

Equity, efficiency and inequality traps: A research agenda

François Bourguignon · Francisco H. G. Ferreira ·
Michael Walton

Published online: 7 December 2006
© Springer Science + Business Media, B.V. 2006

Abstract This paper discusses a research agenda that arises from unanswered questions and unresolved issues considered in the World Bank's *World Development Report 2006: Equity and Development*. After formalizing the key concepts of equity; equality of opportunity; and efficiency, and proposing a definition for an equitable development policy, the paper discusses the concept of inequality traps, around which the research agenda is structured. Four broad groups of research questions are highlighted: those revolving around the measurement of inequality of opportunity and the diagnostics for the existence of an inequality trap; those dealing with the causes of inequality traps; the quantification of their efficiency costs; and those related to how institutions (including governments) evolve to overcome inequality traps.

Key words equity · equality of opportunity · inequality traps.

1 Introduction

The World Bank's World Development Report 2006, entitled *Equity and Development*, discusses the importance of distributional considerations for the process of development. Drawing on recent advances in political philosophy,¹ the report's answer to Sen's [56]

Bourguignon and Ferreira are in the Development Economics Vice-Presidency, The World Bank. Michael Walton is at the John F. Kennedy School of Government, Harvard University. Together, they led the team that produced the World Bank's *World Development Report 2006: Equity and Development*.

¹The most influential references are probably Dworkin [31], Arneson [7], Cohen [30] and Roemer [53].

F. H. G. Ferreira (✉)
Development Research Group, The World Bank, 1818 H Street, NW, Washington, DC 20433, USA
e-mail: fferreira@worldbank.org

F. Bourguignon · F. H. G. Ferreira
Development Economics Vice-Presidency, The World Bank, Washington, DC, USA

M. Walton
John F. Kennedy School of Government, Harvard University,
Cambridge, MA 02138, USA

Tanner Lectures question “Equality of What?” was to focus on the concept of inequality of opportunity. Like all World Development Reports (WDRs), this was neither an original research volume nor an academic survey of the literature. WDRs are synthesis documents, that draw selectively (though hopefully fairly) on various literatures, in order to illustrate and document a set of key messages that World Bank staff feel are important for policy-making in developing countries. They are message-driven documents, and the central message in this particular report was that the pursuit of equity – defined as greater equality of opportunity combined with the avoidance of extreme deprivation in the space of outcomes – was not only intrinsically desirable, but could, if appropriately pursued, enhance economic efficiency.

While message-driven synthesis documents are useful for many purposes, they are not particularly well-suited to a discussion of unresolved issues and debates, or to the formulation of new research questions which the act of synthesis itself generates. This paper is an attempt – by three among the many people involved in the preparation of the WDR 2006 – to reflect on the problems, unknown quantities and research questions that arose from the synthesis of existing evidence and perspectives in that report. Our aim is to highlight the areas in which the evidence and arguments presented in the WDR are at their most tentative or incomplete, and where the pay-off to new research – both theoretical and empirical – is likely to be greatest.

The paper is organized around the notion of *inequality traps* – persistent differences in power, wealth and status between socio-economic groups, that are sustained over time by economic, political and socio-cultural mechanisms and institutions. In some cases, it is possible that these differences lead to efficiency losses, resulting in an economic equilibrium that is inferior to some feasible alternative. It is argued, therefore, that this concept is crucial for an understanding of the link between equity and efficiency.

Section 2 motivates the paper by offering formal definitions of various basic concepts – including equity, equality of opportunities, and efficiency – that are discussed informally in the WDR. Section 3 formally defines an inequality trap in general terms, and provides two specific examples. Section 4 highlights four areas where further research is needed for a better understanding of inequality traps, and of the role of equity in development more generally: (a) the measurement of inequality of opportunity and group-based mobility studies; (b) identification of the causes of inequality traps; (c) quantification of their efficiency costs; and (d) analysis of how institutions evolve to overcome these low-level traps, and of what policies might be helpful in that process. Section 5 summarizes and offers some brief conclusions.

2 Basic concepts: Equity and efficiency

Before defining an inequality trap, we briefly define another two key concepts used in the WDR – equity and efficiency – more formally than was done there. *Equity* is defined in terms of two basic principles: (a) equal opportunities and (b) avoidance of extreme deprivation in outcomes. The first problem in defining equal opportunities concerns specifying what the opportunity is for. Philosophers and economists have offered various responses. For example, Sen [57] proposed the concept of capabilities, or the set of possible “functionings” that people have reason to value, and can choose to pursue. Functionings occupy many dimensions, and can range from living a healthy life, to being able to participate in the life of a community. Roemer [53] assumes away the problem of multi-

dimensionality by reducing it to a single dimension: “Let the members of the relevant population enjoy *a certain kind of success or advantage...*”(p.25, emphasis added). He considers this unidimensional advantage u to be a function of a person’s circumstances c , and efforts e , as well as of a society’s chosen policy, ϕ : $u(c, e, \phi)$. We follow Roemer’s formulation here, since it simplifies the presentation of the basic ideas.

Central to Roemer’s theory is the distinction between the “circumstances” that individuals enjoy, which are independent of their own choices, and the “efforts” that they exert. Circumstances are exogenous to the individual, by definition, and differences in circumstances are argued to be morally irrelevant to outcomes, while efforts can lead to morally justified differences in achievements.² Roemer [53] defines the *equal opportunity policy*, by partitioning the population into J types, such that each type j consists of all individuals i with identical circumstances, c . The partition is $\Pi = T_1 \cup T_2 \cup \dots \cup T_j \cup \dots \cup T_J$ such that $c_i = c^j \forall i \in T_j, \forall j$. Examples of circumstance variables that may be used for such a partition of the population include gender, ascribed socio-cultural group (for example, black, mixed or white in Brazil, caste status in India), or family background (proxied by parental wealth or education, for instance).

