total industries in these wards to assess the extent of industrial pollution and the proximity to major arterial roads to assess the extent of vehicular pollution. The following characteristics have been considered:

a. Density of slum population in the selected wards;
b. Concentration of polluting industries in the selected wards;
c. Intensity of vehicular pollution in terms of proximity to major arterial roads.

High density of slum population implies the wards are congested enough with poor quality of in-house ventilation leading to increased intensity of indoor smoke. In addition, industrial and vehicular pollution worsens the ambient air quality which influences the indoor air considerably. In fact, the dose–response to pollution exposure, thus, depends on (a) outdoor environment, (b) indoor environment, and (c) time spent in different microenvironment at different age by different sex.

In Table 21.7 the classification of the wards is shown in terms of three criteria—slum settlements, concentration of polluting industry, and concentration of RPM.

Wards 40, 43, 45, 61, and 70 are the cleanest wards in terms of all criteria. These are the wards which are not infested by slums. Low concentration of both polluting industries and RPM are found as well. In contrast, with respect to all three criteria wards like 33, 34, 57, 58, 59, 66, and 67 are the most polluted wards with high slum population, high intensity of industrial pollution as well as high levels of RPM. 44, 107, and 60 are non-slum wards with high concentration of polluting industries and low RPM level. This may be due to the fact that in these

[6] Red and Orange category industries as per POLUACT Code.

Table 21.7 Three-way classification of survey wards: slum, industry, and air quality

RPM concentration	Concentration of polluting industries				
	High		Low		Total
	Slum	Non-slum	Slum	Non-slum	
High	33, 34, 57, 58, 59, 66, 67	–	28, 35, 37, 56, 64, 65	42, 49, 50, 55, 68, 69, 91, 108	21
Low	–	44, 60, 107	36	40,43,45,61,70	9
Total	10		20		30

wards industries release wastes that mainly affect soil and water source. Smoke emitted from these industries has less bearing on the RPM concentration of the locality, although it exceeds the stipulated standard as mentioned earlier. Industrial pollution is not a problem for ward 36. Moreover, this slum ward is not also facing any substantial threat of high RPM level; in fact it is very close to the cleanest ward in terms of RPM concentration (21 $\mu g/m^3$). Ward 36 can be considered as having the cleanest slum in the study area. No slum ward exists with high intensity of industrial pollution and low RPM concentration. It implies that if polluting industries concentrate around the slum areas, low concentration of RPM is not likely in those locations. Mainly because of the congested nature of the slum settlements air does not get enough scope to be dissipated in the atmosphere making RPM level high at the time of measurement. In spite of being a non-slum as well as low industry ward 42, 49, 50, 55, 68, 69, 91, and 108 are having a high level of RPM. This indicates the possible presence of a third factor which is responsible for the high pollution concentration in these localities. High levels of vehicular pollution arising from the flow of heavy traffic in roads adjacent to these wards have been found to be a major source explaining high RPM concentration in the absence of any significant contribution of industry and slum.

Following the foregoing discussion the wards have been clubbed into three distinct groups, namely, **dirty** (with high industry, high slum, and high air pollution), **clean** (low concentration of polluting air), and **moderate** (where either industry is not polluting or the ward is mostly non-slum type but the quality of ambient air is showing high concentration). Dirty group contains wards 33, 34, 57, 58, 59, 66, 67 (7 wards in all); clean wards are 36, 40, 43, 44, 45, 60, 61, 70, 107 (9 wards in all); and the moderately polluted wards are 28, 35, 37, 42, 49, 50, 55, 56, 64, 65, 68, 69, 91, 108 (14 wards).

On each member of the sampled household information was collected on the number of illness episode, nature of illness, and types of treatment offered. Figure 21.4 represents percentage of morbid persons in each ward with a black arm and that of healthy (non-morbid) persons with a white arm. In general, the *dirty* wards are having a longer black arm and the opposite is true for the *clean* wards except ward no. 107 indicating poorer health status in *dirty* wards. In the *moderately polluted* wards a mixed picture is observed. Since the wards are heterogeneous in nature, as mentioned earlier and pollution concentration is *high*

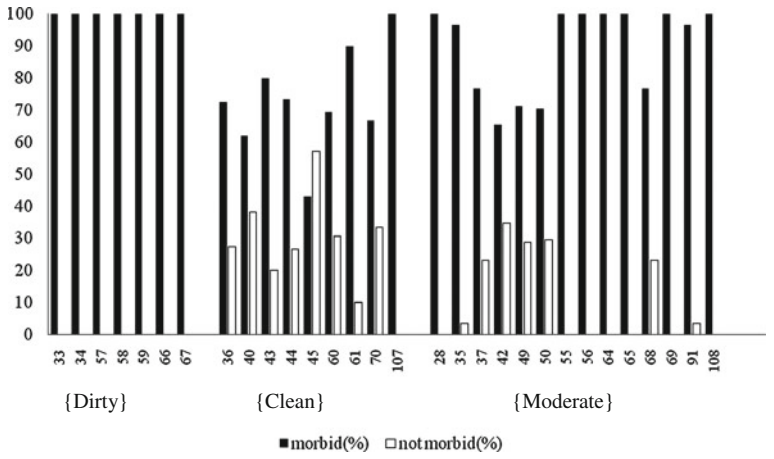


Fig. 21.4 Ward wise incidence of morbidity

everywhere, some of the wards are having higher percentage of morbid people than the non-morbid. This bar-diagram lends support to the basic surmise that the pollution, in any form, had contributed toward an enhancement of morbidity.

Construction of Exposure Index

Ideally speaking, a household index of exposure should be sensitive to (i) the ambient level of pollution concentration, (ii) time spent by the exposed person in a specific microenvironment and (iii) the age and sex of the person concerned. While item (i) and (ii) try to assess the intensity of exposure, item (iii) measures the relative vulnerability of the concerned individual. Denoting ambient level of pollution concentration (in excess of the recommended NAAQM standard) in location k by a_k, exposure weight of the jth micro-environment by c_j, time spent by the ith member of the lth household in the jth micro-environment at location k by t_{ijkl} and age-sex vulnerability weights of the ith member by $w_{i,(\mathrm{as})}$ the final exposure index for the lth household located in the kth Pollution Monitoring Station (PMS) may be developed at three stages.

Stage I. The intensity of exposure for individual i in household l at location k: $P_{ikl} = \sum_j c_j t_{ijkl} \Big/ \sum_j t_{ijkl}$, where $\sum_j t_{ijkl} = 24$;

Stage II. The vulnerability of household l at location k: $V_{lk} = \sum_i w_{i,(as)} P_{ikl} \Big/ \sum i$; where $\sum i = household - size$, and in the PMS k for household l it is V_{lk};

Stage III. The Exposure Index of household l located in PMS k for *RPM*: $a_k V_{lk}$.

We need three sets of weights to construct EXI, the exposure index, viz., a_k, which has been obtained from NEERI and WBPCB records; $w_{i,(as)}$, which has been derived from Kathuria and Khan (op cit.) and c_j, which has been generated from the estimates extended by Suksena and Dayal (op cit.) and Ostro (1994). Table 21.8 presents the weighting scheme followed in EXI construction.

In Suksena and Dayal (op cit) intensity of exposure in cooking as well as non-cooking activities was given for different types of fuel in rural, urban, and slum areas. We have considered urban values to correspond the best possible kitchen (separate kitchen), slum values with the worst possible kitchen (cooking inside the bedroom), and rural value to the intermediate one. They have weights for primary, secondary, and tertiary activities. We have collected information on high risk, moderate risk, and low risk jobs and hence, matched the value of primary activity, where the exposure is direct, with high-risk job, secondary activity with moderate risk job, and tertiary activity with low risk job. Given the high level of air pollution in Kolkata, travel has been equated with moderately risky venture and school time and outdoor leisure time as low risk activities. In the absence of any time-use survey data, Kathuria and Khan (op cit) proposed 5 h exposure for adult male and adult female, 6 h for old female, 8 h for children, and 9 h for old male. Considering the exposure rate of adult male and female as norm we have calculated the vulnerability weights for each age-sex based group as shown in Table 21.8.

Incidence Analysis

On the basis of this weight structure and the methodology described above, the pollution exposure has been calculated broadly in two steps—first for the individuals of the surveyed wards and then for the households located in any particular pollution monitoring station. Individual level exposure, again, is calculated at three stages—first without putting any vulnerability weights (raw, K_i for say, ith individual), then with the vulnerability weights ($K_{iv} = w_{i,(as)}P_{ikl}$) and finally by constructing an index of the exposure obtained with vulnerability weights. The index at the individual level is needed to be constructed because the exposure with vulnerability weights has a huge range of values and requires to be normalized. So we calculated the indexed exposure (KI_{iv}) as

$$[K_{iv} - \min(K_{iv})]/[\max(K_{iv}) - \min(K_{iv})]$$

In the following table (Table 21.9), we report the median values of these three exposure levels for each of the demographic classes mentioned in Table 21.8. Three clear observations can be figured out from the table:

a. First, our sample contains 69 % adult, 9.6 % old, and 21.4 % children. In slum area, the exposure is the highest for the old females and in non-slum area children are facing the maximum vulnerability. The slum–non-slum difference

Table 21.8 Exposure weights

Time-use based relative intensity weights

Kitchen type	Indoor time				Outdoor time					
	Cooking		Other		Work			School	Travel	Other
	Clean fuel	Dirty fuel	Clean fuel	Dirty fuel	High risk	Moderate risk	Low risk			
Separate	0.4	1.1	0.1	0.4	0.5	0.3	0.2	0.2	0.3	0.2
Balcony	0.4	1.4	0.1	0.7	Age-sex-based relative vulnerability weights					
Bedroom	0.5	1.5	0.2	1.5	Adult male	Adult female	Old male	Old female	Children	
					1.0	1.0	1.8	1.2	1.6	

Source Constructed by extracting information from different sources

Table 21.9 Pollution exposure in Kolkata across demographic classes

Raw (K_i)					With vulnerability weights (K_{iv})			With vulnerability weights indexed (KI_{iv})		
	Total obs. (%)	Slum	Non-slum	% difference	Slum	Non-slum	% difference	Slum	Non-slum	% difference
Adult male	1458 (36.95)	0.44	0.41	6.82	0.44	0.41	6.82	13.08	11.82	10.05
Adult female	1264 (32.03)	0.54	0.26	51.85	0.54	0.26	51.85	16.83	6.25	62.86
Old male	202 (5.12)	0.57	0.20	64.91	1.026	0.36	64.91	35.52	10.29	71.03
Old female	176 (4.46)	0.66	0.23	65.15	0.792	0.276	65.15	26.54	6.54	75.36
Children	846 (21.44)	0.58	0.54	6.90	0.928	0.864	6.90	31.79	29.36	7.64

in exposure is relatively less striking for adult male and children, indicating high level of vulnerability for the children in general. For adult female and old people, both male and female, the situation in non-slum area is not so alarming.

b. When this pattern of exposure obtained from the time-use data is adjusted for the vulnerability weights suggested by Kathuria and Khan (op cit.) the absolute exposure level becomes most alarming for old male and the best for the adult female in non-slum areas.

c. Finally, considering the average exposure index, constructed to compare the normalized relative positions, the female in non-slum area are least exposed (both adult and old), children in general are highly exposed, and the old people in slum are revealing serious vulnerability.

Data Exploration

Using the information collected on sample households we have constructed study variables ***M_1, M_2, LCI***, and ***EDU*** as before and instead of ***FUEL*** we have constructed **EXI** as the exposure index. In our survey no information was collected on the income level of the household and hence, we have no variable corresponding to ***MPCE*** of the NSSO data set. Moreover the components of the LCI are marginally different. We had information on house type, house structure, type of latrine, and the source of drinking water; in addition the congestion rate has been taken as the average room occupancy rate of the household. Unlike NSSO we did not collect information on drainage system, so that dimension has been dropped from the LCI. Moreover, information on house type (own, rented, etc.) has not been used as this has a little relation with the living condition of the household. Table 21.10 reports the descriptive statistics used for the choice of proper functional forms of the relevant variables.

The incidence of morbidity, i.e., M_1 is not a very balanced dummy in this case as only 12.83 % of households are free from all ailment episodes and more than

Table 21.10 Descriptive statistics (primary survey)

Descriptive statistics	M_1	M_2	*LCI*	*EDU*	*EXI*
Mean	0.87	28.97	55.83	58.93	119.57
Median	1.00	28.57	51.95	67.00	61.43
Standard deviation	0.33	19.74	20.04	34.03	134.11
IQR	0	27.5	31.32	67.00	147.50
Pseudo standard deviation	0	20.39	23.22	49.67	109.3
Skewness	−2.22	0.71	0.25	−0.36	1.63
Kurtosis (normalized)	2.94	1.24	−0.79	−1.01	2.19
Sample size	896	896	896	896	896
Observations with 0 value	115				
Observations with 1 value	781				

Table 21.11 Pair-wise correlations

Variables	M_1	*M2*	*LCI*	*EDU*	*lnEXI*
M1	1				
M2	0.5634*	1			
LCI	−0.1191*	−0.0883*	1		
EDU	−0.1529*	−0.2332*	0.4159*	1	
lnEXI	0.2191*	0.3071*	−0.4734*	−0.3735*	1

* significant at <5 % level

87 % are victim of morbidity. So, it is better to consider M_2, the morbidity index as the relevant study variable. For M_2, the mean value is approximately same as the median value and the standard deviation is smaller than the pseudo standard deviation indicating thin tail almost symmetric distribution. Here the ladder analysis has shown best fit χ^2 for the raw data. So, no further transformation of M_2 has been attempted. For the living condition index ***LCI***, the mean is greater than median and the standard deviation is less than the pseudo standard deviation ruling out any possibility of thick tail with outliers at the ends. The ladder analysis has shown very satisfactory goodness of fit for the raw from of the variable and we have taken ***LCI*** as it appears from the data. For ***EDU*** median exceeds mean with greater pseudo standard deviation than standard deviation. This is an indication of thin tail. With very small and negative values of both skewness and kurtosis the raw variable turned out to be the best functional form.

Finally, for ***EXI*** mean is higher than median with presence of outliers at the high end; and from the ladder-analysis a logarithmic transformation appeared most appropriate. So, we have taken ***lnEXI***. Table 21.11 presents the pair-wise correlation of all these relevant variables. Interestingly, unlike ***FUEL*** in the analysis of West Bengal (Urban) discussed earlier, exposure index has positive correlations with all the morbidity variables. Both of ***EDU*** and ***LCI*** are correlated with the morbidity variables with desirable signs. Because all the correlations are statistically significant we have regressed $\boldsymbol{M_2}$ on the explanatory variables and check for the values of Variance Inflation Factors (VIF) to detect presence of multicollinearity in the dataset. The mean VIF turned out to be only 1.35 with a range of 0.06. Hence, no further adjustment seemed necessary to avoid inter-cause interdependence.

Regression Analysis

We have run PROBIT with $\boldsymbol{M_1}$ as the study variable and both simple OLS and TOBIT[7] for $\boldsymbol{M_2}$. To check for the possible presence of heteroskedasticity in the

[7] It is more common to run TOBIT regression with any index as a dependent variable (as the value of an index gets truncated at 0 by construction).

cross-section data GLS with robust variance has also been tried. In the PROBIT regression of M_1, since the number of healthy households is only 12.83 %, hence the regression coefficients are not very interesting to be analyzed. We will mostly concentrate on the regression of M_2. The results are reported in Table 21.12. In all the regressions, the prevalence of morbidity goes down with an improvement in the condition of living and the level of education of the head of the household.

It is obvious from the table that relationship between morbidity status and air pollution exposure is direct. All these relationships are statistically significant and the overall statistical power of all the regressions is fairly strong. This is a clear improvement over our earlier results where type of fuel failed to represent the extent of IAP at the household level and came out to be inadequate to explain household morbidity. In all variants of the regressions, exposure index is coming out to influence morbidity with desirable sign and with strong statistical significance (less than 1 %). Except the PROBIT regression, for all other regression the coefficient of ***lnEXI*** takes a fairly large value implying a strong incremental effect of the exposure level on morbidity. Our hypothesis regarding the relation between IAP and morbidity gets strong support in this situation.

Another important observation from the regressions mentioned is that ***LCI*** has a threshold effect, like West Bengal case, on household morbidity of the study regions. To verify the robustness of the estimated regressions in terms of their predictive power we have estimated the mean predicted value of M_2 from each one of the regression. The predicted values for OLS and GLS regressions turned out to be exactly same (34.92) and that from TOBIT regression is 39.065. It is quite apparent from the coefficients of different variables obtained across different models. The more or less similar values of coefficients (for OLS and GLS variation they are exactly similar) ensure that estimated values for M_2 would be close alike. The robustness has been verified with the threshold values of the ***LCI*** as well. Threshold level of ***LCI*** is 70.85 for both OLS and GLS model and it is 69.93 for TOBIT model. So, the results are robust with respect to model specification and we are reporting below some sensitivity analysis carried out by using the estimated model.

Concluding Remarks

Among the wards surveyed, ward number 107 has the cleanest environment where the average RPM level is exceeding the standard by 19 $\mu g/m^3$ only. If all wards can be brought to this level of cleanliness the average value of the morbidity index M_2 will go down by 24.82 % and will attain a lower value of 29.41 only. In fact, a 50 % reduction in the pollution level will lower morbidity by 9.05 and a 10 % reduction will lead to a 1.30 % reduction (Table 21.13).[8] It is apparent from the

[8] All M_2 values are calculated from the threshold TOBIT regression.

Table 21.12 Regression results from primary survey

Independent variables and diagnostic values	Dependent variable M_1		Dependent variable M_2					
	PROBIT		OLS		GLS		TOBIT	
	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
LCI	0.000957	0.01731	0.1249***	0.666***	0.1249***	0.666***	0.1297***	0.7413***
sqLCI		−0.000139		−0.0047***		−0.0047***		−0.0053***
EDU	−0.00471**	−0.00488**	−0.1004***	−0.1050***	−0.1004***	−0.1050***	−0.1109***	−0.1160***
lnEXI	0.2685***	0.2596***	4.7574***	4.5163***	4.7574***	4.5163***	5.3754***	5.1143***
Constant	0.3501	−0.02799	8.3326**	−4.1758	8.3326**	−4.1758	4.7997	−9.4083
No. of observations	896	896	896	896	896	896	896	896
$\bar{R}^2$			0.119	0.128	0.122	0.1323		
Pseudo R^2	0.075	0.077						
F			41.35***	33.95***	50.06***	39.71***		
LR $\chi^2(3)$	51.89***	53.02***					116.91***	126.88***
Breusch-Pagan $\chi^2(1)$								

***significant at <1 % level
**significant at <5 % level

Table 21.13 Sensitivity of morbidity to variation in air quality

Reduction in average excess RPM(%)	Reduction in M_2	Change (%)
10	0.51	1.30
20	1.13	2.89
50	3.53	9.05
To 19 μg/m^3	9.71	24.82

Source Author's calculation

table that the reduction in the level of morbidity following an improvement in the quality of respirable air, given everything else, is a nonlinear relation and if the air quality can be improved up to a threshold level the favorable effect will be significant. If along with an improvement in air quality the level of education and the quality of living condition can also be improved simultaneously, the effect would be magnified further.

Improvement in the ambient air quality is a public affair whereas living condition, fuel choice, kitchen configuration, etc., are matters of private household decision. Although these latter variables have their impact on air-borne morbidity, they are not purely external like the quality of ambient air, rather are endogenous to the household's decision-making process. Education is another public policy issue which plays an important role in creating awareness regarding the adverse health consequences of pollution. So formulation and implementation of policy in the context of IAP mitigation is not related to any single dimension, rather it is a complex combination of state and household affair.

References

Brooks N, Sethi R (1997) The distribution of pollution: community characteristics and exposure to air toxics'. J Environ Econ Manag 32:233–250

Dasgupta S, Huq M, Khaliquzzaman M, Pandey K, Wheeler D (2004) Indoor air quality for poor families: new evidence from Bangladesh. World Bank Policy Research Working Paper 3393

Dasgupta S, Huq M, Khaliquzzaman M, Pandey K, Wheeler D (2006) Who suffers from indoor air pollution? Evidence from Bangladesh, The World Bank, OUP

Duflo E, Greenstone M, Hanna R (2008) Indoor air pollution, health and economic well-being in SAPIENS

Kathuria V, Khan NA (2007) Vulnerability to air pollution: is there any inequality in exposure? In: Economic and Political Weekly, Sameeksha Trust, Mumbai

Lahiri T, Roy M, Lahiri P (2001) Assessment of health impacts of indoor air pollution in Calcutta. DoE, GOW

Martins HC, Fatigati HFL, Vespoli TC, Martins LC, Pereira LAA, Martins MA, Saldiva PHN, Braga ALF (2004) Influence on socio-economic conditions on air pollution adverse health effects on elderly people: an analysis of six regions of Sao Paolo, Brazil. J Epidemiol Commun Health 58:41–46

Mukherjee C, White H, Wuyts M (1998) Econometrics and data analysis for developing countries. Routledge, UK

Murty MN, Gulati SC, Banerjee A (2003) Health benefits from urban air pollution abatement in the Indian subcontinent. Discussion paper no. 62/2003, Institute of Economic Growth, Delhi
Neidell MJ (2004) Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. J Health Econ 23(2004):1209–1236
Ostro B (1994) Estimating the health effects of air pollutants: a method with an application to Jakarta. The World Bank [Policy Research Working Paper 1301], Washington, DC
Ou CQ, Hedley AJ, Chung RY, Thach TQ, Chau YK, Chan KP, Yang L, Ho SY, Wong CM, Lam TH (2008) Socioeconomic disparities in air pollution-associated mortality in environmental research 107
Reddy BS (2004) Economic and social dimensions of household energy use: a case study of India. In: Proceedings of IV Biennial international workshop on "advances in energy studies", June, 2004
Saksena S and Dayal V (1997) Total exposure as a basis for the economic valuation of air pollution in India. Paper presented at 20th annual international conference on energy and economic growth, New Delhi, 22–24 Jan 1997
Suksena S, Singh PB, Prasad RK, Prasad R, Malhotra P, Joshi V, Patil RS (2003) Exposure of infants to outdoor and indoor air pollution in low-income urban areas: a case study of Delhi. J Exposure Anal Environ Epidemiol 13(3):219–230
Smith KR (1996) The Natural Debt: North and South. In: Giambelluca T, Henderson-Sellers A (eds) Climate change: developing southern hemisphere perspectives. J. Wiley & Sons, UK, pp 423–448
Smith KR (2000) National burden of disease in India from indoor air pollution. PNAS 97(24):13286–13293
Smith KR (2008) Wood, the fuel that warms You thrice. In: Colfer CJ (ed), Human health forests: a global overview of issues, practice and policy, Earthscan, London. pp 97–111
Wheeler D (2000) Greening industry: new roles for communities, markets and Governments, OUP, New York
WHO (2000) Addressing the link between indoor air pollution, household energy and human health, based on the WHO-USAID global consultation of the health impact of indoor air pollution and household energy in developing countries (Meeting Report), May, 2000
WHO (2006) Reports of a symposium and a workshop held at the international society of environmental epidemiology. In: Indoor air pollution and lower respiratory tract infections in children, Paris, Sept 2006
WHO (2009) Global health risks—mortality and burden of disease attributable to selected major risks
World Development Report (1992, 2002, 2003) The World Bank, OUP
www.who.int (accessed in April-May, 2011)
www.worldbank.org (accessed in April-May, 2011)

Chapter 22
Efficient Pollution Management Through CETP: The Case of Calcutta Leather Complex

Subhra Bagchi and Sarmila Banerjee

Introduction

India's leather industry occupies a prominent place in the export market, generating scope for foreign exchange earning as well as large-scale employment. The export potential of this industry was recognized by the Government of India in the mid-1970s when it evolved a policy package consisting of a ban on export of raw hides and skins and providing fiscal and other incentives to stimulate the export of finished leather and leather products. Much of the economic benefits derived from leather production and trade; however, have come at a considerable cost to the environment and human health, which should be attended simultaneously to enjoy sustainable benefits from the industry. This poses a serious challenge before the sector and a number of strategies have been contemplated since the early 1990s to give the industry a cleaner shape.

The leather industry may be regarded as a bridge between the production of the hide as by-product of the food industry and its manufacture into shoes and wearing apparel for which it provides the basic raw material. The production of leather is a long and complicated process which requires specialized skills. The first step is the removal of unwanted components leaving a network of fibers of hide and skin, through a series of pretanning operations. The second step is dealing with tanning materials to produce a stabilized fiber structure. This ensures the permanent preservation. The third step is to build into the tanned fibers characteristics of fullness, color, softness, and lubrication and to finish the fiber surface, to produce a useful product (Thorstensen 1969). As the leather industry, particularly the tanning

S. Bagchi (✉)
Vidyasagar College for Women, 39, Sankar Ghosh Lane, Kolkata,
West Bengal 700006, India
e-mail: shubhrabagchi@gmail.com

S. Banerjee
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: sarmila_b@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_22, © Springer India 2013

segment is a pollution-intensive process it has been put under Red category by the Central Pollution Control Board of India. Tannery effluents contaminate air, surface and subsoil water, and undermine the fertility of the soil. The continuous discharge of untreated tannery effluents on land, rivers, and other water bodies has adversely affected the quality of water and agricultural land in the regions in which tanneries are located. Moreover, this industry is characterized by small-scale units which are technologically and financially incapable of undertaking pollution abatement at the individual level. Hence, what is required is a pollution management joint action. The Common Effluent Treatment Plant is a way by which the small tanneries can comply with the norms laid down by the regulatory authority.

With increasing public awareness about environmental damage caused by tanneries, many NGOs in different parts of the country began to protest loudly. In a landmark judgment delivered by the Supreme Court of India in 1996 against a petition filed by the Vellore Citizen's Forum, more than 400 tanneries of Tamil Nadu were closed for their failure to treat generated waste. Similar public-interest litigation cases in Uttar Pradesh and West Bengal have culminated in Supreme Court Orders for relocation of tanneries or closure of tanneries which do not have their own individual effluent treatment plants(IETPs) or which are not connected to common effluent treatment plants (CETPs) .

Although the industry began to tackle environmental pollution rather slowly in the early 1990s, by the turn of the century, its achievements have been acknowledged by many as truly remarkable. The concept of Common Effluent Treatment Plants has taken deep roots in this industry. In an industry dominated by small-scale units, this has been recognized as an effective strategy. The CETP is found to work efficiently in Tamil Nadu; however, the case of Calcutta Leather Complex in Bantala is far from satisfactory. Though the industry is only too well aware that its ability to overcome environmental challenges is a prerequisite for its continued growth, the management of CETP has started posing some special problems. Because of the public good nature of the plant, here the members generally have an intrinsic tendency to free ride. Therefore, designing of incentive compatible cost sharing rules to ensure voluntary participation of all concerned stakeholders is an analytical challenge.

In this backdrop we have concentrated on the design and management practices adopted in the CETP of Calcutta Leather Complex where the system is technically feasible but still not economically viable. An attempt has been made to assess the existing level of efficiency and to recommend measures to improve it to make the project self sufficient in near future. "Calcutta Leather Complex" will briefly present the site planning, design, and location map of the Calcutta Leather Complex (CLC), " The Prevailing State" will describe the present situation, "CETP as a Cooperative Game" will vouch the case for cooperative institution-based mechanism design, "Efficient Cost Allocation for CETP in CLC" will assess the economic prospect of CLC, and finally, "Concluding Observations" will conclude the chapter.

Calcutta Leather Complex

In 1985 a public-interest petition (No/3727/1985, M.C Mehta versus Union of India and others) was lodged in the Supreme Court of India directed against the tanneries located in the city of Kanpur for discharging untreated effluent into the river Ganga. Subsequently, the scope of the petition was enlarged and the industries located in various cities on the bank of the river were called upon to stop polluting the river. In this process, about 550 tanneries located at Tangra, Tiljola, Topsia, and Pagladanga, the four adjoining areas in the eastern fringe of Kolkata (the Old Tannery Complex) were identified. Because of dearth of space for establishment of waste-treatment plant in the old tannery complex, the Supreme Court of India instructed the Government of West Bengal to make space available for the relocation of tanneries with modern facilities for effluent treatment. The Government of West Bengal, so directed by the Supreme Court, acquired 1100 acres of land in Bantala,[1] 24-Parganas (South) by July 1992 and the Central Leather Research Institute (CLRI) of Chennai prepared the blueprint of the proposed Calcutta Leather Complex in May 1993. It was agreed that the installation cost of this CETP would be shared by the Central and the State Government on a 50–50 basis where the Ministry of Environment and Forest of the Government of India would include this project in the second phase of the Ganga Action Plan. Accordingly, the Calcutta Leather Complex (CLC) was developed in Bantala and the common effluent treatment plant (CETP) was activated in July 2005. The CLC was supposed to accommodate not only nearly 550 existing tanneries (CLRI estimate) from the old complex, but it was designed to set up new tanneries with state-of -the-art technology and an *integrated leather hub* to support all types of units connected with leather trade.

Design of CETP

Keeping in mind the requirements of both relocated tanneries with dated technology with a possibility of gradual infusion of upgraded process as well as newly established tanneries with state-of-the-art technology the Government of West Bengal instructed its joint venture partner M.L.Dalmiya & Company Limited to set up a CETP with 30 MLD installed capacity. The CETP has been planned in six modules each of 5 MLD capacity level. The effluent from various tanneries would be pretreated in-house within the individual tanneries through processes like coarse screening to remove large floating matters like flesh, hair etc., and grease

[1] Subsequently an area of 1100 acre was earmarked for the integrated complex and was developed as a joint venture project on a Built Operate and Transfer (BOT) arrangement with M/s M.L Dalmiya & Co. Ltd as the private partner.

traps for removal of grease. This pretreatment would prevent clogging the Effluent Conveying System (ECS) to ensure its smooth functioning.

The individual tanning units would be connected with CETP through two sets of underground pipeline networks. The treatment will pass through three stages: (i) *Physical*: removal of physical impurities through physical treatments like screening, de-gritting, settling, etc.; (ii) *Chemical*: removal of chemical impurities by chemical treatment such as pH correction by alkali/acid addition or nutrient/alum/poly-electrolyte addition; (iii) *Biological*: removal of organic matter and biological impurities by biological process like Aerobic Activated Sludge Treatment System and like. The sludge formed during the physical, chemical, and biological treatment can be further thickened and dewatered to be used as soil conditioner. It is proposed further that after treatment the CETP will make the final discharged water to pass through a microbiological treatment to get back the required purity to be allowed to flow into the water bodies that supply process water for the system.

The Prevailing State[2]

In-House Treatment

This section reports the findings of a primary survey conducted on 201 tanneries in CLC during the winter of 2008–2009. The existing regulation has made it mandatory on the part of each tannery of the new complex to install in-house water meters (though we have encountered a few units without any such meter) at the inlet point. Even the tanneries that have taken water connection from Dalmiya & Co. have expressed serious reservation regarding the water quality. Nearly 68 % of the units interviewed consider this piped water unfit for tanning purposes due to excessive iron content, bad odor, and erratic supply. The CLC Complex is facing additional problems due to frequent power cuts. Since the process of leather production calls for continuous supply of water, to meet such emergency situations each unit was allowed to bore one in-house pump. Only 47 % of the tanneries visited were found to depend exclusively on Dalmiya's supply, whereas the rest depend on bore wells either solely or partly. There is no way to measure the volume of water used in the production process from the bore wells and here the tanneries are not paying any tax for using the effluent treatment facilities. In fact, inside the premise of some of the units the interviewers have found even dug-wells. So, in the absence of any appropriate monitoring mechanism the initial idea of collecting pollution charges through installation of water meter at the inlet point

[2] For a more detailed discussion confer Bagchi (Majumdar) and Banerjee (2012).

(by applying the polluter's pay principle) appears to be ineffective.[3] In presence of multiple sources of water supply, water meters at the outlet point would be more effective than at the inlet point.

Treatment Through CETP

In the CLC, out of the proposed six modules of CETP only four have been completed so far. The CETP has been set up to treat mainly Suspended Solids (SS), Biological Oxygen demand (BOD), Chemical Oxygen Demand (COD), Sulphide and Chromium. In 2010–2011 only 288 tanneries had started operation. The already constructed four modules have a capacity of 20 MLD. The tanneries operating in CLC were discharging around 18.83 MLD. Since the cost of running a CETP is quite high each tannery has to pay Rs.14.50 per KLD effluent generated. The effectiveness of CETP depends on the coordination among all the proposed treatment arrangements. If some of them are absent and some others are improperly done then very little can be expected from the final outcome. Proper in-house physical treatment through screening, settling, etc., of the liquid effluents by the individual tanneries before sending it to the CETP is a necessary prerequisite. However, our primary survey revealed that in many cases the effluents were being discharged directly in the rainwater channels. Moreover, since the presettler units are not used properly and adequately in many individual tanneries solid wastes are entering the underground High Density Polyethylene (HDPE) pipeline directly, leading to clogging, cracking, and chocking of the network.

From a number of press release and media reports on the miserably poor performance of the Calcutta Leather Complex there is not much doubt left about the acute coordination problem among the stakeholders since the inception of the project. There is no incentive to make the effluent treatment process effective and efficient. In the following section we briefly state the requirements for economic efficiency in the context of a common facility. The subsequent section will apply this theory for the CETP in CLC.

CETP as a Cooperative Game

For proper functioning of a CETP one needs to derive conditions under which the players of the game, namely the tanners, are likely to cooperate to minimize the total cost of pollution abatement. For each player the share of cost to be borne under coalition should be less than his/her stand-alone cost (this is the participation

[3] Generally 90% of water inflow is expected to be discharged as effluent (Shanmugasundaram and Murthy 2000).

constraint, where every player would have the interest to participate) and there should be no incentive in terms of incremental cost to defect from the grand coalition by forming any sub-coalition (this is the incentive constraint). The solutions for which both these conditions are satisfied are said to form the **CORE** of the game. If the core is nonempty then CETP would be considered economically viable in terms of individual rationality, group rationality, and social desirability (in the sense of Pareto efficiency).

Definition and Existence of the CORE

For an N player game, let S be any coalition of n members where n ε [1, N] and $x^i$is the share of cost borne by the ith player. Core is a set of feasible payoff vectors X such that it satisfies

(a) Stand-alone cost test: $x^i \leq C^i \forall i; \sum_{i \in S} x^i \leq C\{S\}$where C^i is the stand-alone cost and $C\{S\}$ is the total cost borne by all members of the coalition by trying to operate individually or collectively through any sub-coalition indicating it is always better to join the coalition than to operate individually and

(b) Incremental cost test: $\sum_i x^i \geq \{C(N) - C(N-S)\} \forall S \subseteq N$ where $C(N)$-$C(N-S)$ is the incremental or marginal cost of any coalition S. The core condition eliminates all outcomes that any set of players can improve upon. It would seem reasonable to assume that if core exists then the imputation chosen should be in the core, since all (2^N-1) coalitions are accounted for. If the cost saving of a joint project increases with the increase in the size of the coalition then the core exists. So, the cost function has to be sub-additive, monotonic and consistent. However, the core may be empty. So, to apply this technique, one should check for the existence of core first, and then, go for the optimizing sharing rule for common cost within the core.

Uniqueness of the Solution

To arrive at a unique solution one may have to encounter either the problem of nonexistence of core or its nonuniqueness. In case of the former, one has to relax the constraints in defining the core and for the latter one has to strengthen the constraints to eliminate the relatively less desirable solutions in terms of some additional requirements. Here the Nucleolus of the core needs to be identified.

(a) *Nucleolus*: The idea of the nucleolus is to find a solution in the core that is central in the sense of being as far away from the boundaries as possible. It can be shown that this allocation makes the least well-off coalition as well-off as possible by choosing that imputation which minimizes the maximum complaint that any

coalition could have against any imputation. Coalition S is better off than T relative to an allocation x, if $c(S) - \sum_S x_i > c(T) - \sum_T x_i$. The quantity $e(x, S) = c(S) - \sum_S x_i$ is called the excess of S relative to x, which measures the amount (the size of inequity) by which coalition S falls short of its potential $c(S)$ in the allocation of x. Since the core is the set of imputations such that $\sum_S x_i \geq c(S)$for all coalitions S, the imputation x is in the core, if and only if, all its excesses are negative or zero. So, one has to find out an allocation x that minimizes the maximum excess $e(x, S)$ over all proper subsets $\phi \subset S \subset N$ is a problem in linear programming:

Max ε, subject to $e(x, S) \geq \varepsilon \forall S \neq \phi, N$; and $\sum_N x_i = c(N)$.

(b) *Shapley Value*: By solving a cost-minimizing type programming problem, Shapley Value technique suggests apportionment rule of the common cost among all players by satisfying both stand-alone cost and incremental cost constraints. This is one of the earliest methods of allocation to be based on a consistent set of postulates about how an allocation should be made. All players are assumed to sign up in some particular order. If a group S has already signed up and i was the last member of the group to arrive, his marginal cost contribution to S is [$c(S)-c(S\text{-}i)$]. The Shapley Value is i's average marginal contribution if all orders for signing up are assumed to be equally likely. It is defined as $x_i = \sum_{\substack{S \subset N \\ i \in N}} \frac{|S-i|!|N-S|!}{N!} C^i(S)$, where $C^i(S) = C(S) - C(S - \{i\})$is the marginal cost of i relative to S, and the sum is over all subsets of S containing i, and x_i is the share of i in the total cost. Shapley value allocation is fair in the sense that it is order free, symmetric, and it gives the dummy player his own worth only and not a share of the saving/cost to which he has not contributed. It includes all possible intermediate coalitions while calculating the share of a particular project and, therefore, is a combinatorial version of marginal analysis by protecting additivity, monotonicity, and consistency properties of cost allocation (Shubik 1984). If the core exists the Shapley allocation belongs to core.

Illustration

To illustrate the technique let us derive the characteristic function of a 3-player cost-minimizing game as follows:

Minimize any variable $c(1)$, $c(2)$ or $c(3)$, subject to:

$$c(1) \leq 81.7; c(2) \leq 46.1;\ c(3) \leq 12.5;$$

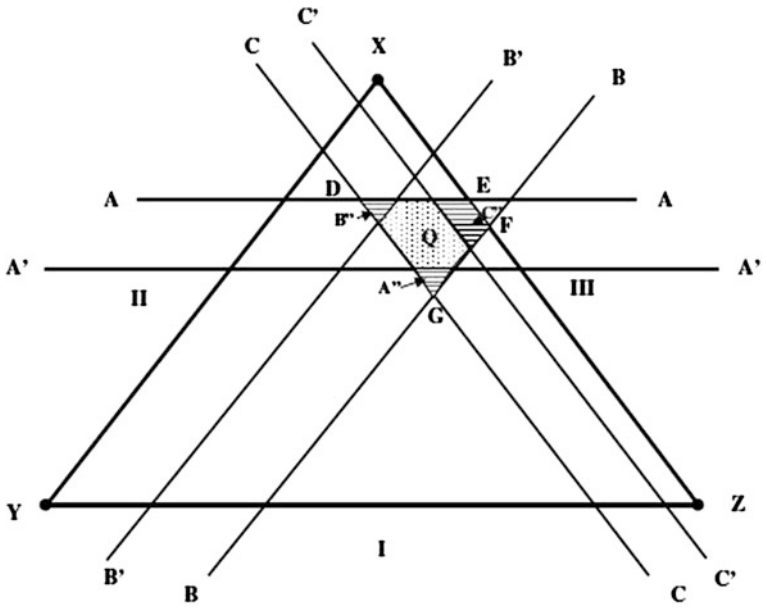


Fig. 22.1 CORE of the game in terms of characteristic functions

$c(1,2) \leq 111.0;\ c(1,3) \leq 89.1;\ c(2,3) \leq 54.0;\ c(1,2,3) \leq 118.1;\ c_i \geq 0 \forall i = 1,2,3.$

The values of the characteristic function can be plotted to show whether the core is empty or not. In Fig. 22.1, an equilateral triangle ΔXYZ is representing this 3-player cooperative game. The length of each arm is 118.1, which is the cost of the grand coalition and each point in or on the triangle is giving a solution to the cost sharing problem. All these solutions may not be feasible and the concept of core will help one to identify the feasible set, provided it is non-empty.

For the first player: stand-alone cost [c(1) = 81.7], represented by the straight line AA; incremental costs of coalition with the second and third players, one at a time, are [c(1,2) − c(2) = 111.0−46.1 = 64.9] and [c(1,3) − c(3) = 89.1 − 12.5 = 76.6] where both 64.9 and 76.6 < 81.7. So, it will be cost-saving option for the first player to form and stay in the two-person coalition. Similarly, if a grand coalition is formed by all three players then the incremental cost for the first player would be [c(1,2,3) − c(2,3) = 118.1 − 54.0 = 64.1] which is even smaller than what he could achieve through two-person coalition. So, the maximum cost saving through coalition that could be attained by player I is 64.1 and it is represented by the straight line A'A' in the diagram. The area in between the two lines AA and A'A' is giving the interest zone for player I. and as the magnitude of cost saving is increasing in the size of the coalition, player I has no incentive to defect.

Table 22.1 Comparison of alternative solutions

Method	Player 1	Player 2	Player 3	Total
Nucleolus	66.8	41.50	9.80	118.1
Shapley	72.18	36.84	9.08	118.1

Following the same logic for player II, the straight line BB represents his stand-alone cost c(2) = 46.1 and he will not be interested in any coalition where his share of cost will exceed BB. Incremental costs of coalition with the first and third players, one at a time, are [c(1,2)−c(1) = 111.0−81.7 = 29.3] and [c(2,3)−c(3) = 54.0−12.5 = 41.5] where both 29.3 and 41.5 < 46.1. Similarly, the incremental cost of the second player from joining a grand coalition would be [c(1,2,3)–c(1,3) = 118.1–89.1 = 29.0] which is even smaller than what could be achieved by joining 2-person coalition This maximum cost saving by player II is represented by B'B' in Fig. 22.1. For player III, the straight line CC represents her stand-alone cost c(3) = 12.5. Incremental costs of coalition with first and second players, one at a time, are 7.4 and 7.9, respectively, and that for the 3-person coalition is 7.1. All these values are less than the stand-alone cost and the maximum possible cost saving is represented by the line C'C'.

Thus, the bargaining zone where each player will have incentive to participate is identified first in terms of stand- alone constraints AA, BB, and CC as the area DEFG. When the incremental cost constraint of player I, viz., A'A' is superimposed triangle ΔA" got eliminated. Similarly, B'B' eliminated the triangle ΔB" and C'C' eliminated the shaded area C". Finally, the dotted area Q represents the core of the game.[4] Since the core is nonempty, two alternative solutions[5] to this cost sharing game have been reported in Table 22.1.

In the following section we will try to estimate the component-wise cost and total cost of running a CETP and the pumping stations in CLC and will try to find out the share of common costs to be borne by each pumping station by following both Nucleolus as well as Shapley principles. Since these allocations are efficient, it is expected that if introduced, they will create incentive for better in-house pollution management through effective and adequate pretreatment.

Efficient Cost Allocation for CETP in CLC

There are at present 288[6] tanneries operating at Calcutta Leather Complex spread over nine zones. There are six pumping stations. Each tannery is connected with some effluent pumping station (EPS) through HDPE pipelines of proper gradient

[4] As the number of players increases the core shrinks.

[5] For detailed derivation of the solutions see Appendix A.

[6] This information was made available by CLCTA and WBPCB officials in March 2012. Volume of effluent entering the CETP is 18.83MLD.

Table 22.2 Zone and EPS-wise distribution of effluent in CLC

Zone →	I	II	III	IV	V	VI	VII	VIII	IX	Total
No of Tanneries	48	62	23	17	48	7	26	32	25	288
Effluent Pumping Station										Volume of Effluent (KLD)
EPS-I	√	√	√		√					6108
EPS-II			√	√	√		√			5046
EPS-III		√				√				2822
EPS-IV								√		2394
EPS-V	√									2144
EPS-VI									√	319
Total										18833

Source WBPCB and CLCTA

whereby the pretreated effluent is supposed to reach the EPS without any external pressure. The effluents accumulated at the pumping stations are sent to CETP for further treatment. If from any zone effluent is sent to more than one EPS then a proportional allocation rule is applied to estimate the total volume of effluent at those EPSs.[7] The primary data on the design of the treatment plant, location map of zones and the position of the EPS, and operation and maintenance cost have been gathered from the West Bengal Pollution Board and the officials of the Calcutta Leather Tanners Association.

Table 22.2 shows the number of tanneries attached to each EPS and the volume of effluent sent to the CETP by each EPS. The first problem encountered in defining the cost function is related to the identification of independent actors in the system. We could have taken either the zones or the Effluent Pumping Stations (EPS) as the actors. If the former is taken, we have to deal with $(2^9-1) = 511$ possible coalitions and for the latter the number would be reduced to $(2^6-1) = 63$ possible coalitions. The second option is more manageable. Moreover, from individual tanneries the pretreated effluent reaches the EPS through underground gradient pipelines by gravitational forces and from there the pumping and treatment process starts. Hence, we have considered the six EPS as the players and the 288 operating tanneries have been grouped accordingly.

The Game Setup

Players	The EPSs are taken as the players in our cooperative game
Common knowledge	All the players have all the information regarding CETP
Negotiation set	The negotiation set is obtained by pre play agreements and it defines the minimum and maximum points (stand-alone and incremental costs) for all the players

[7] This apportionment rule has been suggested by the officials of the WBPCB.

Table 22.3 Cost of CLC project (Rupees in lakhs)

Sl. No.	Cost item	1997–1998 prices	2010–2011 prices
1	Land	1000.00	2020.00
2	Civil	277.00	559.54
3	Mechanical	3378.00	6823.56
4	Conveyance (ETS)	1592.00	3215.84
5	Pumping Station	430.00	868.60
6	Others (technical services)	541.00	1092.82
7	Chrome recovery plant	596.00	1203.92
8	Total cost	7814.00	15784.28

Source Dalmiya ML& Co.Ltd (1998)

Outcomes Determination of fair and equitable cost allocation to member units and the optimal size of coalition.

Cost Calculation

To estimate cost for each EPS, information is required on the following items:

- The capital cost of the grand coalition;
- The operation and maintenance cost of each coalition and the grand coalition.

Estimation of Capital Cost

Table 22.3 gives the cost of the CLC project. To estimate the current cost figures the available dated figures have been transformed into equivalent current price values.[8] The total cost (TC) of the CETP is given by the summation of capital cost (CC) and the operation and maintenance cost (O&M cost). CC has three components: (a) engineering cost like civil, mechanical, and the cost of other technical services, called the *capacity cost* (b) the *conveyance cost* (which depends on the length of pipeline of the effluent transmission system [ETS]) and (c) the *pumping cost* (which is dependent on the cost of pump station).

Capacity cost has been estimated by applying the following formulae $CC(a) = A*(\text{capacity})^{\alpha}$, where A is a constant and α is assumed to be equal to 0.75 (to keep in line with the falling fixed cost per unit of production.) [9] The capacity cost excludes the costs related to conveyance, and pumping. Here the capacity cost

[8] Using the GDP deflator parity has been brought between the two price levels.

[9] Anuradha (2005).

CC(a) is: [15784.28−(3215.84 + 868.6)] = 11699.84 and current capacity is 18833[10] KLD. Hence, from the cost equation for CC (a), A can be solved as 7.27. Given this value of A for each player, i.e., for each EPS, the CC (a) can be estimated corresponding to the volume of effluent pumped.

To estimate the cost of conveyance CC (b), information on the length of pipeline from each effluent pumping station (EPS) to the CETP is required. The total length of pipeline is 24.98 km[11] and the total cost of conveyance is Rs. 3215.84 lakhs, making the cost of conveyance per kilometer equal to Rs. 128.736 lakhs. This total length of pipeline is basically the summation of all pipelines connecting all tanneries to different EPS (18.0 km) and from six different EPSs to the CETP (6.98 km). So, for each EPS, the conveyance cost has been calculated as (specific distance * unit cost).

Finally, the third component of capital cost, viz., the pumping cost CC(c) has been obtained as [868.6/18833]*(volume of effluent per EPS).

Now, [CC (a) + CC (b) + CC(c)] gives the total capital cost for each EPS. To obtain the corresponding annualized value assumptions need to be made regarding the length of the project life and the social discount rate. Following the standard practice applied in case of VANITEC CETP of Tamil Nadu (Anuradha 2005), we have taken the span of project life as 20 years and the social discount rate as 15 %. Then the present value factor would be $PVF = \sum_{t=1}^{20} \left(\frac{1}{1.15}\right)^t = [6.2613]^{-1}$.

Multiplying the Total Capital Cost (TCC) for each coalition by this PVF the annualized capital cost is obtained.

Finally, the operation and maintenance cost is added to this TCC to get the total cost of treatment incurred by each coalition. According to CLCTA, the total O&M cost per month for 18833 KLD is Rs. 81.67 lakhs at CETP and another Rs. 25 lakhs (ETS + EPS) for all pumping stations taken together. So the total operation and maintenance cost is Rs. 106.67 lakhs per month. Hence the variable cost per KLD of effluent turned out to be Rs. 566.4. As the annualized average cost is an annual cost, we represent the operation maintenance cost on a yearly basis by multiplying it by 12.

Ultimately, the total capital cost and the operation and maintenance cost has been added to arrive at the total cost. Since the effluent entering the CETP is measured as Kiloliters per day the total cost (CC + O & M) has been divided by 365 to generate an estimate of daily cost. Tables 22.4 and 22.5 present the break ups of cost.

Table 22.5 gives the total cost of for the six players. With six players we will have [2^6 −1] = 63 number of possible coalitions (including the grand coalition). Given these cost information we want to explore the possibility of cost saving

[10] This information was obtained from WBPCB and CLCTA in 2010–2011.

[11] This information regarding the total length of pipeline and its distribution over different EPS has been provided by the CLCTA.

Table 22.4 Components of capital cost for each EPS (Rs. in lakhs)

Player	Vol. of effluent (KLD)	CC(a) A = 6.66, $\alpha = 0.75$	Length pipeline (km)	CC(b)	CC(c)	Total CC (a) + (b) + (c)	Annualized value; PVF = 6.2613^{-1}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EPS-1	6108	5028.20	0.862	110.971	281.708	5420.878	865.8166
EPS-2	5046	4357.11	1.335	171.863	232.727	4761.702	760.5337
EPS-3	2822	2817.77	0.925	119.081	130.154	3067.008	489.8591
EPS-4	2394	2490.76	1.73	222.714	110.414	2823.885	451.0278
EPS-5	2144	2293.01	0.255	32.8278	98.8838	2424.726	387.2745
EPS-6	319	549.327	1.875	241.381	14.7127	805.4211	128.641
Total	18833	17536.18	6.982	898.839	868.6	19303.62	3083.153

Source Calculated from data collected from WBPCB and CLCTA officials
Notes 1. CC(a) = A*[column(2)]$^{\alpha}$
2. CC(b) = 128. 736*[column (4)], where 128.736 lakhs is cost per kilometer of pipeline
3. CC(c) = 0.04612*[column (2)], where 0.04612 lakhs is the cost of EPS per KLD
4. PVF = $\sum_{t=1}^{20} \left(\frac{1}{1.15}\right)^{t} = [6.2613]^{-1}$

Table 22.5 Total cost per EPS (Rs. in lakhs)

Player	Volume	Capital Cost	O&M Cost	Total Cost	Daily Cost
EPS-1	6108	865.8166	415.1485	1280.965	3.5094
EPS-2	5046	760.5337	342.9665	1103.5	3.0233
EPS-3	2822	489.8591	191.8057	681.6648	1.8675
EPS-4	2394	451.0278	162.7154	613.7432	1.6815
EPS-5	2144	387.2745	145.7234	532.9979	1.4603
EPS-6	319	128.641	21.68179	150.3228	0.4118
Total	18833	3083.153	1280.041	4363.194	11.9539

Source Calculated from the data provided by CLCTA and WBPCB

through economically viable cooperation and for that we want to apply the Shapley Value criterion and Nucleolus criterion of cost sharing. Since for each player the stand-alone cost C_i is greater than the incremental cost of joining any coalition and the incremental cost is consistently falling with an increase in the size of the coalition, hence everyone has the interest to participate in the coalition and once joined, no one would have the incentive to defect. Hence the CORE will exist and the cooperation will be stable.

It is apparent from Table 22.6 that the present cost of running the CETP is Rs. 11.954 lakh and the least cost in core is Rs. 8.244 lakhs. We can now apply the Shapley Value allocation and Nucleolus method to find out the efficient cost allocation over the six players (Table 22.7). Both these allocations belong to core, both are individually rational, group rational, and Pareto efficient. However, the least cost combination in core (i.e., Rs. 8.244 lakhs) is not attainable. The maximum possible cost saving would be (Rs. 11.9540−9.4000) = Rs. 2.5540 lakhs per day,

Table 22.6 Coalition cost combinations and existence of core (Rs. in lakhs)

Players Coalition	Players						
	1	2	3	4	5	6	Game Value
$X_i = C_i$	3.509	3.023	1.868	1.681	1.460	0.412	11.954
$X_{ij} = \min\{C_{ij} - C_j\}\forall j$	2.859	2.373	1.360	1.211	1.489	0.257	9.549
$X_{ijk} = \min\{C_{ijk} - C_{jk}\}\forall j,k$ $j \neq k$	2.742	2.285	1.272	1.134	0.944	0.245	8.662
$X_{ijkl} = \min\{C_{ijkl} - C_{jkl}\}\forall j,k,l$ $j \neq k \neq l$	2.671	2.230	1.244	1.107	0.919	0.241	8.412
$X_{ijklm} = \min\{C_{ijklm} - C_{jklm}\}\forall j,k,l,m$ $j \neq k \neq l \neq m$	2.620	2.190	1.224	1.090	0.903	0.239	8.266
$X_{ijklmn} = \min\{C_{ijklmn} - C_{jklmn}\}\forall j,k,l,m,n$ $j \neq k \neq l \neq m \neq n$	2.613	2.184	1.221	1.088	0.901	0.237	8.244

Source Calculated from the data provided by CLCTA and WBPCB

Table 22.7 Alternative cost allocation options for CETP

Solution	EPS-1	EPS-2	EPS-3	EPS-4	EPS-5	EPS-6	Grand coalition
Nucleolus	2.8900	2.5620	1.4060	1.2200	0.9990	0.3240	9.4000
Shapley	2.9126	2.4615	1.4161	1.2642	1.0705	0.2751	9.4000
Current Allocation	3.5090	3.0230	1.8680	1.6810	1.4600	0.4120	11.9540

which means a 21.365 % cost saving. Comparison of Shapley allocation for each of the six players[12] indicates that players 1 and 2 (the biggest players) are subsidizing others.

At present, the rate charged per KLD is Rs. 14.50. Given the volume of effluent at 18.833 MLD, the total collection is Rs. 2.73 lakhs per day. This creates a deficit of Rs. 0.83 lakhs per day, even when one is inclined to cover only the operation and maintenance cost. However, the Shapley allocation would save Rs. 2.554 lakhs per day; and thereby not only the running cost but a part of the capital cost will also be covered, and eventually the CETP would come out as a self-financed project.

Concluding Observations

Thus we see that to be successful, a CETP institution needs to be organized by following the structure of a cooperative game where all the parties will have incentive to participate and no one will have the scope to violate the rules of participation. If such conditions can be created, the game is said to have a stable core solution. Here the rate structure should be so designed that it would be less

[12] See Appendix B for detailed cost allocation.

costly for each agent to participate in the coalition in the sense that the part of the shared cost borne by her would definitely be less than the stand-alone cost. Moreover, the cost share in this grand coalition would be less than her share of cost in every possible sub-coalition.