Given that effort levels vary within types, each type is characterized by a cumulative distribution of the advantage, given by $F_\phi^j(u)$. The subscript ϕ indicates that this distribution may be affected by, and thus depends on, the set of policies (to which we refer simply as “the policy”) chosen by the government. As effort is the only source of differences among individuals of type j , and as $u(\cdot)$ can reasonably be assumed to be increasing in effort, this cumulative distribution represents increasing levels of effort. Writing an ‘indirect advantage function’ $v^j(\pi, \phi) = F_\phi^{j-1}(\pi)$, as a function of the individual’s rank (or centile) in the cumulative distribution of advantage in type j , Roemer then defines the equal opportunity policy, ϕ^* , as the solution to the program³:

$$\max_{\phi} \min_j v^j(\pi, \phi). \tag{1}$$

Thus defined, ϕ^* is the policy that maximizes the advantage of people in centile π of the distribution in the type for which that particular centile has the lowest advantage across all types. This formulation captures the principle that greater advantage is better, in an extension of the Rawlsian maximin approach across types. In effect, however, the optimal policy ϕ^* depends on the effort centile, π . Recognizing that this optimization is centile-specific (i.e., that the optimal policy that solves Eq. 1 for $\pi=0.17$ will in general differ from the one that solves Eq. 1 for $\pi=0.83$), Roemer proposes an averaging compromise:

$$\max_{\phi} \int_0^1 \min_j v^j(\pi, \phi) d\pi \tag{2}$$

Program Eq. 2 gives equal weight to raising the advantages of the lowest types across all centiles of the effort distributions. This is only one possible weighting scheme; a different,

² What counts as a “circumstance” and what counts as an “effort” is ultimately itself a moral choice, that is perhaps best considered a product of open social and political decision-making. For example, some consider genetic talent a morally justified basis for differences in outcomes (i.e., an effort), while others see it as a circumstance.

³ As indicated, the ‘indirect advantage function’ is simply the inverse distribution function of advantage within group j . It gives the advantage level corresponding to effort centile π in group j , given policy ϕ .

but equally arbitrary, aggregation compromise – which Roemer [54] sees as the operational compromise embraced by the WDR 2006 – would be

$$\max_{\phi} \min_j \mu^j(\phi) \text{ with } \mu^j(\phi) = \int_0^1 v^j(\pi, \phi) d\pi \tag{3}$$

where one seeks the policy that maximizes the mean advantage in the type with the lowest mean advantage.⁴

One could clearly use Roemer’s set-up to define equality of opportunity itself, rather than the equal opportunity policy. Under a *strong criterion*, which appears to be implicit in much of Roemer’s work, equal opportunities would attain if:

$$F^j(u) = F^k(u), \forall j, k \tag{4}$$

The strong criterion requires that the within-type distribution of advantages be identical across all types. This requirement implies that the circumstances by which the population was partitioned into types be immaterial to advantage. Two people in identical centiles of the distribution of effort, but with different circumstances, would have exactly the same advantage level – a rather natural definition of equal opportunities. Intuitively, high-effort Afro-Brazilians would enjoy the same advantage levels as high-effort white Brazilians; low-effort individuals within each type would also have the same outcomes as each other.

A *weaker criterion*, which is implied by Eq. 4 but does not imply it, is that the mean advantage levels across all types be equalized:

$$\mu^j(u) = \mu^k(u), \forall j, k \tag{5}$$

These are the definitions of equal opportunities implicit in the WDR 2006. But what might be the optimal policy advocated by its authors? Although the policy that solves Eq. 3 – as suggested by Roemer [54] – is a good first approximation, we find it useful to be explicit about dynamics. We live in a world in which advantage levels change, as the economy’s production possibility frontier expands. Policies that maximize today’s advantage for a particular group may not be those that lead to higher growth rates, and thus possibly to higher levels of advantage for that group in the future. As time passes, the advantage rank of different types may change, requiring a re-optimization. And so on. We might therefore be looking, at time t , for the policy Φ_t that solves:

$$\max_{\Phi_t} \min_j \int_t^{\infty} e^{\delta(t-s)} \mu_s^j(\Phi_t) ds \tag{6}$$

where δ is a discount rate, which we assume constant for simplicity, and ‘policy’ Φ_t is now defined as the whole sequence of time-specific policies (φ_s) to be undertaken at times $s \geq t$.⁵ Equation 6 simply replaces the static mean advantage of type j in Eq. 3 with the present discounted value of its stream. It is a minor formal change, but it serves to highlight the

⁴ Note that the difference with Eq. 2 lies only in the order in which the last two operators, \int and Min , are applied.

⁵ Note that, in a dynamic context, it is possible that the type of an individual might change over time. Equation 6 must thus be understood either in a model where individual types are invariant over time, or with the index j referring to the initial type – i.e., at time t – of an individual. To simplify, we will implicitly refer to the first case in what follows.

importance of taking a long-term perspective when choosing the policy that maximizes (the stream of) advantage for the poorest group. Policies affect saving and accumulation incentives, which in turn affect the growth process of μ^j . This formulation, which sees the optimization at t taking the future into account, also allows for the possibility that the most-disadvantaged group (min $\{j\}$) changes over time, with corresponding changes in policy choices.