The pollution management issue is mostly approached in India on a case-by-case basis from the perspective of social cost benefit analysis. Here, technical feasibility enjoys greater weight compared to economic efficiency. However, this practice imposes huge fiscal burden on the regulatory authority and the success of the project becomes more contingent on the quality of monitoring, allowing incentive designing no significant role to play. Given the rapid expansion of environmental awareness, more proactive role of the civil society in protecting public interest, time has become ripe enough to incorporate pollution management in the greater agenda of fiscal reform.

A.1 Appendix A

1. The Nucleolus allocation for three players:
   Solving LPP for Nucleolus
   minε
   subject to $x_A \geq -\varepsilon,\ x_B \geq -\varepsilon, x_C \geq -\varepsilon$

$$x_A + x_B \geq 16.8 - \varepsilon, x_A + x_C \geq 5.1 - \varepsilon, x_B + x_C \geq 4.6 - \varepsilon,$$

$$x_A + x_C + x_B = 22.2$$

   The solution is $\varepsilon = -2.7,\ x_A = 14.9, x_B = 4.6,\ x_C = 2.7$
   The corresponding unique cost allocation, the Nucleolus allocation is $(y_A = 66.8,\ y_B = 41.5,\ y_C = 9.8)$;
   2. The Shapley allocation for three players:
   $X_i = \sum_{\substack{S \subset N \\ i \in N}} \frac{|S-i|!|N-S|!}{N!} C^i(S)$, where $i = 1, 2, \ldots$,n and $C^i(S) = C(S) - C(S - \{i\})$
   Using the above equation, the Shapley allocation for the game is calculated as follows

$$\mathrm{X_1} = (1/3)\,[\mathrm{C_1} - \mathrm{C_0}] + (1/6)[(\mathrm{C_{12}} - \mathrm{C_2}) + (\mathrm{C_{13}} - \mathrm{C_3})] + (1/3)\,)[\mathrm{C_{123}} - \mathrm{C_{23}}]$$
$$\mathrm{X_2} = (1/3)[\mathrm{C_2} - \mathrm{C_0}] + (1/6)[(\mathrm{C_{12}} - \mathrm{C_1}) +)[(\mathrm{C_{23}} - \mathrm{C_3})] + (1/3)\,)[\mathrm{C_{123}} - \mathrm{C_{13}}]$$
$$\mathrm{X_3} = (1/3)[\mathrm{C_3} - \mathrm{C_0}] + (1/6)[(\mathrm{C_{13}} - \mathrm{C_1}) +)[(\mathrm{C_{23}} - \mathrm{C_2})] + (1/3)\,)[\,\mathrm{C_{123}} - \mathrm{C_{12}}]$$

   Which gives

$$\mathbf{X_1} = 1/3(81.7) + 1/6(141.4) + 1/3(64.1) = 72.18$$
$$\mathbf{X_2} = 1/3(46.2) + 1/6(70.3) + 1/3(29) = 36.84$$
$$\mathbf{X_3} = 1/3(12.5) + 1/6(15.2) + 1/3(7.1) = 9.08$$

Also $\sum_i X_i = 118.1$

where X_1 = cost allocation of 1st player

X_2 = cost allocation of 2nd player

X_3 = cost allocation of 3rd player

The Shapley allocation satisfies the individual rationality, group rationality and Pareto optimality conditions.

A.2 Appendix B

Comparison of costs under different allocation schemes at CETP in CLC

Player	Volume of effluent (KLD)	Present cost (Rs. Lakhs)/day		Shapley cost allocation (Rs. Lakhs)/day		Cost saving (%)	Nucleolus cost allocation (Rs. Lakhs)/day		Cost saving (%)
		Total (Ci)	Unit	Total (Xi)	Unit		Total (Yi)	Unit	
EPS-1	6108	3.509	0.00057	2.913	0.00048	17.007	2.889	0.00047	17.671
EPS-2	5046	3.023	0.00060	2.462	0.00049	18.581	2.562	0.00051	15.266
EPS-3	2822	1.868	0.00066	1.416	0.00050	24.174	1.406	0.00050	24.713
EPS-4	2394	1.681	0.00070	1.264	0.00053	24.819	1.220	0.00051	27.448
EPS-5	2144	1.460	0.00068	1.070	0.00050	26.694	0.999	0.00047	31.606
EPS-6	319	0.412	0.00129	0.275	0.00086	33.199	0.324	0.00102	21.270
Total	18833	11.954	0.00063	9.400	0.00050	21.365	9.400	0.00050	21.365

Source Calculated from information given by WBPCB and CLCTA

References

Anuradha V (2005) Institutions of collective action for pollution abatement: the common effluent treatment plant. In: Mythili G and Hema R (eds) Topics in applied economics: tools issues and institutions: a festschrift volume for Prof. Sankar U, Academic Foundation New Delhi

Bagchi (Majumdar) S, Banerjee S (2012) Calcutta leather complex and tanning of leather in Kolkata: a blueprint. Artha Beekshan 21(1):68–81

Dalmiya ML & Co Ltd (1998) Calcutta leather complex project revised final version of DPR. (abridged)

Shanmugasundaram S, Murthy DVS (2000) Performance evaluation for common effluent treatment plants for tanneries at Pammal–Pallavaram Tamil Nadu (India). Bioprocess Eng 23:431–434

Shubik M (1984) Game theory in the social sciences: concepts and solutions. MIT Press, London

Thorstensen TC (1969) Practical leather technology. Reinhold Book Corporation, New York

Chapter 23
Embracing the Global Knowledge Economy: Challenges Facing Indian Higher Education

Saumen Chattopadhyay and Rabindranath N. Mukhopadhyay

Introduction

India's growth performance has been impressive by any international standard even if we ignore the recent downturn. However, this high growth has failed consistently to effect transformation in the socio-economic existence of the majority of the masses as indicated by the tardy improvement in the social indicators. As a result the incidence of poverty continues to be a major concern[1] and it is argued that the very character of the growth process has in fact led to a further accentuation in the inequality of distribution of income.[2] In terms of Human Development Index (HDI), India ranks 134th in the world today and in terms of per capita income, India's rank has improved by six notches to 128.[3] In order to strengthen the transmission process from income growth to socio-economic transformation, governance of public sector institutions and the delivery mechanism has to improve and people need to be empowered with good quality education and health to create and seize the growth opportunities and flourish. India, in terms of number, can take pride in having the third largest skilled manpower with the largest number of higher

[1] While nearly 77 % of the Indian population was categorized as poor and vulnerable (in 2003–2004) by the Sengupta Commission Report, India has witnessed a steady rise in the number of millionaires.

[2] Alternative Economic Survey, India 2005–2006, 2006–2007.

[3] The Economic Survey, Government of India (2010, 2012) attribute this inconsistency in ranking to rather poor performance in the realm of education rather naively. The rankings are so low, the difference becomes immaterial. The HDI ranking is reported in Global Human Development Report 2011.

S. Chattopadhyay (✉)
School of Social Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: schatto@gmail.com

R. N. Mukhopadhyay
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: mukhopadhyayr@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_23, © Springer India 2013

education institutions in the world.[4] However in reality, the growing number masks the ground reality in the education sector, the pathetic quality of education at all three levels of education associated with exclusiveness of the ongoing expansionary process. Since the time India embarked on economic reform in 1991, the demand for skilled labor has been going up as the growth is being propelled by capital intensive technology and information technology as knowledge assumes an increasingly larger role in a globalizing world. However, India continues to face shortages of skilled labor force of quality despite having a large stock of graduates. This has made nearly one-third to half of the labor force unemployable. Even for lower level of skill, the overall quality is poor.[5] The government has shown eagerness to raise the gross enrolment rate in higher education to meet the aspirations of the growing number of youths and cater to the growing demand for skill from industry and services by allocating a much larger budget for higher education in the Eleventh Five Year Plan.[6] At the same time, India cannot afford to ignore the emergence of the global knowledge economy. The compulsion is not merely to gain competitive advantage in the world economy in the realm of knowledge generation, but also to realize the huge potential of our human resources and reap the benefits of demographic dividends. The challenge is therefore a total overhauling of the education system to gear India up to embrace the global knowledge economy and at the same time, meet the national objectives.

This chapter seeks to address and assess whether India is prepared to participate in the global knowledge economy. The focus of the paper is therefore mainly on higher education, followed by a critical analysis of the policy initiatives of the government. We begin with an analysis of the emerging global knowledge economy in "The Emerging Global Knowledge Economy". We discuss the main features of the knowledge economy in order to deconstruct the knowledge economy. In "Imagining an Alternative to the Neo-liberal Vision of the Knowledge Economy: Open Science Ecology" and "How is India Placed in the Global Knowledge Market?", we discuss how well India is placed globally and the efficacy of the series of policy measures being implemented and mooted by the government to reform the higher education sector. In "Challenges Facing Higher Education",

[4] The number of higher education institutions would be more than 20,000. In terms of absolute enrolment, it is the third highest in the world though the enrolment ratio is pretty low at around 14 % (Agarwal 2009, p. 36).

[5] As per the 61st Round of the NSS, out of 260 million persons in 15–29 years age group, only about 30 million are trained in any formal or informal ways. The level of vocational skills of labour force was pretty low at 5 percent in comparison with 96 % in Korea and 60–80 % in the industrial countries (Agarwal 2009, p. 192).

[6] The gross enrolment rate was envisaged to rise to 15 % at the end of the Eleventh Five Year Plan. Although the budgetary allocation for higher education was raised by nearly eight times, the actual expenditure may have fallen short of the budgeted amount mainly due to the lack of absorptive capacity of the system and slow completion of the new institutions planned.

we discuss the challenges India faces in the area of professional education, in particular technical education in the context of the challenge of knowledge generation. In the last section, we discuss the role of education in building up an inclusive and cohesive society against the backdrop of India's participation in the global knowledge economy.

The Emerging Global Knowledge Economy

If the objective is to attain equitable growth which is sustainable in the long run, the question is what role education can play to realize this objective. To understand the role of education in the globalizing world in such a growth process keeping in mind the import of sustainability of such a growth path, we need to locate the role of education in such a process.

Education and Sustainable Development

The concept of sustainable development is a multidimensional concept which can be represented as a sustainability triangle with environment, economy, and society as the three vertices of a triangle (Ashford and Hall 2011, p. 9). The importance of the environment needs no explanation, as this has been the starting point of any discussion on sustainable development. The other two dimensions, economy and equity (or employment, the three E's) are interrelated and have linkages with the environment through various channels. Human capital, because of the growing importance of knowledge, plays an important role along with physical and natural capital in contributing to the competitiveness and productiveness of the economy. Through the formation of skills and the determination of wages, education and its externalities would again be a significant determinant of income distribution and social cohesion . Technological change in the context of globalization brings about changes both within as well as between the three dimensions of the sustainability triangle (*Ibid*. 236–237).[7]

Knowledge as a critical input has gained salience in spurring growth. The earlier formulation of the growth model that knowledge was freely available to all the nations has been found to be untenable.[8] The Endogenous Growth Theories or

[7] Education helps spreading awareness about the need to protect environment and can combat corrupt practices, which is the major source of degradation of nature. Environmental education makes individual understand the importance of the nature and the need to distribute natural resources in a sustainable way. Research and development can develop eco-friendly technologies, which can be useful in sustaining the environment.

New Growth Theory developed by Lucas (1988), Romer (1986, 1990) and others assigned a place of prominence to human capital and knowledge to explain growth differences among the nations. Broadly speaking, education and in particular, spill-over effects of knowledge and its nearly public good character has rendered the aggregate production function used for the economy as a whole increasing returns to scale. Romer's key intuition is that the most important precondition for growth lies in human capital dynamics and not in population dynamics. Focusing on knowledge and externalities by making investments in new research, education, and human capital rather than physical capital the country can achieve high growth rates. However, a developing country may remain handicapped because of poor infrastructure, poor quality of human capital, high drop-out rate *inter-alia* (Ashford and Hall 2011, p. 153). Many proxies for human capital have been conceived of starting from enrolment rate in secondary education to R&D expenditure to capture innovation. Further empirical studies indicate that it is not merely some proxy for education but the quality of education that matters. If ideas and innovation are the drivers of growth, then the underlying factors would be development in cognitive capacity which would have bearing on advances made in science and research at a later stage (Hanushek and Kimko 2000).[9] "The New Growth Theory is therefore an open acknowledgement of the role of education in the promotion of economic growth and development" (Comim 2007, pp. 90–91).

The global economy is witnessing rapid and irreversible changes. Facilitated by the expansion of communication technology and cheaper mode of transport, the global space is much larger now than what it used to be. Although there is heterogeneity in the structure as people from different nations participate in the process of globalization and join the global space, there are a great deal of commonalities among the participants in terms of activities and objectives. There are different layers, local, national, and global with different kinds of forces emerging from the spheres of economy, polity, and culture with all the spheres interacting with one another to delineate the emerging contour and the character of the global space. It is hierarchical and inequitable (Marginson et al. 2010). All the nations do not share the same platform in terms of their capabilities. All the citizens of a nation are also not on the same level playing field to participate because of disparities in income and endowments in terms of human capital like education and health. There are socio-economic forces which are pushing some people to the periphery to fend for themselves and there are people who are

[8] The model developed by Solow (1956) assumed constant returns to scale production function which led to diminishing marginal productivity of capital. The nations would therefore reach steady states and technology was assumed to be available to all.

[9] Hanushek and Kimko (2000) conclude that quality of labour force could be measured by comparative tests of mathematics and scientific skills which 'has a consistent, stable and strong relationship with economic growth'.

capable, and are coming forward to participate in the growth process to seize the opportunities. Knowledge is playing a key role in uniting the different forces emanating from different regions. Knowledge generation and dissemination are now truly global. Massification of higher education and the emergence of the knowledge economy have been driving the steady rise in the research output. Professionalisation of the academic profession in view of an increasing focus on quantification of research output in charting an academic career coupled with external incentives to publish has led to this growth (Vincent-Lancrin 2009, pp. 149–150). Growth in the knowledge economy relies primarily on innovation and R&D but universities have a role to play in basic research. This has got to do with the importance and character of knowledge and the knowledge economy. There has been a rise in the dominance of knowledge as the major input in propelling growth in developed countries. The undergoing transformation of the knowledge economy is crucial to understand the neoliberal globalization and its impact on higher education policy (Olssen and Peters 2005).

Furthermore, transmission of knowledge across the borders is virtually a costless process because of the development in communicative technology. It is also a typical commodity, the production of which can take place at different places at the same time through the Internet. The open source software is an example. However, characterizing globalization merely by growing mobility of people, trade flows, communication, ideas and knowledge technologies, and capitals is tantamount to repetition of a truism (Marginson 2010, p. 121). In the realm of communication, information, and finance, there has emerged a one-world system working in real time in communication. Development in higher education is crucial for both knowledge generation and knowledge dissemination. Higher education institutions produce not only trained graduates but also space for undertaking fundamental research. In view of the linkages higher education has with the economy and society, it is beyond dispute that reform of the higher education sector is the need of the hour. However, the question is what kind of reform given the socio–economic situation that prevails, would alleviate the problems that we encounter in higher education and help India to compete in the global economy? International competition and international ranking have defined a new context for countries and institutions. International collaboration has grown as is evident from the growth of internationally co-authored and collaborative science articles. For India during 1996–2003, the number of collaborating countries has risen from 82 to 107 and the share of international collaborations in national output has grown from 10 to 16 to 22 in the years 1988, 1996, and 2005, respectively (Vincent-Lancrin 2009, p. 162).

Imagining the World of the Neo-liberals

The neo-liberals[10] intend to create a quasi-market in higher education (HE) both at the national as well as the global levels. Neo-liberalism promotes choice making

individualism but not the building up of capacities to exercise choice, and so the focus is more on negative freedom (overcoming constraints and encouraging anti-statism) and not on positive freedom (creating new 'imaginings'). Universal commodification of knowledge has also failed because the global knowledge economy is being driven no less by cooperation (Marginson 2010).[11] The intrusion of neo-liberal ideology in mainstream citizen education has been rather limited (ibid.). Globalization is characterized by both competitive and cooperative elements. While subjecting education to a competitive global market is contestable, there are possibilities of definite gains from collaboration in research.

Hierarchy in Space Making

The desire for global synchrony now touches many aspects of higher education initiatives. The Bologna and Tuning projects in Europe are processes of voluntary exchange toward global convergence. There is an attempt to reform higher education institutions (HEIs) and bring them closer to university rankings to become global research universities. On the other hand, national systems want to synchronize effectively with one another as individual institutions do. The process of globalization is marked by policy borrowing and policy convergence. Spontaneous synchronies of individual scholars with other researchers and governments, and conformity to global norms are being driven by university rankings. The application of new public management (NPM) to reform HEIs in line with corporate managerial practices is a striking example of policy synchrony (Marginson 2010, p. 140).

There has been rapid growth in global partnerships among universities as evident from the growth of signed MOUs. Although many MOUs remain inactive, the universities take the opportunity to show their commitment to internationalization and global modernizations. These partnerships between broadly similar institutions in terms of status and functions are common which incorporate subsidized collaborative research and exchange of student and faculty. However, there are partnerships between institutions of unequals with 'twining programs' and franchising of foreign university mostly from the Anglo-American world being on the main agenda with local providers from the Asian nations and Mexico where the teaching–learning process is delivered.[12] There are some universities in the world which are heavily networked.

Although the domain of HE and research is now truly global which includes institutions with cross-border activities, this very process excludes other

[10] The neoliberals advocate a new mode of regulation and a new form of governance. The central presuppositions are self interested individuals, free market economics, a commitment of laissez-faire, and a commitment to free trade (Olssen and Peters 2005).

[11] Neo-liberal policy prescriptions have not been implemented in full even in the UK, New Zealand and Australia. Competition has not led to full cost recovery (e.g., American Ivy League) even in some mass vocational education institutions and also in the University of Phoenix.

institutions.[13] At the same time, there are institutions which are best described as diploma mills and corporate-type universities. The rapid increase in the global flows is dynamic and uneven. When networks expand, the cost rises in a linear fashion but the benefits grow at an exponential rate owing to the increasing number of connections. The cost of remaining excluded from the network or penalty rises concomitantly. The flows of the networked remain confined to the nodal cities of North America, UK, Europe, and East Asia while a large part remains outside of it.

Higher Education Under the Regime of WTO-GATS

Negotiations with respect to higher education held under WTO-GATS framework are designed to create an open global trading regime. It seeks to treat education as a tradable commodity (Bergan et al. 2009, Tilak 2010), encourage enterprise and trade, including the entry of foreign providers.[14] At the same time it closes the possibility of production of education as a national public good . The progress has not been satisfactory as nations grapple with the envisaged role of education under the regime of GATS and with their national objectives despite there being a provision of special treatment. Conjuring up a competitive scenario where local institutions are not well equipped to compete is often an uncomfortable scenario for many developing nations.

Imagining an Alternative to the Neo-liberal Vision of the Knowledge Economy: Open Science Ecology

With the unmistakable tendency to transform every sphere of the economy into a market economy and the overwhelming dominance of the neo-liberal approach, it may seem that global knowledge economy characterized by competition and unleashing of market forces at the global level is accentuating global inequalities. While it is true that the global sphere is more competitive than before, it is also marked by burgeoning cases of cooperation at the global level among HEIs, among the individuals to create knowledge, facilitated by rapid advances in

[12] It is expected that a university of repute from the Anglo-American world would partner with a HEI in a developing world for the purpose of offering twining programme or under the franchisee of the foreign university. The craze for a degree from a foreign university in the developing world, like India can be so high that the university from the developed world need not be a famous one as many such collaborations in India would testify.

[13] Tilak (2010) argues that this ranking leads to hierarchy as the universities with high ranking attract all the attention and others remain sidelined.

[14] There are four aspects of global trade in education: cross border supply, consumption abroad, commercial presence and movement of natural persons.

communication technologies. Universities and research institutions worldwide are collaborating with each other. Students and the faculty criss-cross the borders and collectively participate in the creation of knowledge.

The global knowledge economy is emerging as an alternative to the neo-liberal version of globalization (Marginson 2010). Otherwise, it would be difficult to comprehend the social networking and open source of knowledge or open science ecology. Knowledge has emerged as a global public good as knowledge is not only non-rivalrous, barely excludable (Peters 2010), but is not transparent as knowledge requires some experience of it before one discovers whether it is worthwhile and relevant. Nations participate to maximize gains from global public goods through cooperation in various areas like climate, health, education.

Openness and international collaborations have emerged as alternative modes of social production which defies the law of scarcity. This is essentially a non-market, non-proprietary method of cultural exchange, a new information environment, and a networked economy which threatens traditional models based on economic logic of exchange as it rests on an ethic of participation and collaboration based on the co-production and co-design of capital goods and services. This has the potential to assume a much larger role alongside the traditional mode based on economic logic (Benkler 2006).

This is not to say that economic factors are becoming less important. Creativity has economic roots. It has economic manifestations and outcomes. But new technologies are opening up new vistas for creation, new imaginings, more powerful search engines, new ways of synthesis as it has increased our reflexivity (Marginson 2010). Hence, the vectors of space and time are dramatically altered by new 'communicative ecology and travel'. It is not true that universities, even those largely privately funded are entirely being guided by profit motives. Many offshore research ventures collaboration do not contribute to profit making. The changes in spatial material conditions in terms of capacity to communicate and synchronize across the borders have created new potentials for 'imaginings' and creation of special forays.

In the context of this development, building up capacity for knowledge production, a new set of techniques have to be developed which will be in conformity with communicative globalization (Marginson 2010). Investment in public communication and knowledge production is crucial. It has to be considered keeping in mind that there exists diversity in knowledge goods and it is predominantly of a public good character, to encourage local creative enterprise and various traditions of the knowledge production all over the world have to be factored in. Since knowledge is produced also in open source ecology other than the sphere of formal scientific institutions, the Internet can play a major role in connecting the local researcher with their counterparts. The expansion of OSS holds promise for developing nations as market is bypassed in the production of knowledge.

Is Knowledge a Private Good?

As knowledge converges toward the center of policy focus for developed as well as developing nations, the question is what type of commodity is knowledge? Should we allow the market to determine knowledge production or does the public sector have a role to play? A more fundamental question is whether knowledge is a private good or a public good? Stiglitz argued that knowledge is more like a public good if we go by the definition of non-excludability and non-rivalry in consumption (Olssen 2005). If that is so, it would not be proper to leave the production of knowledge to the private sector as the market would result in its under-provision. This argument is based on the assumption that marginal cost for knowledge consumption is zero and hence it should be freely available to all so as to honour the efficiency conditions. Any restriction on its use is therefore inefficient. The production cost of knowledge, however, is not zero. This has led to the emergence of IPR/patents to incentivize production of knowledge and make sure that the production is sustained with cost recovery without facing resource constraints. In practice, knowledge cannot be classified as either a public good or a private good. It depends on policies to shift the balance in favor of either, by increasing its 'publicness' or 'privateness' (Marginson 2007). In the context of global economy, the concept is now one of global public good (Stiglitz 1999).

Outsourcing: The Phenomenon of BPO and KPO

If labor were cheap in a country compared to the rest of world, in a trade theoretic model, either labor would flow across the border to get higher wages or capital would flow across the border to situate production process where labor is cheap. With capital, i.e., high end technology now moveable across the border with great velocity through Internet, the need for labor to cross the border has diminished. Capital instead reaches cheap labor in developing nations and reaps the benefits of cost cutting. The growth in the urban middle class with a command over the English language has now joined the outsourcing industry and they constitute the source of cheap labor. The global economic crisis albeit may put a restriction on it; in India, even the rural space is steadily getting integrated into the process. The new generation now demands better quality education and command over the English language which has led to a rise in the demand for provision of education by the private sector. At the same time, the phenomenon of drop-outs in government schools has shown little improvement. There has not been any

significant development in the learning outcome as several national and international surveys indicate.[15] The extent of diversity in the classroom has gone down with rise in the number of private schools which offer English education at a price not easily affordable by the majority in the society. This has dented the process of social cohesion as the school system tends to reproduce inequality in the society.

How is India Placed in the Global Knowledge Market?

The question is whether India is well equipped to participate in the global knowledge economy. In a way, India is already a part of the global knowledge economy if we go by the number of students who leave the country to study abroad, the MOUs signed, the contribution of non-resident Indians in knowledge production abroad, and engagement of universities and research institutes in global collaborations. However, in the realm of innovation, R&D, and the reputation of research institutions and universities, there is not much that India can take pride in even in the context of Asia–Pacific nations.[16]

One way to build-up research capacity is to increase global connectedness. Research universities are generally more networked than other HEIs. The growth in such collaborations, student, and faculty exchanges has risen in recent years. Meaningful, effective, and long-term participation in the global knowledge economy entails building up the research capacity. Not only will this make India dependent on knowledge flows, it will not be able to contribute to the development of global science and culture. R&D spending in India was 1.03 % in 2006 (Agarwal 2009, p. 252) compared to 3.44 for Japan, 1.49 for China, 3.47 for South Korea, 2.01 for Australia, and, 1.2 for New Zealand in terms of GDP. Research in universities in Asia is lower than in USA (13.3 % in USA compared to 8.3 % in China). China has a high proportion of university research funded by business. It is beyond doubt that India does not have a promising research university sector. Almost 80 % of all R&D is in the public sector with a high share for government research agencies. The share of the university is minor and basic research is only 17.8 %(Agarwal 2009, p. 253).

Rao (2012) assessed India's competence in scientific research very positively. He said that around 10–15 years ago, Indian universities contributed to around 50 % of the total scientific research in the country and for Chemistry, the Indian

[15] Around 50 % drop out by Class VIII. Only around 55 % of those who complete successfully Class XII, pursue higher studies. The ASER report, PISA score which measures the cognitive capacity in Mathematics and science mainly show that learning outcome is woefully poor. As mentioned earlier, this has the potential to affect growth of a country by affecting skill acquisition and innovativeness in the working population.

[16] Not a single Indian university has been included in the top 100 universities across the world. In comparison, universities from China are doing particularly well in view of their late entry into the global high ranking higher education institutions.

Table 23.1 Selected demographic, economic, communicative, education, and research indicators: Select Asia–Pacific nations

Nation	Population 2008	GNI per head PPP 2008	Internet users per 100 persons 2007	Adult literacy (over 15 years and over) 2007	Public spending as % of GDP 2008	GER in HE 2007	Scientific papers 2007 (Average annual change 1995–2007% in science and engineering papers)
	Millions	$US		%	%	%	
China	1325.6	6020	22	93.3	3.9	22	56,806 (16.5)
South Korea	48.6	28,120	77***		3.1**	96	18,467 (14.1)
Japan	127.7	35,220	69		3.8**	58	52,896 (1.0)
Thailand	67.4	5990	20	94.1	4.0	…	1728 (14.5)
Singapore	4.8	47,940	68	94.4	3.1**	…	3792 (10.5)
India	**1140.0**	**2960**	**7**	**66.0#**	3.0	**13**	**18,194 (5.7)**
Pakistan	166.0	2700	11***	54.9		5	741
Bangladesh	160.0	1440	0***	53.5	1.9	7	235
Sri Lanka	20.2	4460	6***	91.5	2.3		124
Australia	21.3	34,040	56***		4.8 (2005)	75	17,831 (2.6)
New Zealand	4.3	25,090	69		5.1 (2004)	79	3173 (2.2)

Notes *** data for 2008. ** data for 2007. # The adult literary rate for India would be much higher than reported here as the rate shown here pertains to 2001 Census

Source Marginson (2010), The Asia–Pacific region of Higher Education: National Funding and Global Influence, paper presented at U21 Symposium at Delhi

Kapur and Mehta (2012) argue that there has been considerable improvement in Indian research in science and technology. The number of papers filed from India increased from 18,911 to 36,261 during 2002–2008 compared to that of China's 38,206 to 104,968 (UNESCO 2010). Role of Indian universities has been modest. Kapur and Mehta argue strongly for a regulatory revolution. "India's higher education system remains neither fish nor fowl—suspended between over-regulation by the state, on the one hand, and a discretionary privatization that is unable to mobilize private capital in productive ways, on the other". (Kapur and Mehta 2012, p. 334). The success of a few professional institutions masks a deeper crisis

contribution to world research was around 8–9 %. The total number of research papers was between 10–14 thousand of variable quality. India is recognized in the IT and biotechnology fields today. Indians have achieved commendable success in the USA. 'the investment of industry in R&D till recently was marginal and almost all the research was government funded.' (Rao 2012, p. 612). The CSIR laboratories have made contributions which are at best considered as incremental improvements in known technologies. The scientific scenario has undergone a sea change. Quality of science is also very high. India now produces 12,000 research papers and 5,000 Ph.D.s in science every year.