The WDR 2006 explicitly recognizes (in a box on p. 78) that maximizing the lowest level of opportunities or advantage (Eq. 3 in a static set-up, or Eq. 6 in a dynamic context) is a different problem than equalizing opportunities (Eqs. 4 or 5), and that each of these in turn differs from maximizing the sum of advantages across all groups. The report deliberately chose not to take a position on which of these objectives a country should pursue, on the ground that such a central social choice should be the subject of an internal socio-political debate within each country.

Personally, we find it compelling to think of an *equitable development policy* as the one that solves:

$$\max_{\Phi_t \in \Phi} \min_j \int_t^\infty e^{\delta(t-s)} \mu_s^j(\Phi_t) ds \tag{7}$$

$$\text{subject to } u_s^j \geq \bar{u}_s \mid \forall i, \forall j, \forall s \geq t \tag{8}$$

Inspection of the problem (7, 8) suggests the following characterization of the optimal policy:

- (a) accounting explicitly for the dynamic nature of the development process, the optimal policy is the one which maximizes the presented discounted value of advantages for the least-advantaged type;⁶
- (b) provided that it is consistent with the absence of severe deprivation (defined by having no individual advantage below some critical level \bar{u}_t , which is allowed to vary over time, as a society’s definition of deprivation changes);⁷
- (c) and provided that the policy sequence belongs to some permissible set Φ . We interpret this permissible policy set as a subset of the set of technically feasible policies, reflecting social choices about the legitimacy of various policies. Thus, forced labor, forced fertility control, or expropriation of property might all be feasible policies, but might (or might not!) be deliberately excluded from the permissible set Φ by social choice. This incorporates some of the concerns of those who, like Nozick [46], argue that fairness cannot be sought only in final allocations, but in the rules governing the processes by which such allocations are arrived at.

This formulation of the problem captures the two principles which make up our definition of equity: it makes avoidance of severe deprivation a constraint that must be

⁶ The formulation of the problem implies selecting the type with the lowest present value of advantage over time, which is not necessarily the same as picking *today’s* least advantaged type. Obviously, under uncertainty one would replace variables by their expectations.

⁷ The critical deprivation level \bar{u}_t is not necessarily the same as a society’s chosen poverty line: if an individual has a level of income below the poverty line she is *categorized* as deprived within the society, whereas if her level of advantage falls below the critical deprivation level the society takes action to lift her back above this level. In very poor societies this may be a very low level, such as preventing death from lack of food in famine conditions.

satisfied in the process of pursuing the broader objective of equal opportunity. The implication is that the policy that solves this problem (let us call it Φ^{**}) may in general lead to a lower present discounted value of advantages for the poorest group than that which would solve the unconstrained problem. Upholding the principle of avoidance of deprivation in outcomes is not, in general, costless. We find the above formulation appealing in that it makes “poverty reduction” (understood in this context as enforcing a minimum level of advantage for all, regardless of *both* circumstances *and* efforts) a necessary *requirement* for equitable policy, but not its ultimate *objective*. That ultimate objective goes beyond the elimination of absolute deprivation, and is the pursuit of “equal opportunities” in the Rawls-Roemer sense.

The proposed definition of an equitable development policy in terms of types should be interpreted as an approach to evaluation. It does not necessarily imply that policy should be designed around types. For example, affirmative action for Afro-Brazilians or lower Indian castes may or may not make sense. There may indeed be cases where linking policy to types can be harmful, through making policies unnecessarily conflictual or increasing the salience of discriminatory social systems based on the categorization of people. Furthermore, types themselves are likely to be contested and dynamic, and to be endogenous in long-run paths of economic and social change. Policy design requires context-specific analysis that recognizes these possibilities.

Having thus defined two aspects of equity – equality of opportunity in Eqs. 4 or 5, and a variant of Roemer’s equal opportunity policy (which we called the equitable development policy) – let us briefly clarify the concept of efficiency used in the WDR. Roemer [54] questions whether by “efficiency” we mean “technical” efficiency (defined in textbooks as the use of any combination of inputs such that the output lies *on* the production function, rather than below the frontier of the production set) or “Pareto” efficiency. In fact, we meant neither of those, but *economic efficiency* in the sense of the largest possible value of output (or advantage, for this analysis) given a production possibility set.

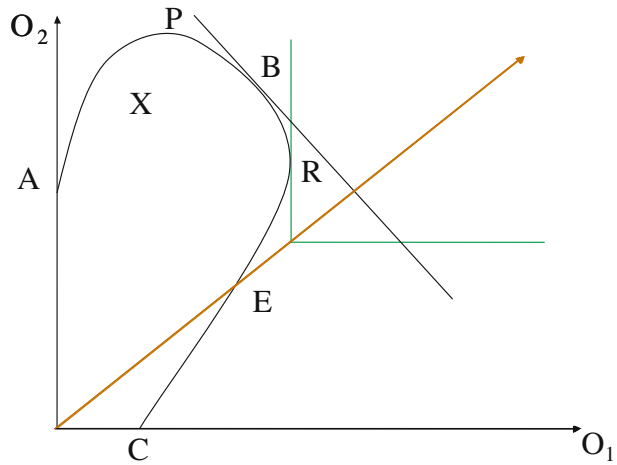
Economic efficiency (in the space of advantages) is clearly more demanding than technical efficiency, in that it requires, in addition to technical efficiency, a set of economic decisions that maximizes advantage. Statically, economic efficiency coincides with Pareto efficiency for some particular distribution of aggregate advantage. But the two concepts differ when used to assess *movements* from an initial vector of advantages A to another vector B. The move from A to B is Pareto efficient (or ‘Pareto-improving’) if and only if no person is worse off at B than at A, and at least one person is better-off. Our concept of economic efficiency requires only that overall aggregate advantage be higher in B than A. It might therefore be termed *Kaldor-Hicks efficiency*, since the Kaldor-Hicks criterion differs from the Pareto criterion exactly in that it requires only that it be possible for a social planner to compensate any losers in the move from A to B, rather than that there *be* no losers. The Kaldor-Hicks test effectively approves a move from allocation A to B if the ‘size of the pie’ is larger in B than in A. This is the concept of economic efficiency which we adopt, here and in the WDR 2006.