Select Asia–Pacific nations are doing well by international standards. Table 23.1 shows how India is placed in the region. India lags behind developed economies of the region having $25,000, or more per capita income like Japan,

South Korea, Taiwan, Singapore, Hong Kong China, Australia, and New Zealand. As Marginson (2010) rightly remarked, though India talks about knowledge economy, it underperforms relative to its reputation as an emerging economic giant. China, though a middle emerging economy, is ahead of India in almost all respects.

India is yet to develop a coherent national strategy having weak national level coordination. Both India and China have gross disparities within the country. India has 32 % population in the age group of 0–14 years. The demand for HE is set to rise, and is supposed to provide a larger pool of productive labor and a larger capacity to generate national wealth.

The quality of education in India, both with regard to school level and higher, by any measure is poor as indicated by various studies. The quality of education imparted in private schools is hardly better. With regard to higher education, the employability of graduates is low as claimed by some studies.[17] It is an irony that Indian higher education institutions do not even feature in the ranking of the top 100 in the world despite the fact that IITs and IIMs are generally regarded as institutions par excellence[18] and with India having the third or fourth largest skilled manpower in the world and the largest number of institutions. With very few universities of excellence, the research base at the universities is rather limited. State universities somehow maintain the day-to-day functioning with state governments refusing to raise the budgetary allocations for state funded universities. The quality of the infrastructure for supporting research activity is somewhat better in the central universities, IITs, IIMs, and some select government funded institutions. In the aggregate sense, in a country with 400 plus universities, the production of knowledge is utterly inadequate by any international standard.

For meaningful integration with the emerging knowledge economy, the higher education sector is confronted with many challenges. The government has of late emphasized the salience of knowledge and the development of human resources and initiated many policy measures to overhaul the system. The question is how good are the policy measures in the context of the prevailing system, the issue of resistance, and proper implementation of programs. This problem is aggravated by stagnancy in the budgetary allocations by governments. The existing ones are suffering from poor governance.[19]

The issue is why even private sector institutions are failing to impart good quality education when the advocates of the market and the private sector

[17] As reported by the EXIM Bank of India (2007, p. 21), a survey conducted by McKinsey Global Institute found that only 25 % of the engineers, 15 % of the finance professionals and 10 % of the generalists are considered to be employable at the pass-out stage by the employers.

[18] The IITs and the IIMs are not full fledged universities and therefore in terms of the criteria for ranking, the IITs and the IIMs fail to qualify. Marginson (2010) that India has only one institution in the top 100.

[19] Generally it is argued that because of the principal-agent problem, the quality of service delivery in the government sector is poor as self-interest dominates the conduct of the employees. The neo-liberals advocate for a public choice approach to reform the governance structure.

dominate policy making; privatization, however, has helped in expansion of technical education in a major way.

Absence of a Well-Defined Production Function for HEIs

Universities are embedded in social and cultural contexts and missions/mandate are determined by local/national needs. This is true for government aided institutions. For the majority of privately funded institutions, the goal is profit maximization although the Constitution does not allow it. As argued by Chattopadhyay (2009), in the absence of a well-defined production function for an academic institution, it has been profitable for privately funded institutions to abuse the input–output relationship to the detriment of quality education.[20] Since quality of education cannot be measured precisely and craze is for the certificate and not skill, it is both profitable and possible for the private sector to compromise with the quality and quantity of input such as infrastructure and teachers.

Failure of Education as a Screening Device

In the theory of screening, it is argued that certificates act more as signaling devices meant for the use of screening. This is based on the fact that students strive to get good grades and teachers/owners are sincere. In the present milieu that prevails in the majority of the private HE sector, certificates have ceased to be screening devices as degrees are up for sale. Students have become clients who are to be appeased by the institutes. In other words, the certificates and grades have ceased to reveal the true quality of the students/job applicants in the market resulting in serious mismatch between demand and supply of specific skill.

Any nation would want a closer correspondence between HE, work, and social opportunity. In India, the rate of unemployment is higher among people with more than 12 years of education than in any other group. The social demand for university places may exceed the labor market demand for skills. The functioning of the labor market does not ensure that supply of skill matches demand for skill. Degrees may be treated more as generic rather than occupationally specific. A high proportion of graduates work in areas other than those they were trained for. The

[20] Profit making in social sector like education is disallowed by the Constitution. However, making reasonable amount of profit can be made and a part of the profit can be re-invested. Through artificial escalation of costs, profit can be siphoned out and that is how the majority of the privately funded institutions are sustaining their business. This also provides enough temptations for the private operators to invest in professional institutions. In fact the growth, or the best way to describe would proliferating of the private sector institutions has been very rapid to the tune of 18–20 % per annum.

graduate unemployed do better over their lifetime and respond better to new job opportunities besides the externalities they generate for building up a better society.

Globalization has ushered in sea changes in the global dynamics of knowledge production and dissemination. India has to position herself in the global knowledge economy and consolidate her competitive strength. Possibly the most meaningful way of achieving this is to focus on higher education through overhauling of HEIs, strategizing to link up with the rest of the world. To what extent our education sector is contributing to the growth of the knowledge economy in India and whether India can participate in the global knowledge economy to reap the benefits to further accelerate growth and distribute the fruits of growth in an equitable way is therefore a pertinent question. Now the issue we need to confront is whether India in the context of globalizing market will always be in a disadvantageous position because of the inherent north–south disparity and a large but substandard education system. The best brains will always veer toward developed countries as migration rises. The developed countries' demand for high-skilled workers will always be present. Their demand for students will also go up because of demographic transitions. Nowadays research does not take place in isolation. Collaboration opens up opportunities and gain from each others' strength.

Challenges Facing Higher Education

If India is to participate in the global knowledge economy today, the higher education system needs to be reformed in a major way. Some of the policy initiatives being mooted by the government are discussed in order to look for some practical solutions.

Recent Policy Initiatives in Higher Education

There have been a series of policy initiatives by the government in recent years starting with the setting up of the National Knowledge Commission (NKC) which marked a new beginning in higher education and in particular knowledge generation, which got its due role in national policy-making culminating in five Bills pending for approval in the Parliament. The overall thrust of NKC was expansion of the HE sector based on public–private partnership (PPP) mode, institutional reforms, and setting up an Independent Regulatory Authority for Higher Education (IRAHE) to essentially regulate and monitor the Indian higher education market.[21]

[21] It is being felt that there are many authorities with overlapping jurisdictions to regulate the chaotic situation which prevails in the Indian higher education sector comprising a growing number of private institutions who defy regulations recklessly.

Unprecedented rise in the budgetary allocation for HE during the Eleventh FYP had a renewed thrust of inclusive expansion of the HE sector along with the focus on achieving excellence. In view of NKC's thrust on expansion, but to make expansion inclusive which could not be achieved by the growing participation of the private sector alone, the Eleventh FYP envisaged a government-led expansion of the HE sector. Since the HE sector was becoming more like an unregulated market with very little power at the disposal of the government, the Yash Pal Committee was set up to examine the proposal of setting up of IRAHE.[22] Another committee was constituted under the direct guidance of PMO to look into the problem of skill development through emphasis on vocational training and setting up of ITIs. Two of the five Bills are on Foreign University Bill and setting up of National Commission of Higher Education and Research . The question is to what extent the policy initiatives serve the purpose of integrating the Indian economy with the global economy, not to forget the paramount importance of achieving inclusive growth.

Inclusiveness and Expansion: The Issue of Fiscal Sustainability

Two of the three major objectives as envisaged in the Eleventh FYP are expansion and inclusion with the overriding objective of attaining excellence. While the HE sector needs to be expanded to accommodate the growing demand for higher education, the expansion needs to be inclusive in view of the low level of income of the majority of the masses, and regional and gender disparities. Banking on private sector for expansion would inevitably make it less inclusive as cost of education would become higher.[23] The fiscal capacity of the state has become weaker despite an apparent improvement in the fiscal health of the states[24] in India. The state still has to grapple with the implementation of the Right to Education (RTE) Act in terms of provisioning of adequate financial resources to the states and overcoming the loopholes in the Act which are now being debated to be utterly inadequate in view of the need. Successive governments have failed to raise the expenditure on education as a whole to 6 % of GDP. Expansion of higher education to be made inclusive entails a significant rise in government support. The government, center, and the states together spend only around 0.8 % of GDP

[22] Chairperson, Prof Yash Pal expanded the mandate of the Report and changed the focus to rejuvenation of the HE sector as he wanted to have a holistic view of the prevailing situation.

[23] The provision of education loans is a poor substitute for scholarship and subsidization. The capital market for education loans is inherently imperfect and it is more so in India (Chattopadhyay 2007).

[24] As per the observations made by the Thirteenth Finance Commission, the states are now having surplus in their revenue accounts with surplus cash being held. The need to comply with the FRBM Act has possibly led to such a situation.

on higher education which is just half of what the NKC recommended.[25] It is argued that adherence to the FRBM Act should get priority and proper implementation of RTE would put the fiscal situation under further strain. Hence the government looks for solutions in the form of PPPs. Reprioritization of education values in the context of globalization should not mean deviation from the objective of equitable growth. Private sector participation in higher education cannot make expansion inclusive whether pure private sector participation or in the form of PPP. In fact, while NKC argues that there exist trade-offs between two of the any three objectives of expansion, inclusion, and excellence, the UGC argues that there need not be any such trade-offs if the government evinces interest in spending the required amount of resources.

Assessing the Policy Options

Research takes place in specialized research institutes as well as in universities. Although there has been an increase in public funding for research institutes, the institutes suffer from paucity of resources and leadership problems. On the other hand, universities are being compelled to become self-financing as they offer market-oriented courses. There is a tendency to ignore fundamental research as return is delayed. Tying up with industry makes research more market-oriented as funding determines research agenda. Universities fail to draw enough talent despite an increase in pay package. There should not be any segregation between a teaching and a research university. Indian universities are far from being able to compete in the international arena as the international ranking testifies.

New Public Management and Governance Issues

There is a convergence in the adoption of NPM to reform the functioning of HEIs in the major parts of the world. The adoption of NPM assumes that corporate managerial practices can be applied to HEIs as schools, colleges, and universities are treated no differently from factories. It is argued that service delivery in the public sector is inherently inefficient as public sector employees neither face an incentivized pay structure nor do they operate in a competitive ambience. Moreover, the public sector suffers from the principal-agent problem (Olssen and Peters 2005).[26]

[25] NKC recommended 3 % for primary education, 1.5 % for school education, 1.5 % for higher and technical education.

[26] NPM derives its rational from Buchanan's public choice approach to the government. Given the self-interest employees and the hierarchical structure of a public sector organization, the principal agent problem becomes inevitable. Some high ranking universities are motivated by the

One way to ensure that public sector employees deliver would be to incentivize their pay structure, measure their output and performance. The approach identifies goals, policies, and strategies to achieve them. NPM emphasizes surveillance, transparency of outcomes, and performance management. HEIs are managed by professional staff like executive managers and entrepreneurial staff. They seek to foster client style relation between the producers (i.e., the teachers) and consumers (i.e., the students). Institutions are managed by a regulatory body through competition, and funding is based on incentivized formula-based agreements.

Although it seems that NPM may achieve efficiency in delivery, it has many adverse implications. It deprofessionalizes the teaching profession as the autonomy of the teachers is seriously compromised (Olssen et al. 2004). It commodifies education, which is treated like any other commodity. It makes the teacher a worker and a student as an investor. The rationale behind the application of NPM implies that HEIs are similar to factories.

Limits of Market Reform

Application of market principle to reform education sector occurs at two different levels. One is at the institutional level with the application of NPM and the other is at the macro level. Neo-liberalism advocates choice-based individualism but not on how to develop capacity to make the choice. Basically they are opposed to state intervention.

Application of market logic to reform the market is fraught with dangers. The market for education which can be constructed can be construed as only a quasi-market. The market would remain essentially an imperfect competitive market with high degree of heterogeneity in the quality of education provisions and limited mobility of resources. What is being attempted is to provide sovereignty to both consumers and producers. As argued by Chattopadhyay (2009), sovereignty of consumers and producers would actually lead to serious compromise with the values that education embodies. Education becomes a commodity, if prices are determined in the market, if information is provided to students to make informed choices, and seats in prestigious institutions are offered for sale. Competition does not lead to efficiency and improvement in quality is not often the outcome. The result of the competition is consolidation of hierarchy. The argument in favor of relying more on market forces also implies that market demand is a proxy for social demand, which is fallacious. To enhance access to education to make the growth process more inclusive, the need is to cater to social demand. The objective should be to expand the 'publicness' of higher education. However, marketization

(Footnote 26 continued)
'prestige maximization' objective which enable the faculty to rise above their narrow self-interest and deliver.

renders knowledge more as a private good as IPR restricts its use. Further, hierarchy in the ranking of institutions assigns social positional values to render higher education a 'positional good'. Degrees from high ranking institutions which make degrees a positional good in the society is actually a zero-sum game (Marginson 2004) where those from the high ranking institutions gain at the expense of others who are from low ranking institutions. The phenomenon of growing share of open source software (OSS) has remained largely unexplained. In fact, synchronicity in OSS is more than the kind of cooperation that usually prevails in HE (Marginson 2010). Further, as argued by Winston (1999), competition in education market leads to consolidation of the hierarchy as the best institutions attract the best minds and funding and the low ranking perforce become fifth generation universities which are no different from a typical corporate entity (Tilak 2010).

The PPP is widely believed to be the panacea for both the resource constraint faced by the government and for ensuring provision of quality education as the private sector is more efficient. There are two questions. One, if the private sector is not allowed to make profit why should they be interested in investing? If the private operator is allowed to recover at least a part of the cost from the students, access would suffer. Otherwise, the financial burden of the state would not necessarily fall. Will an increase in efficiency in resource use improve the quality of education?

As argued by Chattopadhyay (2010) the entry of foreign universities is unlikely to contribute to all the three objectives of expansion, inclusion, and excellence. The expensive foreign universities will only cater to the upper crust of the society and fail to create the academic ambience and quality of education obtained in Harvard or Yale or Cambridge University.

Industry–University Linkage

The mandate of the university has now gone beyond mere teaching and research. The university is being aligned with the industry to legitimate its contribution to society in a more direct way under a new social contract. The government has become proactive to ensure realization of the full potential of research capabilities of the university by giving incentives and thereby encouraging academic institutions to contribute to wealth creation also. In earlier times, only long-term contributions of academic knowledge were considered in a linear model. Now, even short-term contributions are recognized to capture the multiple reciprocal linkages at different stages of capitalization of knowledge (Etzkowitz and Leydesdorf 1997). Shortage of adequate funds has also become a compelling factor behind this new collaboration between universities and the industry. Technology has become complex and multidimensional and therefore, it requires more than one scientific discipline which a university can cater to. The interface in the short run is sought

to be captured in a triple helix model comprising the academia, industry, and the government. This brings about a change in the mission of universities from traditional intellectual philanthropy to intellectual property.[27] As Tilak (2010) argues, universities become interested in immediate return research and client centered research and they imbibe the corporate culture of hierarchy and control, efficiency and standardization, and obedience which are "anathema to the creation of new knowledge, therefore, to the core mission of the university" (*ibid*. 115).

Technical Education: Growing Mismatch

The demand for skilled workers has gone up, pulling up their pay packages along with it. HEIs, particularly under the ownership of the private sector exhibit increasing preference for offering market-oriented professional courses. Course fees rise with the increasing role of the private sector. However, higher studies in technical education fail to attract talents as they are not lucrative enough for research in engineering and for entering the job market as a faculty in engineering. The IITs are already facing an acute shortage of faculty and the dearth of teachers in technical education will tend to accentuate further. Vacancies keep mounting. As a result, research in technical universities suffers. The teaching load of the existing faculty has become heavier. At the same time, the faculty exhibits keenness in consultancy, which is generally market-oriented research funded by the industry to meet their short-run problems. Funding for fundamental research remains insufficient. Although there has been a rise in the supply of institutions, the quantity of quality-skilled professional has risen only marginally. The trade-off between quality and quantity has become acute. The tremendous growth in private professional institutes has not been able to create a larger research base because of several factors. One, there is a growing craze for credentials but not for competence as certificates become more important rather than skill embodied in persons as discussed earlier. The number of BPOs and KPOs has witnessed rapid increase in India providing job opportunities for engineers and graduates from the general stream.[28] This is coupled with the fact that the majority of private institutes impart

[27] Mowery and Sampat (2006) argue in the context of USA that growth in licensing and university based "spin-offs" need not be entirely attributed to Bayh-Dole Act. The university trained graduates facilitated transfer of knowledge of knowledge to industry and other sectors.

[28] The KPO is a higher end version of outsourcing of businesses processes (BPO) that requires domain expertise. India has a share of about 29 % of the world KPO business. Various types, engineering process outsourcing, LPO (legal), India has a large pool of skilled professionals at a low cost with the competitive advantage of English and IT applications. It is estimated that India is getting 30 % of global EPO (engineering) valued at around $6 billion, 20 % of global LPO. Around 300,000 students in engineering per year are churned out in India. India offers 25% cost reduction with salaries 1/10th of the average salary in USA. There are around 300 R&D institutions and national laboratories in India funded largely by the government.

poor quality education as they end up providing students with certificates and not competence. The cycle is vicious; as institutes proliferate in numbers, the demand for faculty goes up, and students prefer to join the job market rather than be researchers and teachers. Salaries are low but the problem is compounded by the fact that private institutes somehow manage with part-time teachers who would be otherwise jobless. The quality of education suffers further as the quality of teachers falls. Education being an 'experience good', students are often taken for a ride by commercially motivated private institutions in the presence of information asymmetry. The adverse selection problem emerges and the hierarchy in the market for technical education gets consolidated resulting in very little improvement in quality of education. As discussed earlier, education as a screening device fails as certificates lose value and credibility in the job market. Frustrations among degree holders mount. Those who can afford to send their children abroad for higher studies, do so. The next best option that remains is relatively higher ranking private institutes within India. As the process of selection has become highly competitive, only the upper middle class ends up competing as expenditure on tuition and coaching classes is on the higher side. In any case, 75–80 % of the students drop out by the time they pass out of Class 10 + 2; India is not in a position to realize its vast human potential.

Transition from Education to Skill: Revisiting the Role of Education

Because of market pressure there is a transition from broad-based value education to narrowly defined skill-based education. The demand for skilled workers can be met but the critical role of education, its democratizing role, and its organic role in the society will be increasingly undermined. Education has a social role (Patnaik 2007). It should have a social perspective. The neo-liberal attack on education has brought it down to the level of a mere commodity. The application of NPM brings down institutions of higher learning to the level of factories churning out certificates assuring the skill possessed by the holder. Teachers are supposed to behave more like workers as their output is measured and monitored for advancement in their career, and an increasing reliance on education loans and increase in the cost of education goad students to behave more like investors. The role of education in inculcating critical thinking and fostering community feeling is sidelined. This is indeed an arduous challenge for a nation. The compulsion of participating in the emerging global knowledge economy is hard to ignore. As HEIs are increasingly inclined to treat HE as a private good, the courses offered are those that are in high demand as the cost of provision has to be recovered. Humanities and social sciences become skill-oriented courses. Even in case of skill-based programs, the focus is often unbalanced as the fashionable IT is preferred to basic engineering

sciences like mechanical, electrical, and civil (Tilak 2010). The best minds are also attracted to IT as pay packages of IT professionals are more lucrative than others.[29]

Concluding Remarks

While India has to embrace the knowledge economy to reap the benefits of global knowledge production and dissemination, she cannot afford to ignore the compulsions of achieving inclusive growth. With the kind of growth process achieved so far, income disparity is getting further accentuated. The government has mooted several policy initiatives to facilitate construction of market with the objective of reaping efficiency gains, both allocative and technical. However, pro-market policies are likely to be inappropriate keeping in mind that linkages between education and society in a developing country context may get weakened. Based on the experiences gathered from world over, there are limits to market-based reform. India has to recognize her strengths and weaknesses and build-up on her strength and overcome her weaknesses. While we need a regulatory authority in view of rapid privatization, the public responsibility for higher education has to go up further to increase 'publicness' of higher education and fund knowledge generation. After all it is a question of prioritization by the government as resource crunch will remain a perennial problem. The government has to lead and HEIs need to adopt strategies to engage in collaborative research with the overall objective of gaining more from cooperation rather than from competition. Ensuring quality of education at all levels remains vital for realizing the goal of inclusive growth with knowledge increasingly assuming the status of being a critical input in the growth process.

References

Agarwal P (2009) Indian higher education: envisioning the future. Sage Publications, New Delhi

Alternative Economic Survey, India (2005-06) Disempowering masses. Daanish Books, New Delhi

Alternative Economic Survey, India (2006-07) Pampering corporates. Paupersizing Masses, Daanish Books, New Delhi

Ashford NA, Hall RP (2011) (ed) Technology, globalization, and sustainable development: transforming the industrial state. Yale University Press, London

Benkler Y (2006) The wealth of nations: how social production transforms markets and freedom. Yale University Press, London

[29] The best minds are also attracted to IT. Even those who do mechanical and electrical from the best of the institutions join the IT industry after being provided with the necessary training.

Bergan S et al (2009) Public responsibility for higher education. World conference on higher education 2009, UNESCO
Chattopadhyay S (2009) The market in higher education: concern for equity and quality. Econ Polit Weekly July 18:53–61
Chattopadhyay S (2007) Exploring alternative sources of financing higher education in India. Econ Polit Weekly 20 Oct:4251–4259
Chattopadhyay S (2010) An Elitist and Flawed approach towards higher education. Econ Polit Weekly XLV(18)
Comim F (2007) Concepts of development: the role of education. In: Bridges D et al (ed) Higher education and national development. Routledge, London
Etzkowitz H, Leydesdorf L (eds) (1997) Universities and the global knowledge economy: a triple helix of university-industry-government relations. Continuum, London
Export-Import Bank of India (2007) Knowledge process outsourcing: emerging opportunities for India, Occasional Paper No. 121. Quest Publications
Government of India (2010) Economic Survey
Government of India (2012) Economic Survey
Hanushek EA, Kimko DD (2000) Schooling, labour force quality, and the growth of nations. Am Econ Rev 90(5):1184–1208
Kapur D, Pratap BM (2012) Higher education. In: Kaushik B, Maertens A (ed) The new oxford companion to economics in India, vol I, Oxford University Press, New Delhi
Lucas, Robert E Jr (1988) On the mechanics of economic development. J Monetary Econ 22(1): 3–42
Marginson S (2004) Australian higher education: national and global markets. In: Teixeira P et al (ed) Markets in higher education, rhetoric or reality, Kluwer Academic Publishers, Dordrecht
Marginson S (2007) Five somersaults in enschede: rethinking public/private in higher education for the global era. In: Enders J, Jongbloed B (eds) Public-private dynamics in higher education: expectations, developments and outcomes. Transcript, Transaction Publishers, New Brunswick
Marginson S (2010) The Asia-Pacific region of higher education: national funding and global influence. Paper presented at U21 symposium 2010, Delhi, on 29 April 2010
Marginson S, Murphy P, Peters MA (2010) Global creation: space, mobility and synchrony in the age of knowledge economy. Peter Lang, New York
Mowery David C, Bhaven S (2006) The Bayh-Dole Act of 1980 and university-industry technology transfer: a policy model for other governments?" In: Kahin B, Dominique F (ed) Advancing knowledge and the knowledge economy, The MIT Press, Cambridge
Olssen M, Michael P (2005) Neo-liberalism, higher education and the knowledge economy: from the free market to knowledge capitalism. J Educ Policy 20(3):313–346
Olssen M, Codd J, O'Neill A-M (2004) Education policy, globalization, citizenship and democracy. Sage Publications, London
Patnaik P (2007) Alternative perspectives on higher education. Social Sci 35(11–12):3–14
Peters MA (2010) The virtues of openness in higher education. In: Marginson et al (ed) 2010 Global creation: space, mobility and synchrony in the age of the knowledge economy. Peter Lang, New York
Rao CNR (2012) Scientific research. In: Kaushik B, Maertens A (eds) The new oxford companion to economics in India, vol II, Oxford University Press, New Delhi
Romer, Paul M (1986) Increasing returns and long-run growth. J Polit Econ 94(5): 1002–37
Romer, Paul J (1990) Endogenous Technological Change. J Polit Econ vol 98(5) pp S71–S101
Stiglitz JE (1999) Knowledge as a global public good. In: Kaul I, Grunberg I, Stern MA (eds) Global public goods: international cooperation in the 21st century. Oxford University Press, New York
Tilak JBG (2010) Universities: an endangered species? J World Univ Forum 3
Solow, Robert M (1956) A contribution to the theory of economic growth. Q J Econ 70: 65–94

Vincent-Lancrin S (2009) What is changing in academic research? Trends and prospects in higher education to 2030, Globalisation, Centre for Educational Research and Innovation, vol 2, OECD, Paris

Winston GC (1999) Subsidies, hierarchy and peers: the awkward economics of higher education. J Econ Perspect 13(1) (Winter):13–36

Chapter 24
HIV Epidemic: Global Response and India's Policy

Arijita Dutta

Introduction

According to the World Health Organization's database of 2004 (WHO 2004), globally 3.5 % of deaths are caused by Human Immuno Deficiency Virus (HIV) and its disease manifestation, the Acquired Immuno Deficiency Syndrome (AIDS), only sixth in position after Ischemic heart diseases, Cerebrovascular diseases, Lower respiratory diseases, Chronic Obstructive Pulmonary Disease (COPD) and Diarrhoeal diseases, while the corresponding share for HIV AIDS among low income poor countries is 5.7 % (4th in rank after Lower respiratory diseases, Ischemic heart diseases and Diarrhoeal diseases) indicating a greater burden of HIV epidemics on the poorer countries. This disease is a special category of communicable disease which is asymptomatic[1] in nature, non-vaccine preventable and its transmission route is more behavioural than environmental. This ailment appeared as a "re-emergence" of infectious disease in an *overlapping of eras* in epidemiologic transition (Bobadilla et al. 1993).

Apart from its epidemiologic characteristics, HIV AIDS has a very special significance for its socio-economic character and linkage to globalisation. While the disease emerged as a zoonotic infection in some of the most economically fragile nations in their post-colonial phase, it soon assumed the shape and size of a pandemic within two decades of globalisation. Starting from a prevalence rate of 0.3 % globally in 1990, it reached a peak of 0.95 % in 2003, hinting towards a galloping pandemic. In 2008, an estimated 2.7 million [2.4–3.0 million] new HIV infections took place. Moreover, this pandemic is extremely uneven in nature: while in 2008 East Asia had an adult prevalence rate of 0.1 % in 2008 and Sub

[1] Asymptomatic disease means without any early observable or perceivable signs among the patients.

A. Dutta (✉)
Department of Economics, University of Calcutta, 56A, BT Road, Kolkata 700050, India
e-mail: dutta.arijita@gmail.com

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_24, © Springer India 2013

Saharan African countries had 5.2 %, Latin American countries, Eastern Europe and Central Asia and South and South East Asia had 0.6, 0.5 and 0.2 %, respectively (UNAIDS 2009). It has been discussed in various studies that globalisation had its severe impact felt on health sector due to fiscal austerity, lack of government support and misplaced macroeconomic policies (Mackintosh Maureen 2005; Mardia Stone 2002; Colvin 2011; Bernett Tony 2007). With the said health reform policies sweeping the countries across geographical regions since 1980s, it resulted in severe roll back in investment in public health in general with lower access to treatments for the poor beyond the *'essential clinical package'* outlined in World Development Report in 1993. Thus, on one hand the HIV catastrophe can be seen as a true harbinger of the global public health crisis (Qadeer Imrana and Kasturi Sen 1998), while on the other hand it emanates from the rising inequality and poverty following the overall economic globalisation. Poverty, in its various manifestations, creates an environment, conducive to the dissemination of HIV infection, while the HIV epidemic poses the greatest threat to every aspect of sustainable human development within an already impoverished region, creating a vicious circle. The role poverty plays in fuelling the epidemic could be highlighted by three observations. First, poor people have less access to information, health services (including treatment) and the means to protect themselves against infection. Poor communities are more difficult to reach with HIV/AIDS prevention messages and services, due to illiteracy, remoteness of location, rural underdevelopment and inaccessibility to communications technology. Second, poor people are less able to act on information they receive, partly because of these social limitations and, clue to the fact that survival needs may propel them into risky survival strategies (e.g., commercial sex). Third, malnutrition and poor living conditions increase the risks of exposure to other infections and health hazards (Stone Mardia 2002).

The disease is unique in another account too. There has been an unprecedented global response to the epidemic and pandemic of this disease primarily because of its huge total welfare loss (both in the current and in the future period), heavy caseload revenue and excess burden transmitted through generations. By 2000, most of the countries have responded to control and reverse the prevalence rate and to offer care and cure for People Living with HIV AIDS (PLHA). Additionally, huge international fund has flown to African and Asian countries to assist the governments in this effort. In 2003, domestic governments, international organisations, foundations and nongovernmental organisations spent an estimated $4.7 billion to address the AIDS epidemic in low- and middle-income countries. This represents a ninefold increase from 1996 to 2003. Out of this total $4.7 billion estimate, international bilateral assistance was of $1.6 billion, multilateral spending and the Global Fund to Fight AIDS, Tuberculosis and Malaria was $1 billion (Commission on HIV AIDS and Governance in Africa 2005). This trend continued and in 2009, international donors and governments together provided US$ 15.9 billion for the global AIDS response (UNAIDS 2010). The United Nations Secretary called for a Special Session of the General Assembly (UNGASS) in New York, in June 2001, which addressed the issue and created the Global Fund for AIDS, Tuberculosis and

Malaria (GFATM), to finance intensified activities around the world (particularly in Africa) and to curtail the continuing and accelerated spread of the epidemic.

According to UNAID's Global Report of 2010; new HIV AIDS infection appears to be declining. In 2009, there were an estimated 2.6 million newly infected people with HIV, which is nearly one-fifth (19 %) fewer than the figure in 1999. The corresponding fall for Sub Saharan African (SSA) countries was from 2.2 to 1.8 million during 2001–2009. However, in this success story one can trace back some big anecdotes of failures. In Eastern Europe and Central Asia, the adult prevalence rate has increased significantly from 0.4 % in 2001 to 0.8 % in 2009. Though overall global death rates declined, there is no evidence of such decrease in Asia and Latin America, while that actually increased in Eastern Europe. In this backdrop, India, with a relatively low prevalence rate is sitting on the tip of an iceberg. In 2008, an estimated 2.27 million people between the ages of 15–49 years of India's population was living with HIV and she carries the largest burden of HIV in the world behind South Africa and Nigeria.

The policy matrix in HIV AIDS, since its very inception, has pointed out that as the transmission route is predominantly behavioural, the paths to control the disease would essentially attempt to alter these behaviours. However, the major determinants of this disease are poverty, existence of malnutrition, lack of livelihood, low literacy, unbalanced urban growth without infrastructure to accommodate rural migration, etc., and so; an attempt to change the behaviours without any focus on altering the environment where these behaviours emerge from, has to face limited success.

In the aforesaid milieu of the crisis, the present chapter attempts to provide an outline of the global, as well as domestic Indian policies to combat the growth of AIDS epidemic. The main research objectives of the chapter are:

- To analyse the global policy matrix to reduce the impact of HIV AIDS, including that of international donor agencies to offer aid for this disease in low and middle-income countries and to locate where exactly these external funds come (in prevention or in cure).
- To focus the political economy of the policy in the light of industrial capabilities to manufacture and export of generic HIV AIDS drug following the provision of compulsory licensing under WTO.
- To identify the specificities of the Indian policy contour and its success and failures.

The International Policy Matrix

Due to its severity in impact, investments in HIV AIDS have been the highest in any particular disease ever. Yet, the gap between investment needs and resource availability is widening and it was US $10 billion in 2009. The Fig. 24.1 shows the

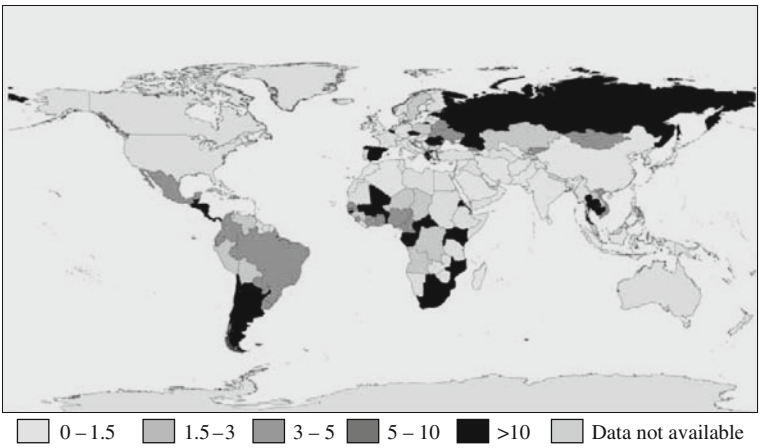


Fig. 24.1 Global map of HIV investment (Both domestic and International spending on HIV AIDS.) per person living with HIV (US $ PPP). *Source* UNAIDS (2010)

distribution of countries with different levels of total investment per capita living with HIV AIDS.

In most of the countries, the sources of the funds have been a complex mixture of domestic government's share, global multilateral agency's intervention as well as bilateral aid. Though the middle-income countries like Brazil, Chile, Argentina, etc., do invest from domestic funds in fighting AIDS, many countries with high prevalence under invest in this area of crucial importance. The new indicator Domestic Investment Priority Index (DIPI) has been calculated which is defined as:

DIPI = (Public Expenditure on AIDS response/Government Revenue)/(PLHA/National population)

This index considers the investment priorities from the domestic government in relation to the disease burden in the country. The Index is calculated by dividing the percentage of government revenue each country directs to the AIDS response by the population share with HIV prevalence. The advantage of this index is that it identifies the priority area of *need* by considering the share of population with HIV prevalence and corresponding priority of the domestic government in dealing with that particular *need.* A high value usually indicates a high level of priority. Table 24.1 reflects that most of the high prevalence countries actually manage to put only a meagre share of their resources in this disease and India is no exception. The median value of the indicator for 104 reporting countries comes out to be 0.35 and majority of them actually have below this average (UNAIDS 2010). Eight of 14 countries in West and Central Africa and six of 16 countries in east and southern Africa appear to be spending less on the AIDS response than might be expected given their disease burden and government resources available.

The other option for HIV response is international fund and majority of it (77 %) comes from bilateral aid. Remaining 23 % comes from multilateral

Table 24.1 DIPI and Persons needing HIV treatment of selected countries

Country	Persons needing ARI treatment	DIPI
Zimbabwe	6,40,000	0.04
Mozambique	5,70,000	0.03
Malawi	4,40,000	0.03
Cote d'Ivoire	2,60,000	0.05
Cameroon	2,70,000	0.06
Vietnam	1,10,000	0.05
India	1,250,000	0.07
Ghana	1,30,000	0.10

Source UNAIDS (2010)

institutions such as the Global Fund to fight AIDS, Tuberculosis and Malaria and UNITAID.

Next, the question comes about where to put the money: in prevention or in treatment and care? Epidemiologic and medical literature suggests that this mix for any disease will depend on a number of things: prevalence rate, absolute number of people suffering, success rate of treatment from existing therapies and inter-generational case loads. If N is the total population in an economy and n is the number of people affected by the disease, then the cost of prevention (which has to cover all population) is $N * P$ (where P is the unit cost of prevention) and the cost of curative treatment is $n * C$ (where C is the unit cost of care and treatment). Now we can say n is actually $N * i$ where i is the incidence rate. From a cost-benefit analysis, the prevention will be preferred if

$$N * P < N * i * C \quad \text{or} \quad P/C < i.$$

For a disease like HIV, the cost of care and treatment is still pretty high with branded Anti Retroviral (ART) Therapy and thus P/C is often small. But at the same time if i is small (which is true in many countries for AIDS), it is more cost-effective to treat the few ill persons, rather than providing prevention for all. Due to the facts that till date therapy of HIV AIDS means just postponement of death, rather than complete cure, many argue that the focus will be more on preventing this issue, rather than cure. However, by the same logic that availability of wonder drug is still not achieved for cancer too, yet the policy for that disease is primarily geared towards care and treatment, rather than prevention.[2] Now let us look at the data on where the countries actually put their resources in fighting HIV. Eighty-one developing countries (with full report of data to UNAIDS in 2010) on the average invest only 34.85 % in care and treatment. Figure 24.2 shows the geographic region-wise share of prevention and care in this disease.

Clearly Latin American countries spend far more on treatment and care, while Sub Saharan African (SSA) countries spend marginally higher in care. On the

[2] The total 11th plan outlay for cancer treatment and diagnosis in India was 2,400 crores and that for tobacco control was 471.92 crores (Annual Report 2012 MH & FW, GoI).

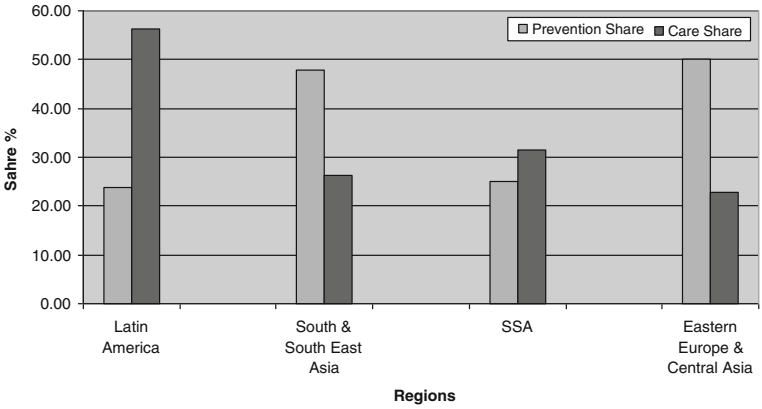


Fig. 24.2 Geographical preference for prevention and care in 2009. *Source Analysed from data of* UNAIDS (2010)

other hand, South and South East Asian countries as well as Eastern European and Central Asian countries prefer prevention more. In order to analyse the reasons behind this particular trend, we need to look deeper into the situation. Nearly 71 countries depend on international sources for funding more than 50 % of their prevention activities (UNAIDS 2010). In contrast, the cost of treatment and care programmes on average appears to be shared equally between domestic sources and international funding. There is a strong correlation between the share of funds invested in care and treatment (C and T) and the share of domestic funds in total investment of the country (correlation coefficient for all developing countries is 0.56 and is significant at 5 % level of significance, while the coefficient for developing countries except Sub Saharan Africa is 0.68 also significant at 5 % level). In fact only in SSA region international funds are invested more in care and treatment (the reasons will be taken up in the next section of this paper). It can be seen that in most of the countries with very high prevalence of HIV AIDS patients per 1000 population, the share of domestic funding is higher in C and T compared to that in prevention (Table 24.2).

Looking at the data on expenditures on prevention per capita show that the overall average figure for 80 countries with data are 1.29 Million US $ per million population, while it is pretty high in Hondurous and Botswana (17.17 and 14.66). Interestingly, the expenditure on prevention per million populations has almost no correlation either with the share of international share in total expenditure or with the spread of the disease in that particular country.

Analysing the causal determinants of the investment share in C and T and then of share in prevention, we run simple OLS regressions with the said share on two regressors: share of domestic country's investment (DOM) and the number of PLHA needing C and T per 1000 population of that country (PATIENTS). While the former represents the domestic country's sovereignty in decision making,

Table 24.2 Share of domestic funds in both prevention and care in eight countries with high prevalence

Country	Prevalence of HIV AIDS per 1,000 population	Share of domestic fund in prevention	Share of domestic fund in care	Expenditure on prevention per million population ($ m)
Botswana	83.70	30	78	14.66
Swaziland	66.50	12	28	7.12
Lesotho	59.25	35	74	4.5
Malawi	33.65	2	5	1.35
Honduras	25.76	45	45	17.17
Mozambique	24.73	3	5	1.67
Kenya	18.39	5	20	4.69
Thailand	5.03	85	98	0.42

Source Analysed from UNAIDS (2010)

the second one stands for the real *need* of the country for treatment. We run the regressions for all 80 countries, then with only countries outside SSA region. The results are depicted in Tables 24.3 and 24.4. We checked VIF tests for multicollinearity and performed ladder analysis for each variable which suggested no change in form is required.

The results confirm that DOM appears to be significant determinant for all countries and for countries outside SSA for the share of C and T. Looking at the partial correlation values for these two variables, we can surely say that increase in domestic country's share in total investment in HIV AIDS increases the share of care and treatment profoundly and significantly. PATIENTS however does not play important role in determining C and T. In other words, the share of domestic funds explains the variation in share of funds invested in C and T more than the actual need. The results are even stronger for the model excluding SSA. Now given the fact that only economically better off countries can afford to have a higher share in total investment in HIV AIDS (the correlation coefficient between GDP per capita and domestic country's share in investment being 0.73), the poorer

Table 24.3 Regression Results of share of care in total investment on HIV AIDS

	For all 80 countries			For countries outside SSA		
	Coefficient	*t* statistics	Partial correlation	Coefficient	*t* statistics	Partial correlation
Patients	0.16	1.11	0.13	0.45	0.64	0.09
DOM	0.37	6.1**	0.57**	0.5	6.39**	0.68**
Constant	18.52	5.47**		9.35	1.81	
Adjusted R^2	0.35			0.45		

Formally, the partial correlation between X and Y given a set of n controlling variables $\mathbf{Z} = \{Z_1, Z_2, \ldots, Z_n\}$, written $\rho_{X\,Y.\mathbf{Z}}$, is the correlation between the residuals R_X and R_Y resulting from the linear regression of X with $\mathbf{Z}$ and of Y with $\mathbf{Z}$, respectively

** Significant at 1 % level

Source Analysed from UNAIDS (2010) data

Table 24.4 Regression results of share of prevention in total investment on HIV AIDS

	For all 80 countries			For countries outside SSA		
	Coefficient	*t* statistics	Partial correlation	Coefficient	*t* statistics	Partial correlation
Patients	−0.36	−2.76**	−0.23**	0.03	0.04	0
DOM	−0.07	−1.29	−0.15	−0.21	−2.75**	−0.37**
Constant	37.48	11.90*		47.73	9.59*	
Adjusted R^2	0.42			0.53		

** Significant at 5 % level
* Significant at 1 % level
Source Analysed from UNAIDS (2010) data

countries who depend on the international funding to confront and fight HIV epidemic, hardly manages to put money for C and T of the vulnerable citizens. As a result only 13 countries within that set of 80 countries have been able to achieve 50 % or higher coverage of ART vis-à-vis the need (according to 2010 WHO guidelines). This is created a crucial *"care gap"* in the policy (Ajithkumar and Irudayarajan 2006).

From Table 24.4, we find that the share of prevention in total investment is determined negatively and significantly by the need variable. The countries with higher prevalence do not invest more in prevention. But for the countries outside SSA, we find that DOM variable has a critical negative role. As the country's domestic share increases, it increases its share in C and T, while decreasing its share in prevention. In those regions the need variable does not play any significant role in determining the share of prevention or care.

Next, we need to look at the components of prevention policies where the countries focus their funds. As the disease has a behavioural path to spread (either by unprotected sex or by injectable drug use), the main focus of such policy was the high-risk groups (HRG) across the globe. Within this group, the main targets are commercial sex workers (CSW), both male and female, injectable drug users (IDU) and men having sex with men (MSM). It was believed in policy documents that targeting these HRG would be able to truncate the spread of the virus considerably (Monitoring the AIDS Pandemic MAP 2004). While the focus on these groups was on the global policy anvil, additionally a few developing countries identified some other sub-sets of population, namely, truckers and migrant labourers as secondary target groups or *'bridge population'* (NACO 2009). However, the general population was left out almost fully from the policy matrix and the vulnerable women folk became victims to the disease quite naturally. Though later the policy to check the Mother to Child Transmission (MTCT) by offering special protection for HIV positive mothers was adopted (UNAIDS 2010), still it has low penetration in developing countries.

Again, while it is universally accepted that prevention programme should be spread among all sections of population almost evenly, the curative system offers

treatment and care only to the morbid persons. Therefore, prevention offered to a small group of people cannot successfully change the behavioural context of the entire population and in fact it is a total mis-specified policy option going against the very definition of economics of prevention. Altogether separate public policy on prevention in two sub-groups, namely, HRG and LRG (Low-Risk Group) may not be effective in creating actual private demand for prevention (here actual usage of condoms for example) owing to different prevalence elasticity of private demand for prevention[3] among them. A higher prevalence among the peers in any group may induce more demand for using condoms and a lower prevalence among peers will result in lower usage of condoms (Philipson Thomas 2000). Therefore, this kind of fragmented policy prescription might alter the behaviour of the non-positive individuals in HRG, but might not be able to do that for non-negatives in LRG. The empirical findings of higher condom use among CSW and not among general population support this hypothesis.

Thus in short, the international policy matrix appears to follow a common intervention framework of strategies with more international money financing preventive care and domestic governments spending more on C&T. The prevention intercession has primarily focused on the HRG groups, though in most of the Asian countries the routes of spread has reached the general population. Continuous high prevalence eventually seeps into lower risk part of the population which has not yet been brought into the coverage nets of the policy. On the other extreme the epidemic is getting more concentrated among HRG in Eastern Europe and Central Asia. Thus, while the contour of epidemiology of HIV AIDS appear to be a kaleidoscopic mosaic within countries and regions, one singular format of policy is bound to leave out the intricate details. In creating this policy format and virtual tying of international aids in the prevention related areas, industrial dynamics too played an important role which is dealt in the next section.

Industrial Capability and Access to Medicine in HIV AIDS

With limited funds for C and T, availability and access to treatment to HIV AIDS in developing countries have always been an issue of concern across the countries. The costs of these drugs have been observed to be pretty high. In 1996, Highly Active Anti-Retroviral Therapy (HAART)—the first effective combination therapy that delays the onset of AIDS from being HIV positive—became available to those living with HIV in developed countries, where within 4 years, death rates for

[3] The prevalence elasticity of private demand for prevention is defined as the rate of change of actual private demand for prevention with respect to rate of change of prevalence. This is a concept related to *Economics of Epidemiology*. It says that the general public will actually utilise the prevention methods (here use of condoms) depending upon the prevalence rate of the disease at that time.

PLHA dropped by 84 %. Due to dearness of these drugs (costing US$ 10,000 to 15,000 per person per annum) they could reach only 2 % of people in developing countries after 5 years of introduction of HAART. TRIPS being accepted by most of the developing countries by that time meant stricter intellectual property regime and escalation of prices of branded drugs worldwide. With seventeen ART drugs under patent protection, the global market of these drugs is oligopolistic with only seven producer firms. At this time, Brazil and Thailand have found HIV epidemic detrimental to their economy and they started producing and marketing ART drugs themselves to provide access to the medicines to their citizens. The US Trade Representative considered this action as threats to American corporate interests and the United States named Thailand and Brazil in US article 301 Special procedures, threatening them with serious reprisals and trade sanctions (Kaiser Family Foundation 2007). Soon Thailand had to adopt a "Super-TRIPS" like arrangement to satisfy some US demands (where exclusive rights for marketing any new drug was given to multinationals for two years). Later when South Africa issued compulsory licenses, 39 big pharmaceutical companies tried to prosecute the South African government which, however, finally had to be stopped on the face of public protest.

At the beginning of the new millennium there was a breakthrough in treatment provision for poor countries when the Indian pharmaceutical company Cipla started to produce generic antiretroviral at significantly cheaper price. This innovation, coupled with bargaining pressure from activists, organisations (such as the Clinton Foundation) and governments of poor countries with severe AIDS epidemics, dramatically reduced the price of ARVs for developing countries. By 2001, *triple combination therapy* was available from Cipla for as little as $295 per person per year (Médecins Sans Frontières 2001). These Fixed Dose Combination (FDC) therapies reduced the number of pills per day and made it easier to manage for both the patient as well as health workers especially in developing countries. Also these cocktail drugs were available in heat resistant forms, which proved extremely valuable for use in the developing world, where refrigeration facilities are often scarce. Cipla could market these cocktail drugs (combination of three drugs) as there was no product patent in India at that time. The patents for the individual drugs were lying with separate multinationals globally, making it impossible to produce one single dose by anybody there (Guennif Samira 2004). Following this, the price of drug went down drastically for both generic as well as branded drugs (Fig. 24.3).

Now till date, India is the largest supplier of generic ARVs to low- and middle-income countries (the main producers being Cipla, Matrix, Ranbaxy, etc.) providing 80 % of donor-funded ARVs, followed by Brazil, Thailand and South Africa. A few African nations like Zambia, Ghana, Tanzania, Uganda, Zimbabwe and Kenya have developed local HIV drug manufacturing facilities. National governments of 94 % of countries with generalised epidemics, and 61 % of countries with concentrated epidemics, had national policies for using generics to promote antiretroviral access (UNAIDS 2008). The generic export from these

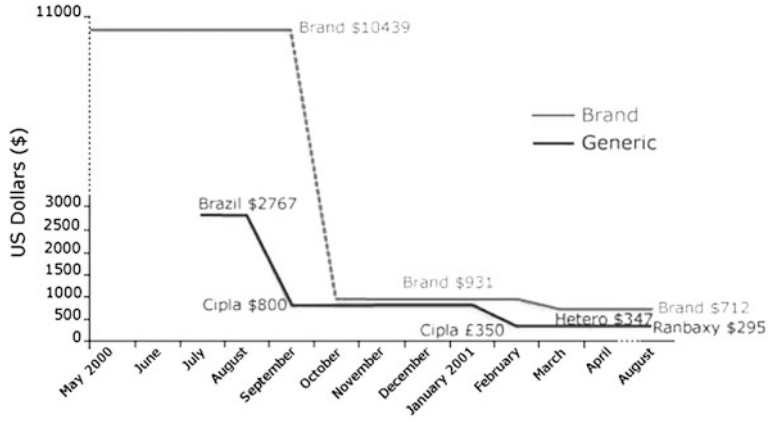


Fig. 24.3 Fall in ARV drug prices in 2000–2001. *Source* www.avert.org as accessed in January, 2012

aforesaid countries contributed around 98 % supply provided through international NGO the President's Emergency Plan for AIDS Relief (PEPFAR) in 2004–2005 (www.clintonhealthaccess.org accessed in February 2012).

This story essentially brings back the arguments of (Bueno et al. 2005) where they assert that foreign aid given by democratic governments is not given selflessly to assist underdeveloped countries, but rather is given to buy policy concessions, typically from autocratic governments. Given the fact that big pharmaceutical multinationals as well as the governments in developed countries severely opposed the introduction of generic in ARV and subsequent price fall, it is not difficult to understand why they do not finance the buying of such drugs from generic producers from India and Brazil.

In reaction, the large pharmaceutical countries (often termed as 'big pharma'), as well as generic producers, entered in price negotiations with organisations namely Clinton Foundations, Médecins Sans Frontières and UNITAID to regain their markets in high prevalence countries by 2001. Clinton Foundation claims to have reduced prices up to 90 % for key drugs for HIV, malaria, and TB, and up to 80 % for HIV diagnostics, saving more than $1 billion in total (www.clintonhealthaccess.org, accessed in January 2012). Not only for first-line drugs, these organisations negotiated with companies like Pfizer, Mylan to reduce the prices of second generation drugs below $500 per year in 2008 (www.avert.org, accessed in January 2012). These international NGOs primarily work in SSA where their share of funds is pretty high. For example, the share of international aids from GFATM, UN and other Multilateral agencies in Botswana was 11.5 cent in total funds, 41 % in Zimbabwe, 10.6 % in Lesotho, 10 % in Mozambique, etc., and this fund is primarily going to C and T. This means that these international foundations acted as a monopsony player in the market of HIV AIDS drugs in certain high prevalence countries, especially in Sub Saharan Africa.

The twist in the story comes as we look at the limited availability and coverage of ARV treatment in India in spite of its proven industrial capabilities. The main reason behind the lack of access to ART in India is sheer cost: while the Indian government spends Rs 7,500 PPY for first-line ART drugs, it needs to spend Rs 1 lakh PPY for second-line drugs. This is, however, not because of the presence of patent on the second-line drugs since Indian companies do produce them all. The Indian pharmaceutical firms like Ranbaxy, Cipla, Matrix, etc., reduce the prices of both first-line and second-line drugs facing price negotiation in fear of losing their market shares in other developing countries. Thus, while Cipla charges Rs 54,000 PPY for a combination of TDF, FTC and EFV in the domestic market, Matrix agreed to sell the same drug to the Clinton Foundation at a cost of $389 (Rs 15,847) PPY and to reduce the price of TDF, 3TC and EFV to $ 339 (Rs 13,899) PPY. But due to lack of that market competition in India and any attempt of negotiation from the government, Indian companies continue to charge higher prices. This, coupled with limited share of funds allocated to C and T in India, results in extremely low access of ARV treatment in India.

As a result, in 2010 only 26 % PLHA and needing ART (according to WHO 2010 guidelines) have access to the drugs in India. But the corresponding figure for children alone is 41.5 primarily owing to The Clinton Foundation supporting the Government of India with $35 million of US$ 50 million needed for providing ART drugs for life to children under 15 years of age. There is also an urgent need for affordable second-line treatment. Estimates of people who have developed resistance to first-line drugs and who need second-line drugs range from 1,800 to 35,000 (Gopakumar 2011). In addition to this scenario of far-from-universal-access for ARV drugs in developing countries, the access to other drugs for opportunistic infection is also limited, which essentially creates further financial burden on the patients. Given this backdrop, countries like India faces the most unfortunate scenario: lesser amount of fund for treatment, collusive oligopoly in domestic market by Indian drug producers and very limited access to medicines in spite of high industrial capabilities.

However, the issue of access to ART treatment and care is not just a concern related to supply side alone. The Chilean experience shows that even in a country with a relatively low epidemic and an intermediate level of development, the economic and organisational consequences of scaling up access to ART are far from minor at both the individual and the community level (Morales Christian et al. 2003). The success of scaling up access to ART will depend not only on financial issues but also on issues regarding the organisation of the health system itself, especially the drug distribution process, availability of trained manpower and above all the awareness generation to avoid stigma among the general population.

India's Epidemic and Response in Policy

Nature of the Epidemic

With the aforesaid policy backdrop and supply bottlenecks, many regions including Asia suffered setbacks in the new incidence rates of HIV AIDS. AIDS-related diseases continue to be leading causes of death in this region. Table 24.5 summarises the prevalence rate in Asia. The epidemiologic spread of the disease also followed a different path in parts of Asia; the epidemic here started among the High Risk Groups (HRG), namely, CSW, MSM, IDU; it then spread to the truckers and migrant labourers and now it is more characterised by significant transmission among couples with monogamous women partners. This path of transmission is called *Path 4 (*Monitoring the AIDS Pandemic MAP 2004). This was confirmed by a recent meta-analysis of available data in Viet Nam, which found that many low-risk women may be at considerable risk of HIV infection due to the high-risk sexual and drug-using behaviours of their male partners (Nguyen et al. 2008).