When suggesting a positive link between equity and efficiency, the WDR 2006 does *not* argue that the equitable development policy sequence (Φ^{**} , which solves Eq. 7, subject to Eq. 8), also solves:

$$\max_{\phi \in \Phi} \int_0^{\infty} u dF(u) \quad (9)$$

the utilitarian objective function (in the space of advantages). Equations 6 and 9 are clearly different problems, and the solutions will therefore also differ, as illustrated by Figure 1 for

Figure 1 Different choices along an “Advantage Possibility Frontier.” Figure 1 is a reproduction of box figure 4.1 in WDR 2006 and, as acknowledged there, is drawn from Buchanan [26], through Atkinson and Stiglitz [8].



the static case. If the “advantage possibility frontier” in a society with two types of individuals is given by the curve AC, then the (static) solution to Eq. 6 will be at point R (for Rawls), while the solution to Eq. 9 will lie at point B (for Bentham). What the WDR does argue is that, since most countries are likely to be in interior positions such as X, Φ^{**} may in many instances lead to an overall level of aggregate advantage which is higher than that observed today.⁸ In that case, the pursuit of equity may enhance (but not maximize) efficiency.

3 Inequality traps

We now turn to the concept of an “inequality trap” that is central to the WDR’s overall argument. An intuitive definition is provided by Rao [51], who first coined the term: “Inequality traps...describe situations where the entire distribution is stable because the various dimensions of inequality (in wealth, power, and social status) interact to protect the rich from downward mobility, and to prevent the poor from being upwardly mobile” ([51], p.11). The WDR’s contention is that such inequality traps are pervasive; are inconsistent with equality of opportunity; and that there is a significant class of cases in which they are *also* associated with inefficiency. Here we first describe the concept formally and then turn to questions of normative evaluation and policy choice.

As the concept of an “inequality trap” is essentially about persistence over time, its definition requires a general description of the dynamics of advantage. As indicated earlier, an individual i of type j has advantage level $u_t^{ij} = u(c_t^j, e_t^{ij}, \phi_t)$. The subscript t denotes a particular time period. The two general assumptions needed to define the concept of inequality traps are: first, that there be persistence in relative positions in a distribution across time periods, and that this be (partly) a product of features of the overall distribution – or of relations between groups. This might come about because the circumstances enjoyed by group j today (e.g., African-Americans in the year 2000), depend in part on the advantage levels enjoyed by the preceding generation (e.g., African Americans in the

⁸ This would happen here provided the line with slope -1 which goes through X lies below the line with the same slope that goes through R.

1960s): $c_t^j(u_{t-1}^j, \xi_t)$, where ξ_t denotes an innovation at time (or generation) t . Similarly, an individual i 's effort levels today may also depend both on past advantage (their own or of a previous generation), and on the *distribution* of past advantages: $e_t^{ij}(u_{t-1}^{ij}, F_{t-1}(u), \zeta_t)$, where $F_{t-1}(u)$ denotes the entire distribution of advantages at generation $t-1$, and ζ_t denotes another innovation at generation t .^{9,10} This innovation should include the element of personal choice that distinguishes an effort from a pure circumstance. Third, it is plausible that the policy in place at time t has also been determined in part by the prevailing distribution of advantages at time $t-1$: $\phi(F_{t-1}, \theta_t)$, where θ_t denotes a third innovation at time t .¹¹

For now, we treat policies as completely endogenous, although some element of exogeneity could be introduced by allowing θ_t – or some component of θ_t – to be a control variable through which policymakers at time t could influence the evolution of policy.¹² Alongside the intertemporal transmission of circumstances, the dependence of both efforts and policies on the previous distribution of advantages are key elements in the transmission processes of the distribution of advantage from one period to another. A very general reduced form of the dynamic process of advantage can therefore be written as:

$$u_t^{ij} = \Gamma(u_{t-1}^{ij}, F_{t-1}(u), \gamma_t) \tag{10}$$

where γ_t is a vector of innovations. Equation 10 simply defines, at a general level, a first-order stochastic dynamic process, in which individual (or lineage) i 's current advantage depends on his own past advantage, as well as on the entire distribution of advantages in the preceding period (or generation).

Still at a very high level of generality, assume that this process is characterized by multiple equilibria, in each of which the distribution of advantages converges to a well-defined long-run distribution $F_{\infty(k)}(u)$. Lest notation become confusing, note that the subscript k here denotes a particular long-run – or limiting – distribution of advantages, each of which corresponds to a particular equilibrium of the dynamic process. It is not to be confused with the superscript j , which refers to the distribution of advantages for a specific type, j . As with any dynamic process with multiple equilibria, the idea is that differences in initial conditions (in the initial distribution of advantages and circumstances, for instance) may imply convergence to different long-run equilibrium distributions. Denote by K the set of equilibria for this process. We classify a particular long-run equilibrium k ($k \in K$) as an inequality trap if for any two types, j and l , the distribution of advantage for type j , $F_{\infty(k)}^j(u)$, is dominated by that of type of l , $F_{\infty(k)}^l(u)$, provided there exists some alternative

⁹ Formally, $F_{\phi,t}$ is a mixture of all type-specific distributions $F_{\phi,t}^j \forall j$.