Though with the exception of Thailand, every country in Asia has an adult HIV prevalence of less than 1 %, the economic consequences of AIDS will force an additional 6 million households in Asia into poverty by 2015 unless national responses are significantly strengthened (Commission on AIDS in Asia, 2008). Table 24.6 summarises the picture of HIV prevalence and the investment response in the major Asian countries. It clearly reflects the same trend as we have discussed in "Introduction" of this chapter. India clearly invests lowest among these countries and its share in C and T is still lower than 40 % threshold. Malaysia and Thailand invests the most from the domestic sources.

Now coming to the disease prevalence particularly in India, the rate has marginally gone down to 0.36 % in 2007 from 0.45 % in 2002. Yet, India is home to about two-third of the people living with HIV in Asia, posing a serious public health challenge. In fact India has the second highest number of HIV AIDS patients all over the world. However, what is alarming is the increased share of adolescent population and that of above the age of 50 years. This indicates that the vulnerability for the disease is not only higher among the young people in their reproductory ages, but the general populations in their teens and in their childhood too are getting affected.

Table 24.5 HIV prevalence in Asia

Indicator	2001	2008
People living with HIV	4.5 million (3.8–5.2 million)	4.9 million (4.5–5.5 million)
New HIV infections	400,000 (310,000–480,000)	360,000(300,000-430,000)
Children newly infected	33,000 (18,000–49,000)	16,000 (11,000–21,000)
AIDS-related deaths	280,000 (230,000–340,000)	300,000(260,000–340,000)

Note The ranges around the estimates in this table define the boundaries within which the actual numbers lie, based on the best available information
Source UNAIDS 2009

Table 24.6 Prevalence and investment for AIDS in Asian countries, 2009

Country	ARI per 1000	Total inv per HIV patient	Domestic government share	Share of C and T in total investment
Bangladesh	0.01	17.96	0	8.91
Cambodia	2.99	1.30	10.1	28.56
India	1.03	0.11	16.5	37.32
Indonesia	0.31	0.68	40	14.78
Lao	0.32	3.00	1.9	16.04
Malaysia	1.52	0.64	98.4	37.71
Myanmar	2.48	0.27	4.7	38.66
Nepal	1.16	0.57	67.8	16.62
Pakistan	0.20	0.56	78.4	7.61
Philippines	0.02	5.23	16.2	8.70
Sri Lanka	0.05	1.41	33.2	4.89
Thailand	5.03	0.61	93.3	76.06
Viet Nam	1.27	0.94	2.1	46.01

Source UNAIDS (2010)

Geographically, seven Indian states have high prevalence rate of the diseases and they are Tamil Nadu, Maharashtra, Andhra Pradesh, Karnataka, Manipur, Nagaland and Delhi. Eight more states are identified as vulnerable; they are Uttar Pradesh, Rajasthan, Chhattisgarh, Jharkhand, Bihar, Madhya Pradesh, Uttaranchal and Orissa (Care 2006). Studies reveal that other states like West Bengal, Madhya Pradesh, Gujarat, etc., are also moving towards the tipping point in vulnerability. Higher HIV prevalence among IDU is specifically observed in North Eastern States, though new sites of IDU prevalence have been identified in Punjab, West Bengal, Tamil Nadu, Kerala and Maharashtra, representing a *dual nature of the epidemic in the country* (NACO 2007). Figure 24.4 shows that there are several districts in the country with high prevalence among general population attending ANC clinics, while there are some distinct areas where the spread is primarily among the HRG.

India's Response

Since the first HIV/AIDS case was traced in Chennai in 1986, National AIDS Committee was formed under the Ministry of Health and Family Welfare in 1987. In 1992, National AIDS Control Organization (NACO) was formed and the first phase of National AIDS Control Programme (NACP) was launched primarily with the help of extensive external funding to target the HRG to control the spread and incidence of AIDS in the country. NACP I (1992–1999) focused on expansion of infrastructure of blood banks and for treatment of Sexually Transmitted Diseases (STD). State AIDS Cells were commissioned since 1999, along with large-scale Non Government Organizations (NGO) to be linked with the programme for

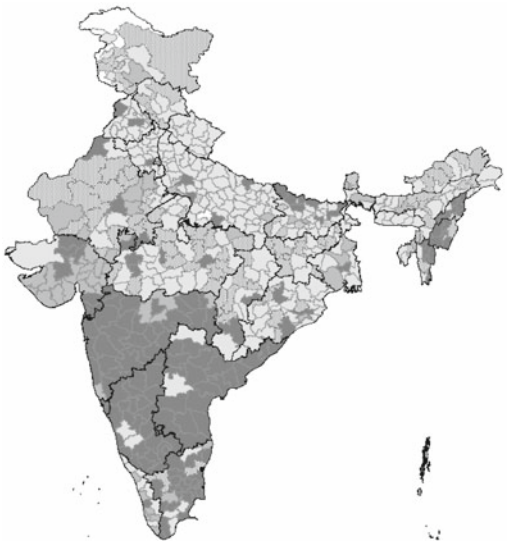


Fig. 24.4 District-level epidemiology, risk and vulnerability in India 2010. *Source* NACO (2006)

spreading awareness and proving care and support to the PLHA section. NACP II (1999–2006) targeted the HRG people, and stressed on condom promotion and voluntary counselling facilities. School AIDS Education was conceptualised to build up life skill and awareness among the youth and adolescent students.

With these two extensive programmes, NACO claim to have created awareness about HIV AIDS among 84.6 % of rural population, put 47,693 people on ART, distributed 1.6 billion condoms and covered more than 93,000 schools under life skill generation and awareness creation network (www.nacoonline.org, accessed in January 2011). The sources of funding have been overwhelmingly external, as clearly identified in Fig. 24.5. Government of India spent only 25 % of total expenditure under NACP III. The largest contributor was World Bank (15.71 %), followed by GFATM (15.42 %).

NACP III has been launched in 2006 for 2006–2011, with a total package of Rs 7,009 crores. Table 24.7 shows that even in the third phase, prevention programmes grabbed an overwhelming share of funds, while treatment and care continued to be in the shadow.

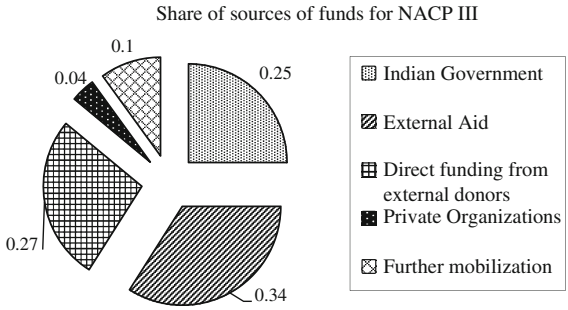


Fig. 24.5 Sources of funds under NACP III by their shares. *Source* Analysed from data from NACO (2008)

Table 24.7 Shares of different components in NACP III

	Item	Share in total
1	*Prevention*	67.2
1.1	TI for HRG	19.7
1.2	Others (migrants, truckers, prisoners, etc.)	1.1
1.3	Package of services	12
1.4	Blood safety	8.2
1.5	Communication etc.	8.8
1.6	Condom promotion	17.3
2	*Care, treatment, support*	16.9
2.1	ART	11.5
2.1.1	ART children	1
2.2	Care and support	4.3
3	*Capacity building*	7.9
3.1	Mainstreaming, PPP	1.1
3.2	Training	1.9
3.3	Capacity strengthening	2.4
4	*Strategic info management*	3.1
4.1	Monitoring and evaluation	1.7

Source NACO (2006)

In spite of such large scale and extensive effort from the government, it has failed to create substantial improvement in behaviour, awareness and prevalence of the diseases and the overall policy framework can be criticised from a number of view points.

- *First*, owing to specific socio-cultural environments as well as the nature and transmission routes across the districts (depicted in Fig. 24.4), a uniform policy schedule for all the states can be a wrong method altogether to start with. The programmes have been able to increase condom usage among HRG in significant way: in West Bengal it jumped from 52 % in 2001 to 88 % in 2004 while in

Andhra Pradesh it increased from 62 % in 2002 to 97 % in 2005 (NACO 2006). However, the huge amount of money spent on condom promotion has not been successful in penetrating the general population in India. The total share of married women in India whose partners were using condoms in 2004–2005 was only 5.2 %, while the condom use among the poorest class, undoubtedly the most vulnerable section of the population was found to be as low as 1.2 % (National Family Health Survey 3 2005). Survey for 2005 also points out that in only 9 % acts of sexual intercourse with non marital/non cohabitating partners condoms were being used, a decline since 2003 (UNAIDS 2006). This identifies that improving the supply of free condoms has not been effective in a country like India where the literacy rate has been rather poor. The government recognised that the 'ABC'—or Abstinence, Behavioural change and Condom use approach is of little effect, given the varied risk factors for infection (Panos India 2007). The increase in condom use among HRG while Low-Risk Groups remaining at the same low level proves the theoretical issue of prevalence elasticity of private demand for protection discussed in the "Introduction" in the chapter.

- *Second*, out of the total prevention programmes, the largest share of expenditure has been received by the HRG, including Commercial Sex Workers, IUD, MSM, etc., while a smaller share has gone to other *bridge population*, like truckers and migrants. In spite of the fact that several reports, including that from NACO itself identifies that the path of spreading the infection has taken different contours (path 4) in recent years and monogamous married women are being infected by the diseases more, focus just on the HRG population cannot result in expected gains. Given the low social status of the women in India, they hardly enjoy any right to choose their contraceptive methods and given the choice between the immediate threat of domestic violence and relatively hypothetical spectre of HIV, women often resign themselves to sexual demands increasing the risk of being affected by HIV. NACP's prevention programme leaves out this huge low risk, but extremely vulnerable general women out of their prevention net.
- *Third,* NACP relied more on international funds and hence the policies followed were dependent on the expertise and experience of the Northern countries, rather than identifying the local resource positions. Therefore, focusing on 'blood safety, alone, neglecting the overall precautions for sterilisation of all medical equipments, has partly neglected the lacunae in the health care market on the whole. By this NACO created a management structure parallel to the already existing one, at a high cost (Priya et al. 2006).
- *Fourth,* the issue of access to ART medicine has been already discussed in earlier sections. With public roll-out of ART drugs in India in 2006, it was expected to cover a target of 300,000 by the end of 2011 and planned to create 250 ART centres for this. It also waived the testing cost of CD4 counts to improve access. However, the lower than the target coverage highlighted the problems of drug resistance and unavailability of second-line ART drugs in

public programme, lack of government's negotiation capabilities with pharma companies and limited access to drugs and treatment for Opportunistic Infections (OI), like Tuberculosis, Malaria, etc.

- *Fifth*, though HIV AIDS has a separate commission under NACO in the MH&FW, the delivery of services has to be done through the health department mechanism only. While health is primarily a state subject, the efficiency, expenditure, etc., is formally controlled by state governments and hence there exists a vast difference in character of health departments in different states. Moreover, the overall Public Health situations in India is extremely poor and state governments typically spend only 8–9 % of total health expenditure (GOI, MH and FW 2005). In West Bengal we find a shocking drop in the share from 11.18 % in 2004–2005 to 8.99 in 2009–2010, in spite of the launch of NRHM (Dutta A and Ratan Khasnabish 2011). Given this infrastructure, it is unviable for the health department to run the huge prevention policies of NACP III with their existing staff strength. Also the entire manpower of the department is trained in different types of curative system and suddenly thrusting them with the burden of HIV prevention seems to me like asking an economist to teach physics in the colleges! 'An unprepared public health system with no transparency is in no position to handle such an intensive programme' (Panos India 2007). A prior change of orientation of the Public Health management is a necessary, though not sufficient condition for successful results.
- *Last* but not the least, the NACP has not only overlooked the conditions of malnutrition, it also neglected the crucial part of alternative methods of care for the infected and affected families, including the children. Given the fact that a big proportion of people in India are severely malnourished, more so among the women, they become extremely vulnerable in fighting HIV AIDS. Specific ART drugs needing a threshold body weight for application makes the access to treatment more difficult and hence, though informed and aware of the disease, they suffer from inaccessibility of right medicine at right time.

In parallel to NACO, some NGOs have developed their own model of support for low-risk and highly vulnerable community where focused targets have been well integrated in a broader framework, keeping in mind the trajectory of the disease and its related exposure to risk (Dutta 2010). While health care support is necessary for treating HIV AIDS, Opportunistic Infections, as well as those diseases and symptoms which occur commonly among vulnerable population, it cannot bring the desired result, unless it is backed by nutritional support programme as malnutrition reduces immunity severely. However, nutritional support is not a desirable as a long-term strategy unless it is complemented by a creation of alternative livelihood plan. Thus, a holistic approach is necessary to support the PLHA, while ensuring awareness to truncate further flow of the disease.

Conclusion

From the above discussion it becomes clear that HIV AIDS has received unprecedented response and policy attraction across the globe. However, due to critical politico-economic lobbying from various interest groups the greater part of the huge international aid has remained tied to prevention programmes, that also among HGRs and Bridge Population in most of the developing countries. While this not only contradicts the economics of epidemiology and prevention, it also leaves the so-called 'low risk groups' almost untouched by the programme nets. The generic export of ART from Indian companies and systematic bargaining of big NGOs have played instrumental role in reducing the prices, thus increasing the ART coverage. However, some countries yet face unreasonable pricing and hence the access to latest potent drugs has remained limited. India exactly matches this very scenario with highest number of HIV affected people, low priority from the domestic government and low investment for care and treatment, coupled with relatively high prices of ART in absence of tight public negotiations.

The central problem of Indian programmes under the tying international aid is that by introducing a set of external changes, but without any change in the endogenous context of the people, they are trying to bring shifts in their behavioural patterns. As most of the determinants of being affected by HIV AIDS are endogenous in nature (like lack of nutrition, endemic disease burden, lack of access to healthcare, absence of alternative employment generation packages, etc.), there is a strong need for upstream interventions that can change the biomedical risk environment. The overall poor infrastructure of health care system and a critically low level of public health facilities truncates the possibility of maintenance of blood safety and reduction of MTCT possibilities. What we now offer is the opportunity to make a somewhat unattractive decision: to choose surviving through risky behaviour in a risky environment until they succumb to death (Stillwaggen Eileen 2006). Offering just information without any alternative is surely not going to change the situation much.

References

Ajithkumar K, S Irudayarajan (2006) Medical care for HIV: an opportunity or crisis? EPW 15:1428–1429

Barnett T (2007) HIV/AIDS and globalization: what is the epidemic telling us about economics, morality and pragmatism? Working Paper. School of Development Studies, UEA, Norwich

Bueno de Mesquita B, Morrow JD, Siverson RM, Smith A (2005) The logic of political survival. MIT Press, Cambridge

Bobadilla JL, Frenk J, Lozano R, Frejka T, Stern C et al. (1993) Cardiovascular disease. In: Jamison DT, Mosley WH, Measham AR, Bobadilla JL (eds) Disease control priorities in developing countries. Oxford University Press, Oxford

Colvin CJ (2011) HIV/AIDS, Chronic diseases and globalization. Glob Health 7:31

Commission on HIV/AIDS and Governance in Africa (2005) Economic Commission of Africa Index No. CHGA-B-I1-0002
Dutta A (2010) Mid term evaluation of FXB villages in Chaipat in West Bengal (Unpublished Report)
Dutta A, Ratan K (2011) Health sector reform in West Bengal: a GP level analysis. In: Chattopadhyah PK & Bhattacharya S (ed) Challenges of livelihood & inclusive rural development in the era of globalization. New Delhi Publishers ISSN 9789381274255
Guennif S (2004): AIDS in India: public health related aspects of industrial and intellectual property rights in a developing country. CSH Occasional Paper. New Delhi
Gopakumar KM (2011) Why are AIDS Drugs unaffordable in India? www.infochangeindia.org. Accessed Jan 2012
Government of India, Ministry of Health and Family Welfare (2005) National Health Accounts. Central Bureau of Health Intelligence
Kaiser Family Foundation (2007) U.S. Trade representative places Thailand on priority watch list in annual report
Qadeer I, Sen K (1998) Public health debacle in South Asia: a reflection of crisis in the welfarism. J Public Health Med 20(1):93–96
Mackintosh M, Koivusalo M (2005) Health systems and commercialization: in search of good sense. In: Mackintosh M, Koivusalo M (ed) Commercialization of health care: global and local dynamics and policy responses. Palgrave Macmillan, Basingstoke
Stone M (2002) Globalization and HIV policies in Africa. Aid Transparency, Regional think tank in HIV AIDS
Monitoring the AIDS Pandemic MAP (2004) AIDS in Asia: Face the facts
Médecins Sans Frontières (2001) A matter of life and death: the role of patents in access to essential medicines
Morales C, Cid Pedraza C, Souteyrand Y (2003) Expanding access to antiretroviral therapies in Chile: economic and financial issues for patients and the health system. In: Moatti J.-P , Coriat B, Souteyrand Y, Barnett T, Dumoulin J, Flori Y.-A (ed) Economics of AIDS and access to HIV/AIDS care in developing countries: issues and challenges ANRS, Paris
NACO (2006) National AIDS control programme III'. MH & FW. GOI
NACO (2007) Annual report. Ministry of health & family welfare, GOI
NACO (2008) UNGASS country report. India
NACO (2009) Annual report. Ministry of health & family welfare, GOI
National Family Health Survey 3 (2005) Full India report. IIPS. Mumbai
Nguyen TA et al. (2008) A hidden HIV epidemic among women in Vietnam. BMC Public Health, DOI:10.
Panos India (2007) Antiretroviral drugs for all? Obstacles in accessing treatment: lessons from India
Philipson T (2000) Economic epidemiology and infectious diseases. In: Cuyler A, Newhouse J P (ed) Handbook of health economics, vol 1. Elsevier, North Holland
Priya R, Kumaraswami N, Qadeer I, Ganesh AK, Solomon S, Crowe SM (2006) India. In: Beck E, Mays N (ed) HIV pandemics and health Systems: an introduction. OUP, UK
Stillwaggen E (2006) Determinants of the HIV pandemic in developing countries. In: Beck E, Mays N (ed) HIV pandemics and health Systems: an introduction. OUP, UK
Mardia S (2004) Globalization and HIV AIDS policies in Africa. Aid transparency. Regional think tank on HIV/AIDS
WHO (2004) Burden of diseases. www.who.org
UNAIDS (2006) Global report
UNAIDS (2008) Commission on AIDS in Asia
UNAIDS (2009) AIDS epidermic update
UNAIDS (2010) Global report on HIV AIDS

Chapter 25
The Political Economy of Mental Health in India

Anup Dhar, Anjan Chakrabarti and Pratiksha Banerjee

> Psychology is in need of its own *Das Kapital*.
> Lev Vygotsky, *The Historical Meaning of the Crisis in Psychology.*

Introduction

Marx writes in the section on 'The Method of Political Economy' in *Grundrisse*: "even the most abstract categories, despite their validity—precisely because of their abstractness—for all epochs, are nevertheless, in the specific character of this abstraction, themselves likewise a product of historic relations, and possess their full validity only for and within these relations" (1993, p. 105). Marx had in mind the category of labor. Mind, psyche, reason, madness, and health are, however, no exceptions. They are also the product of historical relations and they possess their full validity only for and within these relations. Medicine, psychiatry, psychology, and the clinic also need to be seen within the "historical constitution of [their] own practices" and the "political-economic conditions in which [they] became possible" (Parker 2011, p. 1) as also the discourse within which they grew roots (Foucault 2006). The "invention of the mind" by Descartes as "mirror of nature" (Rorty 1979, p. 357) and the "Cartesian progression of doubt" as the condition of the "great exorcism of madness" (Foucault 2006, p. 244) would therefore serve as philosophical signposts and historical roadmaps in our foray into the political economy of mental health. The account of the Indian state, its 'first transition' into a developmental regime post-1947 and its 'second transition' into neoliberal forms of globalization post-1989 (Chakrabarti et al. 2012, pp. 200–282; Samaddar and Sen 2012, pp. 1–8), the movement from 'welfare medicine' to 'development medicine'

A. Dhar
School of Human Studies, Ambedkar University, Kashmere Gate Campus, Lothian Road, New Delhi 110006, India
e-mail: anup@aud.ac.in

A. Chakrabarti (✉) · P. Banerjee
Department of Economics, University of Calcutta, Kolkata 700050, India
e-mail: chakanjana@yahoo.co.in

P. Banerjee
e-mail: chakanjana@yahoo.co.in

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2_25, © Springer India 2013

to 'neo-liberal medicine' (Zachariah et al 2010, pp. 9–23) shall also serve as relational pointers to the understanding of the political economy of mental health. The Marxian maxim: "political economy is not technology" (Marx 1993, p. 86), but is an (psycho)analytical *fort-da*[1] between a tracing of the "simplest determinations" and a retracing of the whole as a "rich totality of many determinations and relations" shall also be our guiding principle in this foray. We shall also work at the overdetermination of the axes of the 'patient' (which includes experience and suffering), the 'professional' (which includes listening, diagnosis, cure, care, and medical education), the 'service provider' (which includes public institutions with tertiary [medical colleges], secondary [district hospitals] and primary [rural health centers] care delivery systems as also private clinics) and 'industry'[2] [which includes, on the one hand, circuits of global capital and global markets and, on the other, circuits of local capital and local markets (Chakrabarti et al. 2012, pp. 133–199)]. The axis of the patient will, however, be our entry point. This paper shall therefore begin with a take on the experiences of the patient and shall end with a take on the suffering the patient goes through.

Experience

> In one sense, the expropriation of experience was implicit in the founding project of modern science. … modern science has its origins in an unprecedented mistrust of experience as it was traditionally understood (Bacon defines it as a 'forest' and a 'maze' which has to be put in order). The view through Galileo's telescope produced not certainty and faith in experience but Descartes's doubt, and his famous hypothesis of a demon whose only preoccupation is to deceive our senses (Agamben 2007, pp. 19–20).

[1] Freud (*Beyond the Pleasure Principle* 1920) understood *Fort* and *Da* (German for *gone* and *there*) as the dialectic between disappearance and return, absence, and presence. We see the dialectic between 'whole' and 'simple determinations'.

[2] India has a vast pharma market. Approximately, 50 % of the essential medicines that the UNICEF distributes in developing countries come from India. About 75–80 % of all medicines distributed by the International Dispensary Association (IDA) to developing countries are manufactured in India. "The strength of the pharmaceutical industry is evident in the existence of 5,877 manufacturing units in India producing 20,000 formulations under 8,000 branded names. … This rosy picture of the pharmaceutical industry however is reflected neither in availability of drugs for important public health problems nor in the affordability of drugs for domestic users [especially an estimated 649 million people, people below or near the poverty line, most of them not covered by insurance or social security mechanisms]". Note worthy in this context is the (not-so-secret) connection between on the one hand, the indigenous drug industry and on the other, the space of medical education, research, and the majority of the national pool of psychiatrists. In India, the drug industry through its aggressive marketing mechanism and through a "structure of incentives, enticements and perks" influences even the thinking and practice of psychiatrists; such that the philosophy of suffering-healing come to be colonized by *only* neurobiological causalities and pharmaceutical remedies; it becomes difficult for most psychiatrists to see anything other than biology-as-causing-suffering and medicine-as-contributing-to-healing; economic-political-cultural causations and psychotherapeutic contributions to healing get severely devalued in Indian hospitals.

The "mistrust of experience" is deepened when the experience is the 'experience of unreason'; and this paper is precisely about *that* experience.[3] However, it is not just about mistrust. It is about the "caesura that establishes the distance" between reason and unreason; it is about the historical absence of an *oeuvre* of unreason; it is about the silence that envelops unreason and the "archaeology of silence" (Foucault 2006, pp. xxvii–xxxvi). It is about both the disavowed and the near-forgotten experience of unreason. It is also about the history of *a* relation; about a relationship we have (had) with unreason, which is also the *absence* of a relationship, which is also the history of a *turning away* from unreason; of a kind of 'double forgetting', forgetting the fact that one has forgotten unreason. Yet one cannot forget unreason. Unreason resides in the very pores of Reason; deferring its solitude; hence the disavowal. It is also about a twist in the tale, perhaps. Perhaps because one can never be certain about the relation the contemporary is setting up with unreason; one can only 'know' retrospectively. Twist in the tale, in the sense that the contemporary is throwing up unthought of relations with unreason. The contemporary is bringing its 'private world' and its 'exiled experience' or the 'experience of exile' within the ambit of the hitherto incomplete 'psychiatrization of the self', as if, we have never been psychiatric enough. Psychiatry has thus become one of the most powerful forms of social control in late capitalism. It operates efficiently precisely "because it is double-edged. … On the one side it is a technical instrumental apparatus now at one with the pharmaceutical companies and enrolled into the delirious search for neurological certainties. … On the other side is a system of patronage and disdain for any kind of dissent, all the more dangerous when it senses that its power is precarious" (Parker 2011, p. 24); the political economy of mental health in India is determined by this "double-edge". Further, the contemporary is bringing unreason under and within modern institutions, institutions which were earlier called 'mental asylums' and are presently called 'mental hospitals'. The contemporary is subsuming it within a largely burgeoning pharmaceutical industry and somewhat secondarily within the ambit of psychotherapeutics (including all kinds of counseling services, psychological testing, and pop-spirituality). It is also bringing its more public expressions under psychological scrutiny: psychiatry as an ideological state apparatus enforcing "adaptation", which normalizes certain behaviors and pathologizes others, and "mental hygiene" has come to operate as "one of the names of politics in contemporary culture" (Parker 2011, p. 50). Psychology and psychoanalysis—because they unknowingly borrow from psychiatry its diagnostic logic and framework—find themselves trapped in the very psychiatrization they sometimes oppose.

[3] However, we do not wish to make this 'turn to experience', function "as an all the more seductive ante-room to the psychiatric clinic", that will in turn intensify 'psychiatrisation' (see Parker 2011, p. 34). "What could work [perhaps] is a *return*—from the perspective of the excluded as resource—not authentic or originary, but appropriate. What could possibly work is an attention to possession as momentary cognizance, a momentary gift of abnormal vision that could help describe the dominant in terms different than its own, as also point to other possibilities" (Achuthan et al. 2007, p. vii).

The new mantra under late capitalism: *psychologize*. Psychologize all kinds of behavior and mental states that are not in tune with the requirements of a corporate culture; psychologize all kinds of non-standard deviations; psychologize anything out of sync with the neoliberal ethos; psychologize all kinds of existences that lack 'killer instinct'; psychologize lack of competitiveness, or efficiency; psychologize any kind of attention-deficit in the child; and when the child or the adult wilts under an incessant pressure to perform, when one fails to deliver or live up to the efficiency index, and to expectations, psychologize even the breakdown; call it 'stress' and therapeutize in turn the breakdown, the vulnerability. The trope of one's own 'patienthood', deployed by post-Freudians acquires new meaning and power under "banalised versions of programmes of mental and moral improvement designed to get people back to work and ensure that they are docile" (Parker 2011, p. 7).

In other words, the history of unreason is no more the *mere* history of either absence-*ing*, or stigmatization–pathologization. It is no more a simple history of the ghettoization of unreason. It is no more a history of the 'putting outside' or the putting aside of unreason. It is no more the history of putting into the asylum or prison or reformatory any trace of unreason in the social order. Modern European man had reduced unreason to first the '*lacking other* of reason' and then to florid madness; the history of unreason was thus largely the history of madness. The late modern man, the globalized man treats unreason at home, at school, at workplace, and only at times in the clinic; he treats unreason even over telephone. He treats unreason not as madness. He does not relegate unreason to the asylum or the ghetto. He does not forget unreason. He *remembers* unreason every moment, at every site: home, family, school, college, and workplace. He, in actuality, never lets the leash off unreason. *Counsel*: counsel the single for their singlehood, counsel the couple for their coupledom, counsel the married, counsel the unmarried, counsel the sexual for being sexual, counsel the asexual for being asexual, counsel the rearing mother, counsel the rear-unhappy mother, counsel the child reared in school, counsel the teenager for being erratic, counsel the student for absenteeism and low grades, counsel the worker for being discontent and breakdown, and counsel the old for being old. Enter the bedroom, the inner world of every expression or irruption of unreason. Enter its bedrock; its bed. Do not drag it out. Do not throw it out of the social. Enter instead its private world. Do not exclude it. Include it instead. Exclusion is a thing of the past. Inclusion is the new avatar of hegemony. Not inclusion per se, but exclusion by inclusion. It is through incessant remembering, and an incitement to discourse around unreason that it forgets the experience of unreason.

Unreason is thus not just reduced to madness anymore. Its place is taken by an elaborate nosology of 'substitute signifiers' that "map the field of human distress [or suffering] as it comes to be organized in Western culture … [and] that is then populated by subjects outside the West" (see Parker 2011, pp. 41–51); and in the process the experience of unreason, its suffering is lost sight of. Its place is taken by a host of 'diagnostic categories': hysteria, anorexia, depression, mania, obsession, perversion, psychosis, schizophrenia, anxiety, and stress; each with a symptomatology, pharmacopeia, and psychotherapeutics. What happens in the process is that the patient's experience (of suffering), in a word, the experience of unreason is

chiseled out in terms of the theoretical framework that is at work or is being deployed; the therapist's experience of the patient's experience of suffering or of unreason is not allowed to bear upon the theoretical framework of the mental health professional. A procrustean act[4] takes shape: unreason is chiseled according to the theoretical framework; theory is not chiseled according to experience: the training of mental health professionals, "the transmission of methods, models and concepts from one generation to the next, and, perhaps most importantly, the sense of the reliability of the method from one analysis to the next—all rely on a kind of *repetition*, rather than on a *making anew* in each instance",[5] rather than a *renewal* of the analytic or of theory (see Abraham and Torok 1994), in its relation with experience.

With the advent of the phase of globalization under neoliberal conditions, the relation with unreason becomes even more complex; and the complexity constitutes the political economy of mental health. Neoliberalism sets up two encounters with unreason, one through the standard of human capital and the other through the trope of rational conduct. Both contribute to the new definition of *homo economicus*; the two also intersect, reinforce, and compensate for one another. Let us first discuss human capital.

In the neoliberal frame, *homo economicus* is not fundamentally a person involved in exchange but is an *entrepreneur*—entrepreneur *of oneself with exchange value* (Foucault 2008, pp. 215–288). He can rationally choose, deploying cost–benefit analysis to allocate scant resources that satisfies self-defined objectives. He is an ability-machine who is an "entrepreneur of himself, *being for himself his own capital*, being for himself his own producer" (Foucault 2008, p. 226); where the human *is itself* capital, where capital is embodied not in goods and services but in the human itself; such that labor power of individual humans does not just generate capital—*the human is itself capital begetting capital*; education and training, skills, values, etc., are some of the constituents of this new kind of capital (Becker 1976, 1993). In this model, "wage laborers are no longer employees dependent on a company, but are *autonomous entrepreneurs* with full responsibility for their own investment decisions and endeavoring to produce surplus value; they are the entrepreneurs of themselves" (Lemke 2001a, b, p. 199). In this new imagination of the human, marked by a certain self-maximizing calculus of rationality, unreason emerges as especially a suspect register. More than suspect, it becomes a register that has to be perpetually taken note of and kept within wraps; i.e., kept in check for productivity and efficiency conditions to be constantly met; the value of labor power under neoliberal conditions depends not just on the mental health/sturdiness of the laborer, but on how well one guards oneself against unreason.

---

[4] The question of the procrustean act can also be raised in the context of the question of 'cultural difference': "Soltani et al. (2004) caution that 'diagnostic categories for common mental disorders, usually developed in Western countries, may have limited validity' in other parts of the world, and Patel et al. (2003) warn that variance in culture and health systems may affect key parameters such as rates of medication adherence" (Knapp et al. 2006).

[5] Laurence Johnson, internet article, "'I wish to dream' and Other Possible Effects of the Crypt".

Two conditions of failure must be emphasized in this context: (i) the inability of the human to compete in terms of productivity, efficiency, and corporate values become a signal of the failure of his embedded capital or of his ability to adequately create and cultivate capital, and (ii) the inability of the human to be rational in terms of a psyche that abides by, cultivates and unequivocally accepts the principle of cost-benefit to choose over competing ends. Failure of each on its own or in tandem is sufficient to signal a *lack* in the human. Squashed by certain logic of interiority (which transmutes the psyche along lines of rational conduct) and a certain need to acquire the necessary means to become appropriate human capital, *lacks* gesture toward cracks in human reason; as also a *fall* into the abyss of unreason. Here, there is every chance that 'professional failure' will merge with 'personal unreason'; such that the failure of the alignment of mental soundness becomes palpable, even to the concerned human.

In the corporate environment, where competition has become the new buzz word, one thus has to guard against all minor irruptions of unreason and ward off all traces of unreason. The involvement of psychologists (trained in 'industrial psychology' and 'organizational behavior') in human resource management is a case in point. The task of contemporary psychology is to optimize efficiency and see how to make the worker produce more. The mental health of the worker is directly related to the efficiency of the worker; the efficiency of the worker depends on the mental health of the worker. One can see from the above what an appendage to late capitalism psychiatry-psychology is or could be (see Parker 2011).

Neoliberalism's economic perspective presents the problem as more acute with respect to developing countries since they are considered a repository of both lacking subjects and lacking structures; because of the mindset of the masses that is not appreciative of rational conduct as the organizing principle of behavior and the social milieu that is hopelessly short of the infrastructure and environment necessary to create and cultivate human capital. If, in the case of corporate environment, the failure is individualized, then in the case of developing countries the failure is a result of backward structures and reactionary or immobile cultures. The assumption of entire civilizations being entrapped in unreason underscores the entire corpus of mainstream development discourse which professes modernization; in fact, development discourse cannot be operationalized without a prior contempt for the 'psyche of the Other':

> Every intervention in the name of a civilization requires an initial contempt for the situation as a whole, including its victims. And this is why the reign of 'ethics' coincides, after decades of courageous critiques of colonialism and imperialism, with today's sordid self-satisfaction in the 'West', with the insistent argument according to which the misery of the Third World is the result of its own incompetence, its own inanity—in short, of its subhumanity (Badiou 2001, p. 13).

This aspect of an innate subhumanity of the other perhaps defines all work of neoliberal rhetoric and intervention, in which even well-meaning discourses on mental health get caught. Loss of economic value rather than actual mental states and the nature of suffering seem to be a matter of concern; as if any discussion on

mental states must take into account the aspect of resource, efficiency, and loss in economic value. The following quote is one example of such unconscious complicity:

> The sizeable and growing public health burden imposed by mental health problems across the world has been well documented (Murray and Lopez 1996; Whiteford et al. 2001; WHO 2001a), as have low treatment rates (Kohn et al. 2003). Mental health problems have considerable negative consequences for quality of life, and in many countries, particularly low- and middle-income countries, they also contribute to continued economic burden and sub-optimal productivity at the individual and national levels, through their reinforcing relationships with poverty (Saraceno and Barbui 1997; Bir and Frank 2001). Moreover, some of the most serious mental health disorders have their onset in early adulthood when people might be expected to be at their most economically active (Ustun 1999). Tackling this public health burden is a global challenge, for mental health systems in many countries remain seriously under-resourced and under-developed (Tausig and Subedi 1997; WHO 2001a).

The quote shows how mental disease is looked at under neoliberal conditions. The discourse is one of "economic burden", "sub-optimal productivity at the individual and national levels", and economic inactivity in early adulthood.[6] As if, mental disease is not about individual suffering,[7] but is about loss of productivity (and by default about efficiency and competitiveness). There is thus a shift in the discourse around mental disease. From a more 'moral problem' in late medieval European Christianity (where mental disease was seen as the curse of God, or the result of sinful thought or activity; where the mentally diseased was a 'moral monster'; see Foucault's *Abnormal* and *History of Madness*; where the ideational proximity of leprosy and mental disease is noteworthy, albeit without the perspective of contagion), there was first a shift in the eighteenth and nineteenth century to a more 'socio-political problem' (where the mentally diseased was seen as *threat* to society and to the peace of the polity and was to be put behind the more 'medical bars' of the asylum). There is now a further shift to mental disease as an 'economic problem' and this shift marks, nay dents, the political economy of mental health. Treatment regimes are thus moving from the protracted or the long-drawn (exemplified best by psychoanalysis) to a combination of pills and the magic pill: cognitive behavior therapy. The combination offers short term and quick relief to the mentally diseased and brings them back as soon as possible to surplus generating activities. Optimization of human resource is here the broad role of psychiatry reflecting in turn a certain psychiatrization of human experience. This is, of course, not to say that mental disease is not seen as a moral or a

---

[6] This is one of the reasons why health systems in low-income countries are poorly equipped to meet the needs of the pediatric, the geriatric, and the adolescent population.

[7] Which is also why we shall deploy suffering as trope for understanding the vicissitudes and complex predicaments of mental disease in India; in that sense, we have a twofold critique of 'mainstream mental health': one, premised on the 'lack of access' and the other, premised on the 'nature of access'; we thus emphasize, on the one hand, the need to attend to suffering, and on the other, the need to attend to it in a manner that does not pathologize suffering or make attention an 'adaptationist enterprise'.

sociopolitical problem anymore. All three—the moral, the sociopolitical, and the economic—mutually constitute the political economy of mental health in India.

There is, however, another discourse, in the same quote: the discourse of "low treatment rates" (which is about how mental health systems have remained "seriously under-resourced and under-developed" in our country) and "negative consequences for quality of life" (which could be translated into the language of inner suffering). In this paper, we thus develop a double argument, an argument around an inadequate attention to suffering in interpersonal, family, community, and governmental contexts *and* an argument around the nature of attention to suffering, attention in terms of pathologization and adaptationism. This is important in the context of the finding that almost 90 % of global health expenditure occurs in high-income countries which have only 16 % of the world population (Schieber and Maeda 1997). This disparity in expenditure is likely to be worse in the area of mental health because one-third of countries do not have a specific mental health budget, while 36 % of those countries that do, allocate less than 1 % of their public health budget to mental health (WHO 2001b). However, while these are national figures there are nations within nations, and colonies within erstwhile colonies: even the "new nationalism … encapsulated by the economic utopia of 'inclusive development' … is marked by a fundamental split [in] India's national cartography into the 'expanding circuits of global capital' and its *constitutive outside*: 'world of the third'" (Chakrabarti and Dhar 2012, p. 106). Which is why while scarcity of resources is endemic, it is not equally experienced. Resource scarcity takes many forms, and interferes with the relationship between funding, services, processes and user and societal outcomes. There are thus economic barriers in all health care systems. Knapp et al. (2006) refer to them generically as *resource barriers* because their impacts are most commonly and most immediately felt at the resource end of the care or treatment process. Knapp et al. (2006) discuss those under six headings:

(i) Information Barriers

   (a) like generalizing the evaluation results of a service intervention or a clinical research in the west to the mental health system of India is simply not sensible without a lot of qualifying assumptions
   (b) like the lack of epidemiological data limits the ability of decision-makers to build a framework to allocate and distribute resources according to need; Sarin and Jain (2012, p. 36) have tried to argue that the omission of primary data about insanity in the census from 1941, and the lack of epidemiological data have handicapped adequate state planning for care of the mentally diseased in India.

(ii) Resource Insufficiency

   (a) Recent estimates put the percentage of the total health budget spent on mental health as 1.5 % in low-income countries, 2.8 % in middle-income countries, and 6.9 % in high-income countries (WHO 2003b)

(b) WHO [*Investing in Mental Health*] points out that a global median of 2 % of health care budgets is spent on neuropsychiatric disorders that account for 13 % of the global disease burden (WHO 2003b)
(c) even in a well-resourced mental health system such as Australia's, only one-third of all people with a mental disorder consult for treatment (Andrews et al. 2001)
(d) in the US, '30 % to 50 % of adult primary care patients with depression do not have their condition recognized or treatment initiated' (Wells et al. 2002, p. 658)
(e) surveys by the WHO World Mental Health Survey Consortium (2004) found that 36–50 % of people with serious mental illness in 'developed countries' and 70–85 % in 'less developed countries' had received no treatment in the previous 12 months (see also Kohn et al. 2003).

(iii) Resource (Mal)distribution

(a) Like concentration of mental health services and professionals in urban areas
(b) highly institutionalized, at times, even asylum-centric services such that community mental health gets neglected: developments in Cuba illustrate how mental health services can evolve, making the transition from an asylum or hospital model in the mid 1980s to the promotion of a community mental health model that is integrated with primary care in the 1990s (WHO 2003a)
(c) neglect of particular disorders, like diseases of women and subalterns as also mental retardation.[8]

(iv) Resource Inappropriateness (like over-investment in expensive technology)
(v) Resource Inflexibility (like centralized planning and funding[9]) and
(vi) Resource Timing in the context of 'clinical conservatism', unaffordability of newer drugs (Srinivasa Murthy et al. 2005) (also funding in the context of a general lack of mental health professionals,[10] and the occasion of the transfer of resources from bed-based services to community services).

[8] In some countries, there is debate about whether learning or intellectual disability ('mental retardation' in North American parlance) should be included under mental health.

[9] The highly centralized mental health services of the former Soviet republics were of poor quality, and the monopoly of the Snezhensky school for the training of psychiatrists provided the preconditions for the widespread abuse of psychiatry associated with the Soviet regime (Polubinskaya 2000 as quoted by Knapp et al. 2006).

[10] The paucity of skilled mental health practitioners in low-income countries constitutes a severe resource limitation, a problem exacerbated by the migration of skilled professionals to countries offering better salaries and quality of life. India has a very limited numbers of mental health facilities and professionals (one bed per 40,000 population and three psychiatrists per million).

One therefore has to disaggregate the discourse and experience of scarcity. Scarcity in the context of whole Southern nations as against North America, Western Europe, and Australia is one kind of scarcity. The other kind of scarcity is scarcity within; scarcity in Dark Continents within the nation-state [we have reconceptualized them as 'world of the third' as against the given discourse of third world nations and third worldisms (see Chakrabarti and Dhar 2009, 2012a, b)]. Also there is no reason to believe that there is no space or scope of scarcity having differential social presence within the North–Western part of the globe; pockets of poverty exemplify this difference. The division between the affording populace and a struggling populace, symbolizing two different worlds of scarcity, is however sharper in countries such as India, which is why the question of the political economy of mental health in India needs to be split into at least two kinds of conceptual-spatial registers. One, the register of primarily the elite in urban locations, where flows of global capital are in tune with the flows of newer discourses on health and novel health technologies, both in their production and consumption capacity; in other words, registers within the ambit of the 'circuits of global capital'. In contrast, the narrative in the 'world of the third'—in what could be called the constitutive outside of the circuits of global capital—would be different. Let us constitute for the purpose of this paper a somewhat amorphous and inclusive outside, an outside constitutive of more sharply the world of the third, but more generally of spaces marked as rural, as tribal, as subaltern, and as urban at the margins of urban hubs, all of these 'distance away' from urban hubs. Such spaces continue to languish in extreme scarcity. Or perhaps continue to escape the gaze of global capital and western medicine. It is, however, difficult to choose between languish (which is more existential) and escape (which is more epistemological), because both are extremely powerful suggestions, and one sharply contradicts the other. It would perhaps be more balanced to say that world of the third, and more generally the amorphous and inclusive outside generated above continue to mark *differance*[11] with respect to the overdetermined language-logic-ethos of global capital and western medicine. However, one will still have to see what happens 'there'? There the experience of unreason is both stigmatized and attended to.

(a) There, at a more general level, when they manage to attend to unreason, the state/public health institutions continue to pathologize unreason. However, state/public institutions lack mental health services, which is why the attention is minimal. Even if they have psychiatrists, they do not have psychologists. Even if they have psychologists, they do not have psychiatric medicines. There unreason is reduced to madness and madness is not taken care of. The cure or the care of the mentally diseased is thus negligible. This is, of course, not to

[11] *Differance* represents in one turn the moment of *difference-differing-deferring*. It represents a passive difference already in place as the condition of signification, and an active act of differing, which produces-introduces differences as also deferral as silent, secret, and discreet suspension.

say that madness is not taken care of in just secondary care centers (i.e., district hospitals). Madness is not taken care of in urban/tertiary centers as well. Even at tertiary mental health spaces madness is still an object of ridicule, hatred, and violence.[12]

(b) The above are, however, generalities. From a political economy perspective, what is of crucial importance is "that the idea of universal curative services at a primary level was gradually relinquished" (Zachariah et al. 2010, p. 11) by the Indian state in the 1950s. The withdrawal of the Indian state from 'curative function' at the level of primary health care (and the shifting focus of medical education on tertiary care careers) meant that mental disease along with other illnesses were no more attended to, both in rural areas and at the margins of urban spaces.

(c) The community, on the other hand, at times stigmatizes unreason and puts it outside or puts it at the margins of everyday life; at other times, it is (social) cause of mental disease. Mahasweta Devi's short story *Banyen* is a noteworthy explication of the communal exclusion of unreason by rural communities; however, the nature of the stigmatization is different; the trope of witchcraft and witchhood or evilness is deployed here.

(d) At other times, the community is seen to care for the mentally diseased. The tireless work of *Iswar Sankalpa*—an NGO working with homeless mentally ill women—has shown that some members of the community do indeed take care for the mentally diseased. This is why it is difficult to either defend or criticize the community for the relation it sets up with unreason.

(e) The family and community in India also takes recourse to what is usually called sites of 'faith healing' or 'healing by faith'. Faith healing sites have a political economy different from modern institutions, either state or private. Such sites are usually owned by temple trusts (or trusts in charge of *dargas*) or by individual priests. Around the temple or the *dargah,* a 'local market' takes shape which provides offering-related services to the mentally diseased and their families.

(f) And last, but not the least, would be the ones who have neither family nor community, neither state nor the private sector to look after them. Abandoned by all kinds of familial, social, and institutional support networks, they form an unknown and nebulous mass. Abandoned even by the National Census, they

[12] The National Human Rights Commission in India investigated 37 public mental hospitals (which are actually large psychiatric asylums, often colonial relics, starved of funding and decaying since independence) that housed 18,000 people. People were frequently required to sleep on the floor, and in some male wards people were required to urinate and defecate in an open drain. Water availability was a problem in over 70% of the hospitals, resulting in a lack of safe drinking water and washing facilities (National Human Rights Commission of India 1999). The work of *Anjali*, Kolkata and *BapuTrust*, Pune on the rights of the 'mentally ill' has further demonstrated the somewhat difficult processes through which the mentally dis-eased go through in state/public health institutions.

have come to be known in NGO circles as the 'homeless mentally ill'.[13] They are the ones who have neither home nor remand home; they roam the streets of cities or the fields of rural areas. They live nomadic lives and remain perpetually open to the vagaries of weather gods and male sexual predators. They are the nomadic embodiments, perhaps victims also, of an insensitive and selfish social and State formation. One could also see them as the *outside* (and not the margin) of a more formal framework of political economy.

The above only shows that there is no *one* political economy of mental health in India. The way a state mental health system functions is quite different from the way a private institution functions. The way a modern mental health institution functions and the way a faith healing site functions are also different. The way institutions function and the way the community functions are also different. The way a State institution functions and the way an NGO functions are different. The way an NGO hooked to the circuits of global capital functions and the way an NGO driven by a more 'cooperative community ethic' functions is also different. *Each requires separate and detailed ethnographies.* It would also not be incorrect to say that the state/public institutions and the community set up two or more than two kinds of relations with unreason, outside the circuits of global capital. These differences need to be taken into account when one is discussing the political economy of mental health.

Let us interrogate further the question of the political economy of mental health by looking at mental health not from the perspective of disease and cure/care of the diseased; but from the perspective of health and well-being. While all of the above discussion is about individual suffering and mental disease, one will also have to take into account the sheer psychologization of the capacity to labor and produce under neoliberal conditions. While the first is about state initiatives, about discourses on disease and treatment, about the hegemony of drug industries over psychiatrists, the second is about the vitality of labor. The first is about service to or for the mentally diseased. The second is about labor's capacity to serve, to serve efficiently and competitively, so as to generate maximum surplus.

Political economy of mental health thus cuts in two ways: on the one hand, it is about the politics and economics of mental health services; on the other, it is about the 'mental health of labor', the psychic resilience of labor (or of the service provider or of the performer) to incessant and mounting stress and strain. How 'tough' the IT sector worker is. How 'sturdy' the informal sector worker is, is increasingly becoming a mental health question. To understand further this double-discourse of the political economy of mental health in India, we shall take two more steps. The first would be about shifting perspectives to medicine and mental health in post-independence India. The second would be about the many contradictory relations the contemporary is setting up with unreason.

[13] "They, these mentally ill vagrants are homeless in the registers of mental hospitals but may well overlap with the 'missing woman' records published in newspapers daily. One can imagine that someone is waiting for her somewhere but tragically also imagine that someone is glad to see that she has wandered away into oblivion herself" (Lacroix and Siddiqui 2012).

History: Welfare Medicine to Development Medicine

The Bhore Committee Report was formulated at the zenith of the freedom movement in 1946 as blueprint for right to health and *welfare medicine* in India. The tragedy is that the Bhore Committee Report was never fully implemented. The welfare model of the Bhore Committee Report was in fact replaced in practice after the 1950s by what Zachariah et al. (2010, pp. 10–11) call *development medicine*. "The primary symptom of this change was that the idea of universal curative services at a primary level [meaning Primary Health Centers] was gradually relinquished[14] … in this new modality the issue of health security was addressed like economic growth – or modernization more generally – as something to be planned centrally by experts and policy makers in the long term interests 'of the nation'" (Zachariah et al. 2010, p. 11) as also to maximize economic potential and add to the "vitality of labor". Mental health, however, did not feature in either model. What featured as public health initiatives were state-sponsored mass campaigns around select health risks: malaria, tuberculosis, polio, etc. and now HIV-AIDS. Primary care or care organized around Primary Health Centers became top-down preventive health campaigns organized along the triple axes: (biological) threat, risk (of contagion), and (corporeal) security. Later, the Medical Council of India and the 1977 Srivastava Committee had called for curricular changes that shift the focus to primary care and thus reorient medical education toward community needs. However, there have been few takers for such recommendations. Even the state-sponsored medical education (along with higher education in general) is geared toward promoting tertiary care careers in cities for a very small segment of the Indian population. Medical education thus does not just create doctors, but doctors of a particular kind, doctors suited to become 'local guardians' of the circuits of global capital and the commoditization of (mental) health. The present state of neoliberal health policy argues that barring the treatment of cases like HIV, tuberculosis, swine flu etc. (which are a potential threat to public life),

[14] Iran and Thailand are two countries promoting mental health in primary care settings. WHO's Mental Health Atlas (2005) says that, as far as community care for mental health is concerned, India and south eastern Asia lag behind the rest of the world. We, however, remain a little skeptical of the process of making the primary health center also a mental health care unit. While one can never oppose or detest the possibility of the delivery of mental health care at any level, be it primary, secondary, or tertiary, given the general dearth of mental health services in the country, one still remains wary of the potential slippage of what is thought to be the expansion and deepening of care (and cure) to a paradoxical expansion and deepening of control and surveillance. We hope the District Mental Health programme (DMHP) and mental health care at the primary level indeed becomes care (and cure) and an alleviation of suffering; and not a psychologization and a consequent pathologization of (rural) life hitherto untouched by the mental health apparatus. We hope community mental health indeed becomes mental health for, of, and by the community and not community outposts or microversions of tertiary mental health institutions in rural areas. One thus remains wary of the possibility of what is increasingly becoming a pattern in urban areas or areas hooked to the circuits of global capital, becoming even a minor pattern in rural areas or areas outside the circuits of global capital.

"curative (or clinical) services should be procured in the market. … it [thus] trims public health into a series of centralized, single axis initiatives, and … it has no fundamental quarrel with the costs of the gold standard medical treatment. It accepts that such treatment and cure should be accessible to those who can pay, while advocating health education and prevention for the many who cannot. Given its market focus, neoliberalism is necessarily indifferent to the failure of primary health care in India." (Zachariah et al. 2010, p. 13)

The 'nature of financing' of the mental health system is demonstrative of the above. In India, health financing is undertaken by a number of sources: (i) the tax-based public sector that comprises local, State, and Central Governments, in addition to numerous autonomous public sector bodies; (ii) the private sector including the not-for-profit sector, organizing and financing, directly or through insurance, the health care of their employees and target populations; (iii) households through out-of-pocket expenditures, including user fees paid in public facilities; and (iv) medical insurance. Nearly, three-fourths of the total source of health expenditure was out-of-pocket from households. The public funds accounted for only about 20 %. The private funds from the companies accounted for 5.3 % and that of NGOs was 0.1 %. Almost 60 % of hospitalization and four out of every five short-duration ailments are treated in the private sector. Public spending on health as a share of GDP is marginally higher than 1 % as opposed to National Common Minimum Programme (NCMP) health expenditure of 3 % of GDP. This is much less than developing country average (2.5 % of GDP) or the world average (see Fig. 25.1).

Given the history of the movement of the Indian state from welfare to development to neoliberal medicine, let us now take a closer look at the nature of the health budget of the country. Public spending on health in India (i.e., spending by only Central and State health departments) gradually accelerated from 0.22 % in 1950–1951 to 1.05 % during the mid-1980s, and stagnated at around 0.9 % of the GDP during the later years. Of this, share of recurring expenditures such as salaries

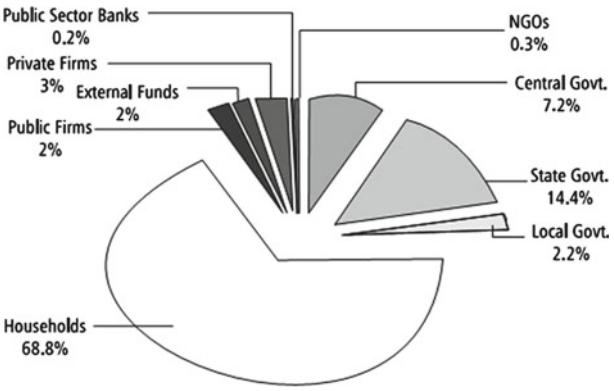


Fig. 25.1 Sources of finance in the health sector in India during 2001–2002 *Source* MHFW 2005, p. 240

and wages, drugs, consumables, etc. account for more than 90 % and is on the rise in recent years. On the one hand, in terms of per capita expenditure, it increased significantly from less than Re 1 in 1950–1951 to about Rs 215 in 2003–2004. However, in real terms, for 2003–2004, this is around Rs 120. On the other hand, under a system dominated by out-of-pocket expenditures, the poor, who have the greater probability of falling ill due to poor nutrition, unhealthy living conditions, etc.,[15] pay disproportionately more on health than the rich and access to health care is dependent on their ability to pay; and when they cannot pay they fall outside the health apparatus altogether. And outside, there is no alternative on which they can fall back. WHO's Mental Health Atlas (2005) states that, as far as community care for mental health is concerned, India lags behind the rest of the world. Political economy of mental health in India is thus split into two:

(i) One is within the circuits of global capital, where the imperative is to be part of the incitement to discourses around psychologization-neurologization (which exempts the overdetermined triad of the economic-political-cultural of critique and makes all problems individualized in either an intrapsychic sense or an intraneuronal sense), where one has to be part of the commoditization of mental health, where one has to see the solution in either paid counseling-psychotherapy or market-based pharmacotherapy.
(ii) The other is outside the circuits of global capital; where one can already be outside the circuits; or one can fall outside the circuits; fall outside because one does not have the means or resources. Outside there are three options:
(a) The first is to be part of a secondary or tertiary care service, but with a begging bowl.
(b) The second is to be part of either non-allopathic western services: homeopathy, or non-western services altogether: ayurveda, unani etc.; services like homeopathy and ayurveda can, however, also be accessible within and in terms of the circuits of global capital now; the *medicalization* of ayurveda through mass production of drugs and the context-insensitive nature of current ayurvedic therapy is an example of how services hitherto outside the circuits of global capital could become a component part of the circuits.
(c) The third is to be part of services offered by sites of faith healing; faith healing sites, however, can also become hooked to the circuits of global capital.