¹⁰ Piketty [47], for example, writes a model in which beliefs about the pay-off to effort in terms of mobility depend on one's family history, and on that family's outcomes relative to those of others. This is directly analogous to sociological work in the tradition of Bourdieu [22] and to the concept of a "capacity to aspire" that is unequally distributed across groups (Appadurai, [5]), as discussed below.

¹¹ A large class of political economy models yield results in which policy variables depend on the distribution of wealth or incomes at an initial period. We will return to these examples below.

¹² Such a view would be intermediate between the full exogeneity of policy choices implicit in the optimization problems in Section 1, and the full endogeneity implicit in most political economy models. It would allow for some "policy space," recognizing both the dependence of policy on historical and political factors, but also the possibility for policymakers to make a difference on the margin. A proper model of political economy in which policy makers had some "residual rights of control" over policy decisions not stipulated in a "political contract" with voters might be one interesting research idea.

equilibrium m ($m \in K$), in which no dominance relationships exist between $F_{\infty(m)}^j(u)$ and $F_{\infty(m)}^l(u)$, for any j, l .¹³

In plain English, this classification defines an inequality trap as a long-run distribution of advantages in which a particular social group does persistently worse than some other social group, even though an alternative equilibrium exists where no two social groups can be similarly ranked. “Social group” is made precise in terms of a Roemerian type, defined by exogenous circumstances. To “do persistently worse” is made precise in terms of either first- or second-order stochastic dominance of the long-run distribution of advantages. In the case of first-order dominance, this implies that for each centile in the distribution of effort, people in the dominated group enjoy lesser advantages than in the dominant group, exclusively as a result of their different circumstances. The salient feature of an inequality trap is this permanent “non-convergence” in the opportunities of some social group.

Two remarks are in order at this stage. First, the existence of inequality traps is by no means guaranteed. The foregoing definition involves a number of assumptions, including the existence of multiple limiting distributions. Even more demanding, it requires that the limiting distribution one observes be characterized by a particular form of inequality of opportunity, which is not present in some feasible alternative distribution. Whether or not a particular society is in an inequality trap is therefore an empirical matter.¹⁴

Second, the dependence of efforts and policy choices on the *entire* distribution of past advantages is one of the factors that differentiate an inequality trap from a poverty trap. In a poverty trap, the incomes of the poor do not grow beyond some fixed threshold: the poor remain forever poor. In one typical poverty trap story, the poor remain poor because they are undernourished, and thus lack the energy to be very productive in their fields. Such a mechanism need not involve any interaction between the various social groups (or types). An inequality trap, on the other hand, does allow for the advantages of the poor to grow over time, as long as patterns of unequal relative advantage persist in the long run. The dynamics of such persistent differences in opportunities are affected by the entire distribution of advantage, reflecting (economic, political and socio-cultural) interactions *across* groups.

To make the concept of an inequality trap a little more concrete, consider two alternative examples that illustrate different mechanisms that might generate inequality traps. These mechanisms flow either from political processes or socio-cultural conditions, which were captured by the endogenous determination of policy and effort in the above formalization. In a first story, suppose productivity (and thus wages) are determined by the quality of the school one attends. Capital markets are highly imperfect, so that children from poorer families cannot afford to attend private schools which charge high fees, and go instead to free public schools. Richer families send their children to private schools, and are prepared to pay fees that enable these schools to provide high-quality schooling. Under certain conditions, one may observe a sorting equilibrium, in which all families above a certain wealth threshold send their children to private schools, while all families below it send their children to public schools. Now suppose that the budgets for public schooling are determined by the level of taxation, which is voted upon by all citizens. If political power is somehow related to wealth, it is possible that the pivotal voter is wealthy enough to send

¹³ Dominance can be by a demanding first-order stochastic dominance criterion; or by the weaker, and probably more reasonable, second-order stochastic dominance criterion.

¹⁴ Since the definition involves a comparison between an actual and a counterfactual long-run distributions, such empirical testing is unlikely to prove straight-forward.

her kids to private school, in which case she might not value public schooling at all, and vote for very low taxes.

An equilibrium arises in which the children of the poor stay poor because they attend bad schools, and the children of the rich stay rich because they attend good schools, and in which the quality of both sets of schools reflects the economic and political power of the parents. Publicly educated workers in this equilibrium are: (a) a Roemerian type, in that they share the common circumstance of having poor parents; and (b) worse-off than they would have been in a feasible alternative equilibrium, which can be shown to arise from a redistribution of wealth and political power. These two features qualify this equilibrium of the model as an inequality trap. Further, it can be shown that, in such a model, overall income levels are lower in the inequality trap equilibrium than in the counterfactual. This story comes from Ferreira [34], and reflects many of the insights in Bénabou [16]. It is an inequality trap in which the nature of the inter-group relationship determining the trap is to a large extent political.