Each of the above has a political economy specific to itself; but each is different from the one within the circuits of global capital. There is, however, a fourth option; which is to have no option. The mentally diseased in India at times fall into an option-less option. The category 'homeless mentally ill' or the "missing mad"

[15] See Zachariah et al. (2010, p. 188) for the new epidemics of the poor: "the poor are also affected by cardiovascular diseases—and that problem is a different one and of greater magnitude. Among the poor, IHD is a result of migration, changes in diet … smoking and urban stress."

is such an option-less category, where even the political economy perspective comes to a standstill, where unreason in its most disheveled form stares back at political economy, and renders its inner reason, its inner logic most vulnerable.[16]

Unreason in Contemporary India: Economization of Mental Health

Mental health in contemporary India is haunted by two sets of questions: one is coming from the 'present' and the other is coming from the 'past'. One is coming from the 'west' and the other is coming from the 'east'. The 'present' is about questions the contemporary is throwing up (questions coming from neurobiology, cognition studies, disability studies, critical legal studies, critical psychology, post-Freudian approaches, qualitative/quantitative debates, public health/private health service related debates, the State-pharmaceutical continuum, incitement to discourse around mental states [like stress, anorexia, and ADHD in a post-globalization world], deinstitutionalization, depathologization, informed consent, living will, ECT, client's perspective, and bioethics etc.); and all these are affecting the political economy of mental health. The 'past' is about questions the 'Indian past and the Indian context' is throwing up (questions pertaining to 'faith healing' and nonwestern approaches to mental health which still thrive in India, debates around '*Indian* mental health', debates around 'medical pluralism' and the possible plurality of approaches to mental health (Sujatha and Abraham 2012), and debates around institutional and community mental health). The question of the 'political economy of mental health' shall need to be taken up in the overdetermined context of the above sets of concerns. However, as soon as one attempts to do so, the question of the political economy of mental health soon splits into two and the paper turns bifocal. On the one hand, one realizes the necessity of taking a close look at the political economy of present day mental health service—spanning private forms of mental health delivery, contemporary biomedical psychiatry and the pharmaceutical industry, and the burgeoning need for psychotherapy and counseling in school, university, and industrial settings. On the other, one needs to see how a certain characterization of mental health (marked by competitiveness, efficiency, aggression, confident body language, articulateness etc.—terms incumbent upon a cricketing or IPL-like context) serves the emergent neoliberal kind of economic imagination. One can ask in this context, what kind of mental health in the labor force does the current economic organization require? How would it be ensured that the current labor force has *sound* mental health

[16] Since the introduction of the National Rural Health Mission (NRHM), some efforts of government to strengthen the public health system in the rural area are observed. However, will it indeed reach the rural poor (apart from the limited space of maternal and child health)?

(unsoundness of mind has been the dominant trope for defining the mad)? How to ensure that labor force sees itself as personification of the rational man and of human capital? Who would ensure it? Marx writes in *Capital*, Volume 1:

> It is plain that commodities cannot go to market and make exchanges of their own account. We must, therefore, have recourse to their guardians, who are also their owners. Commodities are things, and therefore without power of resistance against man. If they are wanting in docility he can use force; in other words, he can take possession of them. In order that these objects may enter into relation with each other as commodities, their guardians must place themselves in relation to one another, as persons whose will resides in those objects, and must behave in such a way that each does not appropriate the commodity of the other, and part with his own, except by means of an act done by mutual consent. They must therefore, mutually recognize in each other the rights of private proprietors. This juridical relation, which thus expresses itself in a contract, whether such contract be part of a developed legal system or not, is a relation between two wills, and is but the reflex of the real economic relation between the two. It is this economic relation that determines the subject-matter comprised in each such juridical act. (www.marxists.org)

Commodities cannot go to the market and make exchanges of their own account. We must, therefore, have 'guardians'. The psychiatric establishment has emerged as the undisputed guardian of both the mentally diseased and the treatment regime in the latter half of the twentieth century. If economics can objectify productivity and efficiency *a la* human capital, then psychiatry appears as a discipline that can objectify (un)soundness of mind. In the domain of the economic production of goods and services, especially in a corporate set up, the twin objectifications are aligned such that (un)soundness of mind is tested and analyzed in terms of the standard of productivity and efficiency. Economic performance and mental performance gets intimately connected; not just the customization of the product, but the customization of the mind too is an important marker of sound economic culture today. This is not to say that customization of mind denies the possibility of differential economic performance (some doing better than others). Rather, what customization of mind means is the universal acceptance of the norm or standard of what economic performance is and how it is to be judged (by measures and markers of productivity and efficiency). Such that it is not just the human resource managers inside the enterprise who can identify a case of failure, but where even the failed human can see and will be willing to accept and internalize failure and subsequently surrender his or her mind–body continuum to a rectification experiment comprising of drugs, counseling, and cognitive behavior therapy. Here, only the individual is to be taken to task for failure, not the larger social; not the conditions which drive him or her to fail, but the mind that, as if, fails to surmount the conditions. This makes the mind the *object interminable* of self-assessment and self-correction. One could call this the reinvention of the mind post-Descartes; where the mind is not a 'mirror of nature' (Rorty 1979, p. 357), but a 'mirror of human capital'. This also creates conditions for the larger psychiatrization of human resource or labor power in late capitalism.

Under the guardianship of psychiatry, mental health service has been subject to a variety of seemingly disparate trends: the neurobiologization of the knowledge of mind and mental health, the medicalization of clinical practice and research, the privatization of services, the fractionation of virtually all needs and activities of the severely mentally ill and the transformation of care into an array of commodities to meet these needs, the increased use of governmental controls such as involuntary outpatient commitment and limitations on rights and entitlements, and the intensification in interprofessional rivalries (the devaluation of psychology, psychotherapy and especially psychoanalysis by psychiatrists is only one example). In a word, the connection between knowledge and economy (and not just between knowledge and power) is increasingly becoming intimate. From the perspective of political economy, many of the transformations within psychiatry (and psychotherapy) can also be understood as reflecting fundamental changes in corporate organization and profitability.

Parallel with the above is the shrinking of the 'safety net' with declines in the real value of welfare exercises, elimination of State support, and attacks on the utility of welfare. Support for public mental health programs has begun to decline precipitously. Increased privatization of psychiatric services, public contracting of private services, transformation of services into commodities, and marked growth of the pharmaceutical industry have only made mental health services increasingly unavailable to the vulnerable and the poor, found mostly among the informal sectors workers and the distant rural workforce.

A further consequence of these economic upheavals has been the mass of unemployed and underemployed including fellow humans caught in zones of displacement-dislocation and (illegal) migration, making it difficult for more marginal workers like those with severe mental illness to obtain work and become tax-paying citizens. Paralleling shifts in discourses on social welfare, and curtailment in public funding for mental health, was the increased receptivity to a new ideology of mental illness that focused exclusively on internal biological elements rather than on broader or holistic social factors (it not only narrowed the approach toward mental suffering—a narrowing consonant with the needs of private capital—but also converted 'social suffering' to individual pathology). This biologization of mental illness was further fueled by the expansion of the pharmaceutical industry, which since the 1980s has enjoyed extremely high profits and has heavily invested in sponsoring symposia, journals, and other promotional products, as also psychiatrists. Thus, the changes wrought by alterations in public funding and ideology about social welfare and mental illness has had a profound impact on persons who were both indigent and mentally diseased, such as those with schizophrenia and other chronic disorders. Deinstitutionalization, paradoxically, provided the initial impetus for states to transfer public services to the private sector. Indeed, coupled with the growth of private insurance plans, there was a mushrooming of proprietary psychiatric hospitals and use of psychiatric units in general hospitals. Increased privatization of public services is also fostered by investors' needs to find new sites for capital, as neoliberalism expands the scope of the economization of society. These trends have had several effects.

The transformation into commodities of various aspects of everyday life and the commoditization of mental health service/care has created a greater potential for economic and social control of the mentally diseased. Privatization has also meant that mental health services are undergoing the same cost-modeling and rationalizing pressures as other businesses. One can ask: what happens, due to all of the above, to the *ethical economy* of mental health service? What happens to the right of the sufferer to receive treatment? Historically, public hospitals and clinics have been part of the mechanisms by which all persons are assured cure/care. What happens to that 'cure/care function'? What happens to the 'care of suffering' under neoliberal conditions? In the context of the question of care of suffering, the paper shall also need to ask: what happens in the backyard of modern western mental health services? What happens in faith healing sites? What is the political economy of such healing sites? Such sites are also going through unthought of economic and cultural transitions that are putting them in tune with the profit-making ethos.

Rethinking Mental Health Through Suffering

> In India, mental health law is a kind of net to catch and institutionalise people perceived to be of 'unsound mind', falling through various incapacity laws. So actions made on incapacity, lead to further actions on deprivation of liberty, through incarceration. Mental health law serves many vested interests in India, and private agencies find it profitable to simultaneously exercise legal incapacity law and mental health law to haul people into a private or public institution. With every divorce filed on insanity petition, at least 1 forced institutionalisation and a few shock treatments, along with the karmic label of 'schizophrenia', comes free! This way, both lawyers and doctors in India make much money.
>
> Bhargavi Davar (2012)

Given the experiences of pathologization and incarceration, how about rethinking mental health from the perspective of 'suffering' and 'care and healing'? This changes the framework. Without this change in the framework, there can be not substantial changes in the political economy of mental health. This is of course not to say that the triad of suffering-care-healing is the only conduit through which mental health can be revisited. There can be other, perhaps more enabling perspectives. Ours is, at best, a *partial perspective*.

Suffering can be of many kinds. Individuals can suffer. Entire cultures/societies can suffer (Tibetans suffering under the Chinese Communist Party is a poignant example of entire cultures suffering in exile). Consequently, attention to suffering can also be of many kinds. One may attend to the individual sufferer or to a suffering culture/social as such. One may attend to the cultural/social sea of suffering or to puddles of it. One can also say that attention to puddles (in the form of institutional/clinical attention) is never enough. One need to attend to the sea of suffering and this is not just a quantitative argument. It is not just about attending

to a larger pool instead of micropiles. It is also a qualitative argument; because for some the puddles are a *symptom* of the sea of suffering. For yet others, the puddles are a by-product of existing social processes. In this argument, the nature or the structure of the social is the *cause* for suffering puddles. This turns the table on the social. The social, and not the individual, becomes the object of interminable critique.

Mental health in India has not managed to attend to the social as cause of suffering; this inattention to the social has been secured and bolstered by a rampaging neoliberal political philosophy that has generally displaced social problems into an individualist metaphysics. Preventive mental health care has thus never been a preoccupation of the Indian mental health apparatus. Also, the larger social, comprising primarily of the rural has remained outside the purview of the Indian mental health apparatus. This is because there is no mental health component in the services the primary health centers provide in India. Mental health has never been a concern of the primary care givers. Mental health has more or less been a primarily tertiary care level issue. Neither the rural nor the poor have been subjects of attention. To be subjects of attention, one has had to travel miles with job loss on days when one comes to tertiary care centers like say Medical College, Kolkata or Ranchi or NIMHANS. Sites of faith healing (and occasionally mediocre private psychiatric service) remain to this day the only option available to the mentally diseased from poor rural backgrounds.

The Indian mental health system is thus primarily urban centered and is strictly individualistic (even family circumstances or dynamic are not considerations). The pharmaceutical industry focuses on only the individual sufferer; the perspective of the social is ignored. It has to be ignored for the mass production and consumption of long-term psychiatric drugs to continue unabated.

Further, suffering can be understood in biological terms, it can also be understood as a product of culture/social. One can attend to suffering through the medical model; one can attend to suffering through a nonmedical model. In the medical model, one can attend to suffering through the pharmaceutical route; but also 'through words'—through 'language' (for example psychotherapy, talking cure, counseling). In all these approaches, the problem is located in the individual, more specifically, in the biology of the individual as understood by pharmaceutical psychiatry and in the growing up or development of the individual as interpreted by psychotherapy; at most in the self, or in the telos of drives or the libidinal, or in object-relations, or in the syntonicity of the ego. Of course, none of these approaches deny the importance of the cultural/social. The somewhat unthinking reiteration of the 'bio-psycho-social' as the structure of causation has become fashionable these days in mainstream psychiatry and psychology. However, the acknowledgement of the importance of culture/social and the non-denial of the same are not the same. The acknowledgment by way of foregrounding the social (and the historical) will radically transform both the understanding of suffering and the method of care-healing, and the political economy of mental health.

Interestingly, the nonmedical model—a model not marked by the *birth of the clinic* (Foucault 2003), the anatomical body of *three-dimensional tissue spaces* (Foucault 2003, p. xviii), and the *asylum*—can also make use of drugs (however, the nature of use is different and the understanding of the mind–body complex in their respective imbrications and exclusivity are different). *Ayurveda*, with a distinctly 'context-sensitive epistemology', would be one example of the nonmedical approach to suffering. Faith-healing would be another example; where the use of the pharmaceutical approach would be tentative; instead, here the use of the language of suffering-faith-healing would need to be tracked. One would need to see how the economy functions in these two cases. Interestingly, Ayurveda is increasingly moving toward a western medical model. The mass production of Ayurvedic drugs stands counter to the context-sensitivity that was the difference of Ayurveda with 'universal medicine'.

One could also have a community approach to suffering and the community approach can again be subdivided into two. First, where the approach is one of developing an imagination of mental health and social well-being through certain learning from the community. In this approach, there is attention to the relation between reason and unreason in the community. The relation is revisited and remapped from a critical-cultural psychology perspective and it is in terms of this remapping that healing in the community context is reconceptualized. The other is where the medical approach does not restrict itself to the hospital-clinic complex, but makes its way into the public sphere and into community life. This approach that makes the medical-psychiatric model imperial (mini-hospitals are set up in the micropores of the community followed by *pathologization* and *psychiatrization*) has been put to critique by Bhargavi Davar of Bapu Trust, Pune, Ranjita Biswas of Iswar Sankalpa, Kolkata and the tradition of critical psychology in India. Bhargavi Davar (2012) discusses two models for community mental health, the medical model and the social model:

What is a community mental health model? There is the WHO given *medical model* of CMH, and a *social model*… [the] medical model (also followed by several NGOs in Nepal) kept to identification of disorder, providing medical service in community, and referral to hospital for tertiary care. A social model of care would be comprehensive, with strategies of

1. Understanding community needs (marginality, resilience) through research.
2. Assessments using psychosocial tools.
3. Developing mental health as a strategy of overall community development.
4. Provision of different types of community services (like community awareness, home-based psychosocial support, neighborhood alert systems to intervene in case of extreme states, such as victimization of people with disabilities).
5. Human resources to be sourced from community itself, to enhance ownership of the work and to leave behind capacitated people in the community with a good exit policy.

It is in the social model, that there is an attention to social suffering and community healing. While the above was about the nature of suffering and the possible approaches to alleviation, it will not be out of place to glean at the quantified extent of suffering; while this quantification does not in any way cover all facets of suffering that we have explicated here, the figures does give us some idea of the magnitude of the problem. The extent of suffering can be gauged from the following information culled out of the *Basic Needs Policy Study 2009: Mental Health Financing in India.* The prevalence of major mental and behavioral disorders at any given point of time in India was estimated at 65/1000 population in all ages and both sexes based on the average value of two pooled studies (Ganguli 2000). For the year 2001, an estimated 67 million people with major mental disorders, 20.5 million with Common Mental Disorders, and 10 million with alcohol dependency problems required services in India (Gururaj et al. 2005). The estimates in 2000 say that mental disorders accounted for 12.3 % of disability adjusted life years (DALY) and 31 % of years lived with disability. It also projected that the health burden due to mental disorders will increase to 15 % of DALY by 2020 (Murray and Lopez 1996). In 2008, a joint publication by India's National Human Rights Commission and the National Institute of Mental Health and Neuro Sciences noted that: "Morbidity on account of mental illness is set to overtake cardiovascular diseases as the single largest health risk in India by 2010". The preliminary findings of a WHO-supported multicentre study on mental health in India indicates that about 10 % of the population in India have mental health problems, according to a recent report by *The Lancet* (Chatterjee 2009).

The above description of mental health in India through the discourse of suffering and alleviation of suffering makes the remapping of the mental health sector/field in India in all of its complexity a necessity. Here is an attempt. The understanding of the political economy of mental health requires this cartography. The cartography of mental health in India is marked by the triad of medical models, community approaches and nonmedical perspectives to suffering-care-healing. The medical model can be further disaggregated into institutional sites (hospitals, public, and private), (private) clinics and public health interventions. In the medical model, the mode of treatment could be drug based. It could also be nonpharmaceutical. Community approaches could be split once again into medical models and social models. Mental health could be curative, could be preventive. It could be based at tertiary care centers. It could be located in the context of primary health care meridians. Nonmedical perspectives to suffering-care-healing could come from Indic, Arabic, African, Latin American, and Chinese sources. Ayurveda, *Unani-Tibbi*, Siddha, Acupuncture are only a few examples. It could also be based at sites of faith healing practices. The history of colonialism in the context of India complicates this already-complicated picture further.

The economic structure in conjunction with each of these mental health service systems and further with the disaggregated existence within each will lay bare a decentered and heterogeneous economy of mental health in India. While the conceptual production of this economy awaits further research, it can only be reasonable to make a conjecture that mental health service in India will be in

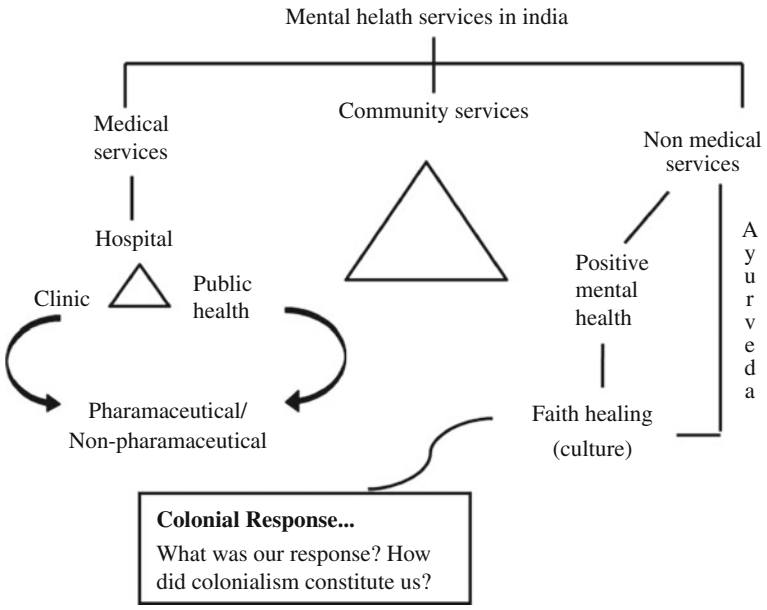


Fig. 25.2 Mental health services in India

articulation with various organizations of surplus—capitalist, communitic, and independent/individual based (see Fig. 25.2).

Limits of Political Economy: In Contexts of Unreason and Cultural Difference

We began with a Lev Vygotsky quote from the *The Historical Meaning of the Crisis in Psychology*: "psychology is in need of its own *Das Kapital*—its own concepts of class, basis, value etc.—in which it might express, describe, and study its object". If psychology is to be in search of its own *Das Kapital,* political economy is also to be in search of its own *History of Madness* (*a la* Foucault) and its own *Interpretation of Dreams* (*a la* Freud). The writing of the 'political economy of unreason' requires the received 'reason of political economy' to be put under erasure simultaneously. Thus, if one has to revisit mental health from a political economy perspective, one has to also revisit political economy from the perspective of the mental; the question of the 'mental' and of 'health' in political economy would perhaps open the way to another critique of political economy. We have not done that in this work. It is a limitation. It would also not be out of context to remind ourselves that Marx left behind the perspective of the mental

(or perhaps the psychic, and the subjective) as the first footnote of *Capital*.[17] This footnote that was ignored thereafter by Statist Marxism has left a big hole in the effort of scholars and activists to build *just* political economies. Or do the setting up of relations with unreason and relations with labor require further the kind of reflexive work on the self that is paradigmatic of Gandhi's *My Experiments with Truth*? Further still, this paper has looked at the political economy of mental health along the axis of time. It has not looked at it along the axis of space. That in itself is another limitation of this paper. Would the questions change a little if one tries to conceptualize both mental health and political economy in other spaces, from experiences, texts, and conceptual frameworks at work in other spaces?

References

Achuthan A, Biswas R, Dhar A (2007). Lesbian standpoint. Sanhati, Calcutta

Agamben G (2007) Infancy and history: on the destruction of experience. Verso, London and New York

Andrews G, Issakidis C, Carter G (2001) Shortfall in mental health service utilisation. Brit J Psychiatry 179:417–425

Badiou A (2001) Ethics: an essay on the understanding of evil. Verso, London

Becker G (1976) The economic approach to human behaviour. University of Chicago Press, Chicago

Becker G (1993) Human capital: a theoretical and empirical analysis with special reference to education, 3rd edn. University of Chicago Press, Chicago (originally 1963)

Bir A, Frank R (2001) Mental illness and the labor market in developing nations. Commission on macroeconomics and health. Working paper series WG1:6. World Health Organization, Geneva

Chakrabarti A, Dhar A (2009).Dislocation and resettlement in development: from third world to world of the third. Routledge, New York

Chakrabarti A, Dhar A, Cullenberg S (2012) World of the third and global capitalism. Worldview, New Delhi

Chakrabarti A, Dhar A (2012) Gravel in the shoe: nationalism and world of the third in rethinking Marxism: a journal of economics, culture and society, vol. 24, No. 1. Routledge, New York

[17] Marx begins Capital with the following entry point: "The wealth of those societies in which the capitalist mode of production prevails, presents itself as "an immense accumulation of commodities", its unit being a single commodity. "Our investigation must therefore begin with the analysis of a commodity". Fair enough as a contingent object of enquiry or origin. Now let us take a look at the second paragraph of *Capital*: "A commodity is, in the first place, an object outside us, a thing that by its properties satisfies human wants of some sort or another. The nature of such wants, whether, for instance, they spring from the stomach or from fancy, makes no difference. Neither are we here concerned to know how the object satisfies these wants, whether directly as means of subsistence, or indirectly as means of production". What if political economy also requires an investigation into what Marx had set aside? Is a commodity only an object outside us? Or does it have a life of its own? What if one also needs to look 'inside us'? What if one needs to turn to the first footnote of *Capital*: "Desire implies want, it is the appetite of the mind, and as natural as hunger to the body…. The greatest number (of things) have their value from supplying the wants of the mind". (www.marxists.org).

Chatterjee P (2009) Economic crisis highlights mental health issues in India. Lancet 373(9670):1160–1161
Davar B (2012) Mission of mainstreaming rights of persons with psychosocial disabilities—Nepal -1. A Report
Foucault M 2003 (1963) The birth of the clinic: an archaeology of medical perception—translated from the French by A.M.Sheridan—Routledge: Taylor & Francis e-Library
Foucault M (2006) History of madness. Routledge, New York
Foucault M (2008) The Birth of Biopolitics: Lectures at the College de France 1978–1979. Translated by Burchell G. Palgrave Macmillan
Ganguli HC (2000) Epidemiological finding on prevalence of mental disorders in India. Indian J Psychiatry 42:14–20
Government of India, Ministry of Health and Family Welfare (MHFW), National Commission on Macroeconomics and Health (2005) 'Financing and delivery of health care services in India'. Background Papers for the National Commission on Macroeconomics of Health. New Delhi. Government of India
Gururaj G, Girish N Isaac MK (2005) Mental, neurological and substance abuse disorders: strategies towards a systems approach in burden of disease. NCMH Background Paper.a
Kohn R, Saxena S, Levav I, Saraceno B (2003) The treatment gap in mental health care. Bull World Health Organ 82:858–866
Knapp M, Mangalore R, Simon J (2004) The global costs of schizophrenia. Schizophr Bull 30:279–293
Knapp M, Funk M, Curran C, Prince M (2006) Economic barriers to better mental health practice and policy. Oxford University Press in association with The London School of Hygiene and Tropical Medicine
Lacroix K & Siddiqui S (2012) Cultures of violence: a case study of a woman without a past or a future - paper submitted to Economic and Political Weekly for publication
Lemke T (2001a) The birth of bio-politics—Michel Foucault's lecture at the collège de france on neo-liberal governmentality. Econ Soc 30(2):190–207
Lemke T (2001b) Foucault, governmentality, and critique. Rethinking Marxism 14(3):190–207
Marx K (1993) Grundrisse: foundations of the critique of political economy - penguin classics - translation Martin Nicolaus
Murray C, Lopez A (1996) The Global burden of disease. World Health Organization, Harvard School of Public Health and the World Bank, Geneva
National Human Rights Commission of India (1999) Quality assurance in mental health. National Human Rights Commission of India, New Delhi
Parker I (2011) Lacanian psychoanalysis: revolutions in subjectivity. Routledge, London
Patel V, Kleinman A (2003) Poverty and common mental disorders in developing countries. Bull World Health Organ 81:609–615
Polubinskaya S (2000) Reform in psychiatry in post-Soviet countries. Acta Psychiatr Scand 101:106–108
Ranciere J (2007) On the shores of politics. Verso, London (Trans. Liz Heron)
Rorty R (1979) Philosophy and the mirror of nature. Princeton University Press, Princeton
Samaddar R. Sen SK (ed) (2012) Political transition and development imperatives in India. Routledge, London
Sarin A Jain S (2012) The census of india and the mentally ill. Indian J Psychiatry 54(1):32–36
Saraceno B, Barbui C (1997) Poverty and mental illness. Can J Psychiatry 42:285–290
Schieber G, Maeda A (1997) A curmudgeon's guide to financing health care in developing countries. Innovations in health care financing. World Bank Discussion Paper No. 365. World Bank, Washington, DC
Soltani A, Moayyeri A, Raza M (2004) Impediments to implementing evidence-based mental health in developing countries. Evid Based Mental Health 7:64–66
Srinivasa Murthy R, Kishore Kumar K, Chisholm D et al (2005) Community outreach for untreated schizophrenia in rural India: a follow-up study of symptoms, disability, family burden and costs. Psychol Med 35:341–351

Sujatha V, Abraham L (eds) (2012) Medical pluralism in contemporary India. Orient Black Swan, Hyderabad

Tausig M, Subedi S (1997) The modern mental health system in Nepal: organisational persistence in the absence of legitimizing myths. Soc Sci Med 45:441–447

Thornicroft Graham, Maingay Samantha (2002) The global response to mental illness. BMJ 325(21):608–609

Ustun T (1999) The global burden of mental disorders. Am J Public Health 89:1315–1318

Whiteford H, Teeson M, Scheurer R, Jamison D (2001) Responding to the burden of mental illness. Commission on macroeconomics and health. Working paper series WG1: 12. World Health Organization, Geneva

World Health Organization (2001a) The world health report 2001- mental health: new understanding, new hope. World Health Organization, Geneva

World Health Organization (2001b) Atlas: mental health resources in the world 2001. World Health Organization, Geneva

World Health Organization (2003a) Mental health in emergencies: mental and social aspects of health of populations exposed to extreme stressors. World Health Organization, Geneva

World Health Organization (2003b) Investing in mental health. World Health Organization, Geneva

Wells K, Miranda J, Bauer M et al (2002) Overcoming barriers to reducing the burden of affective disorders. Biol Psychiatry 52:655–675

Zachariah A, Srivatasan R, Tharu S (ed) (2010) Towards a critical medical perspective: reflections on the dilemmas of medical culture today (on behalf of the CMC-Anveshi collective). Orient Blackswan, Hyderabad

Index

S. Banerjee and A. Chakrabarti (eds.), *Development and Sustainability*,
DOI: 10.1007/978-81-322-1124-2, © Springer India 2013