A second story is one in which a distinct group – which could be a majority group, such as women, or a minority ethnic or caste group – is (erroneously) believed to be in some sense “inferior” to a dominant group. There may be both rational and less rational mechanisms through which members of the stigmatized group adopt behaviors that confirm the prevailing perception of its inferiority, contributing to its persistence, and leading to a low-level trap. One rational mechanism is the internalization of discrimination in the future, when making an investment decision in the present: if a person expects that discrimination will reduce the returns to her schooling in the future, the optimal amount of schooling she will choose today will be less than that for an otherwise identical individual that does not expect to be discriminated against. Less rational mechanisms may also be at work, as when individuals internalize beliefs about their own inferiority, and adjust their ambitions downwards accordingly – a phenomenon that is an element of a reduced “capacity to aspire” [5]. Both types of mechanism are inequality traps, in which the nature of the inter-group relationship determining the trap is to a large extent cultural. Anthropological and sociological accounts see these phenomena as intrinsically a product of unequal social relations, perpetuated via the unequal (informal) institutionalized patterns of cultural and social capital.

Such are the kinds of phenomena that our definition of inequality trap attempts to formalize. The formalization may appear somewhat abstract, but the point is that a generalized analysis of inequality traps in the intuitive sense of the previous examples is presently not available. Better understanding the general structure of inequality trap models, and possibly better characterizing those equilibria, should rank high in theoretical research. This is all the more important given that the preceding examples are just two examples from a much larger literature that spans various disciplines. We refer to a few more examples below in the section on future research.

3.1 Equity and efficiency properties of inequality traps

Note that there is no simple logical mapping between inequality traps, equitable development policies and efficiency. In Section 1, we discussed three normative criteria – equality of opportunity; “equitable development” and maximum Kaldor-Hicks efficiency – which were illustrated in Figure 1 by points E, R and B, respectively. Inequality traps, by definition, involve inequality of opportunity: since individuals in group j (and their descendants) face a worse long-run distribution of advantage than those in group l , owing to circumstances outside their control. A point X that corresponded to an inequality trap could never coincide with point E.

But the relationship with the equitable development criterion (corresponding to the policies that solve Eqs. 7, 8) is more ambiguous. It is *logically possible* that the long-run distribution corresponding to the choice of an equitable development policy (represented by R in Figure 1) is itself an inequality trap. This is because the Rawlsian maximin criterion leads to the long-run equilibrium with the “best” distribution of advantages for the lowest type, comparing lowest types across all feasible equilibria. This could occur in a distribution where the lowest type is dominated by some other type. If this happened at a point such as R, and a distribution corresponding to a point such as E was also feasible, then the equitable development outcome (R) would be classified as an inequality trap. This is just another way of re-stating the fact that it is possible that there be a trade-off between the Rawlsian maximin, and perfect equality of opportunity. Similarly, an equilibrium with an inequality trap may be more or less Kaldor-Hicks-efficient than one without such a trap. A “trap” equilibrium corresponding to an interior point such as X may lie on an iso-advantage curve above or below that of E.

Although it is important to recognize these logical possibilities, the concept of inequality trap would be an interesting one if breaking such a trap – by removing the economic, social or political mechanisms that reproduce inequalities across types – moved an economy closer to either the equitable development outcome (of Eqs. 7 and 8, and point R in Figure 1), or to the most efficient equilibrium (in a Kaldor-Hicks sense, as in point B in Figure 1). In fact, for any inequality traps that corresponded to some interior point X in Figure 1 that lay below the line with slope -1 going through point R, a policy that broke the trap and moved the economy in a “Northeasterly direction” would achieve *both* of these objectives. In fact, the “equitable development policy” of Section 1 might be just one such policy.

4 Research questions on inequality traps

Whether or not inequality traps provide a useful lens through which to view persistent inequalities depends on the ability of future theoretical and empirical research to shed further light on them – in terms of diagnosing their existence (and measuring the inequality of opportunity associated with them); understanding their causes; quantifying their costs; and understanding how to switch equilibria, by transitioning out of them. We sketch research issues in each of these four areas.

4.1 Measurement

Our ignorance begins with the very first component of our definition of equity, namely the idea of *inequality of opportunity*. While the concept has been much discussed in political philosophy, attempts to *measure* the degree of inequality of opportunity in a population are still relatively rare, and very recent. A few papers by John Roemer and co-authors have addressed the closely related question of the impact of specific policies on the distribution of advantages (typically proxied for by either incomes or education levels) across types. Betts and Roemer [21] do this for a hypothetical educational finance reform in the United States, and Roemer et al. [55] compute the contribution of fiscal regimes (tax and transfer systems) in a number of countries to the equalization “of opportunities among citizens for income acquisition.”

Others have sought to measure inequality of opportunity directly. Bourguignon et al. [25] decompose observed earnings inequality into a component due to opportunities (determined by exogenous circumstance variables, such as race, region of birth and family

background), and a residual component. Their “inequality of opportunity index” is given by $\Theta_j := \frac{I(w) - I(w|C_i = \bar{C}, v_i)}{I(w)} = \frac{I(w) - I(\bar{w})}{I(w)}$, which measures the share of earnings (w) inequality (as measured by some standard inequality measure $I(w)$), which is eliminated when the circumstance variables are counterfactually held constant across all individuals – an attempt at generating a counterfactual distribution in which there was a single type (with circumstances \bar{C}).

Pistolesi et al. [48] depart from a definition of equal opportunities like that in Eq. 4, and test for it across a number of OECD countries by comparing conditional distribution and “deficit” functions.¹⁵ A failure to reject the null that two distribution functions (conditional on type) are identical would be a failure to reject Eq. 4, and hence equality of opportunity. They also allow for a weaker test, in which a rejection of second order dominance between two distributions (in both directions) is taken as a failure to reject equality of opportunity. This is a very interesting approach, because it derives directly from the strong criterion for equal opportunities: $F^j(u) = F^k(u), \forall j, k$. The same authors also propose a scalar index to measure inequality of opportunity, which is essentially a Gini aggregator of the areas under the Generalized Lorenz Curves for each type.

Checchi and Peragine [28] develop a non-parametric approach to obtain the same basic decomposition as Bourguignon et al. [25]): that of overall observed inequality into inequality of opportunities and a residual. A ‘types approach’ consists basically of scaling up or down the within-type distributions until they all have the same mean. The difference between the inequality in this counterfactual distribution and the actual inequality is one measure of inequality of opportunity. Conversely, a ‘tranches approach’ consists of replacing the incomes of all individuals in homogenous effort groups (which are called ‘tranches’ and within which all differences are due to circumstances) with their mean incomes. The inequality in this latter counterfactual distribution is another measure of inequality of opportunities. In yet another approach, Van de Gaer et al. [62] develop an index to measure inequality of opportunity in the context of intergenerational mobility, which is essentially a measure of inequality in the expected destination columns, across the different rows of a transition matrix.

These various approaches have only recently been proposed. They all derive from Roemer’s theory, and rely on the definition of types of individuals with identical circumstances. Yet, they differ in the exact manner in which they compare distributions, or construct scalar indices. Some – like Bourguignon et al. [25] and Checchi and Peragine [28] – seem to seek exactly the same decomposition, but through different methods. The moment seems ripe to take stock of these methodological differences by comparing them both analytically and empirically. Additionally, these studies have drawn primarily on data from OECD countries. To our knowledge, Brazil is the only developing country to which formal measurement of inequality of opportunity has been applied (both by Bourguignon et al. [25] and by Cogneau and Gignoux, 2005, unpublished manuscript). If the WDR is right that inequalities of opportunity are an integral part of the underdevelopment story, with large segments of the population stuck in traps due to a lack of opportunity to invest and grow, then one would expect developing countries to be prime candidates for empirical studies of this kind – hopefully in a way that distills the lessons from and compares the strengths and weaknesses across the aforementioned pioneering studies. Additionally, all of these approaches have so far focused on one-dimensional measures of advantage. It may be interesting to explore the implications of allowing for multiple such dimensions.

¹⁵ A deficit function is the integral of a distribution function: $G(y_k) = \int_0^{y_k} F(y) dy$.

Beyond measuring inequality of opportunity, perhaps a greater and more pressing challenge is diagnosing the existence of inequality traps themselves. Given the definition of an inequality trap, diagnostics would almost certainly entail an understanding of socio-economic mobility and, more specifically, of differential mobility patterns by ‘type.’ As a first pass, one might compare transition matrices across different groups (defined by ‘circumstances’), so as to forecast limiting distributions of some measure of advantage (such as income, earnings, wealth, or education). Ascertaining differential mobility patterns by groups, including a lack of convergence, is not a sufficient condition for the existence of an inequality trap, but it is necessary.

If the absence of convergence across ‘types’ in a society – on the basis of observed mobility patterns – can be established, then the next step in diagnosing an inequality trap is to establish that there exists a feasible alternative equilibrium in which mobility patterns are such that there is no trap. Given its counterfactual and dynamic nature, such a test may prove impossible to implement empirically in a fully satisfactory manner. But suggestive evidence may be obtained from two sources: comparisons across countries (or geographic regions within countries) in terms of mobility patterns for advantaged and disadvantaged groups; and model-based simulation of the dynamics of advantage (of education, income or other dimensions) under alternative assumptions. Models for ex-ante policy evaluation, that simulate counterfactual income distributions for different groups, such as those recently proposed by Attanasio et al. [9], Bourguignon et al. [24], and Todd and Wolpin [61], are illustrative of this approach. This takes us to questions of causation.

4.2 Causal mechanisms

Ascertaining the existence of non-convergent dynamics across a society’s ‘types,’ is only the first step in a research agenda on inequality traps. Next, and perhaps more importantly, comes a search for the causal mechanisms that underpin these processes. This involves theoretical and empirical exploration of the nature of the economic, political, social and cultural interactions between dominant and subordinate groups which sustain an inequality trap in a particular dynamic process, in comparison with an alternative dynamic equilibrium.

In Section 2, we organized possible channels of causation into three categories: around circumstances (such as initial wealth or power), the endogenous component of policy choice, and culturally shaped behaviors (that influence “effort” in the abstract representation used there). These, of course, interact in a general equilibrium that determines distributional dynamics. This raises challenges for modeling inter-relationships, and even more severe difficulties for the identification of causal relationships. It is hard to think, for instance, of randomized experiments that one could conduct in order to attribute impact. Some authors have drawn on comparative historical studies, treating colonization as a natural experiment, as in Acemoglu et al. [2]. There is also a series of comparative historical studies by Engerman and Sokoloff [32, 33] on the interaction between factor endowments, political inequalities and a range of choices over political and economic institutions, for example related to suffrage and education provision.¹⁶ An illustrative area of interactions with dynamic processes concerns the patent systems of the US and the Britain in the nineteenth century [44]. Patent fees, rights to assignment, application processes and access to information on existing patents were all designed to be open to all in the US, but were costly and effectively restrictive in Britain. The result was that a majority of patentees came from outside the elite, often from people with little schooling in

¹⁶ Hoff [40] surveys this literature.

the US. In Britain patents were much more confined to the elite, and ran at much lower levels. These kinds of historical analyses are valuable exercises, but are often depressingly long-term for policymakers wishing to effect change in the medium term.

The most promising avenue for identifying causation might therefore be detailed analysis of particular mechanisms through which a specific inequality hampers the investment levels or the productivity of individuals that belong to a subordinate group or type. With respect to the first category of causes, around initial circumstances, the WDR 2006 cited a large number of studies that offer evidence of how poverty (more specifically, lack of assets) combines with imperfect capital markets to hamper the investment opportunities of the poor. In the absence of corrective redistribution, the resulting curtailment in the investment opportunities of the poor leads to the persistence of deprivation and to aggregate inefficiency. This is a classic inequality trap, the interaction between market failure and the distribution of wealth or status prevents an underprivileged group from converging towards a higher level of advantages under prevailing policies, whereas there exists another equilibrium in which policies that either tackled the market failure (for example through credit programs for small farmers or entrepreneurs) or various forms of redistribution (for example through titling for the poor or redistribution of land) would have led to a convergent path.

Much less has been done, however – and the research payoffs are therefore likely to be correspondingly greater – in the field of non-economic interactions that affect the economic opportunities and actions of disadvantaged groups. We focus on two mechanisms that are potentially pervasive, but as yet understudied, corresponding to the endogenous component of policy choice, Φ , and to effort choice, e , in the formulation in Section 2: the capture of political and judicial institutions by the powerful; and how beliefs are formed and shape behaviors for advantaged and disadvantaged groups.

Consider first the capture of institutions, which has been the subject of long literatures in sociology and political science. Economic theorists have suggested various channels through which political decisions may deviate from some notion of the “socially optimal,” so as to favor powerful groups. Bénabou [16] and Ferreira [34] consider the capture of electoral systems, when voting power differs with wealth. The capture of political decisions through lobbying has been studied by Grossman and Helpman [36], Besley and Coate [19], and many others. The choice of predatory political institutions by powerful groups has been explored, for example by Bourguignon and Verdier [23] and Acemoglu and Robinson [1]. The design of financial systems that protect incumbents at the cost of both efficiency and access of outsiders is a central theme of Rajan and Zingales [49]. In other work, the same authors argue that political coalitions will often form against efficiency-increasing reforms to expand educational endowments and open markets, because of loss rents to existing oligopolists and educated groups, especially where education is unequally distributed [50]. Moving beneath the national sphere, the potential for local elites to misappropriate resources in a decentralized allocation has been highlighted by Bardhan and Mookherjee [14, 15]. But empirical documentation of whether these processes can indeed be observed, and a quantification of their importance, remains limited.¹⁷ As a discipline, we are in fact remarkably ignorant of the actual decision-making processes within the institutions which we seek to model – be they national parliaments, courtrooms in provincial capitals, or village meetings. A greater understanding of the institutional processes from which policy

¹⁷ Both Acemoglu and Robinson [1] and Rajan and Zingales [49] provide some descriptive support for their work. Galasso and Ravallion [35] and Araujo et al. [6] are examples of specific empirical analyses.

decisions originate would be required to allow us to understand whether – and if so, how – the powerful capture institutions, and what the effects on the subordinate might be.

The second area we highlight is the persistence of socio-cultural inequalities, in which differences in status are manifest in patterns of interaction, behaviors and beliefs that tend to sustain difference (in what Appadurai [5] characterizes as unequal “terms of recognition” between groups). One consequence can be the internalization of self-depreciating beliefs by members of disadvantaged groups as a product of relations with dominant groups, and the impact of these beliefs on their actions and outcomes. This is another area to which economists have come recently, and where social psychologists and sociologists have long been active. In economics, Piketty [47] was one of the first to write a theoretical model of how beliefs about the relationship between effort and mobility may influence one’s own learning and investment decisions. If these beliefs are acquired by the observation of one’s own family, then “empowering” and “fatalistic” beliefs may be inherited, and contribute to the persistence of inequality – a “culturally driven” inequality trap. More recently, others have suggested that beliefs about the importance of personal effort (*vis-à-vis* luck or circumstances) may affect not only how hard one works, but also how one votes for redistribution, leading to different social contracts in different societies.¹⁸

Empirically, there is already some evidence that beliefs about one’s own abilities affect behavior and rewards. Steele’s [58] work on “stereotype threat” has documented the impact on African Americans of negative perceptions about their own abilities in specific domains (for example in relation to academic versus athletic ability.) Hoff and Pandey’s [41, 42] work on *dalits* in India suggests that lower-caste children perform cognitive tasks unrelated to their social position (such as completing a maze) less well when they are explicitly reminded of their inferior status, than when they are not. Appadurai [5] suggests that such culturally shaped processes may lead to reduced ambition and aspiration among those that share them, thereby further reducing effort and achievement, and leading to an apparent confirmation of the original beliefs. This is a new variant of the long-known phenomenon of self-fulfilling beliefs, which may contribute to the existence of inequality traps.

But relatively little is known about the scope of the problem: do Afro-Brazilians suffer from similar ‘stereotype threats’ to those that appear to afflict African Americans? Does performance loss caused by the salience of caste differences persist over time, or do *dalits* “get over it” as they grow older? To what extent are drug-abuse, gang-membership and violence “escape strategies” adopted by some members of disadvantaged groups to deal with stereotypes of inferiority in ‘mainstream’ activities? Once the Pandora’s Box of the effect of culture and beliefs on education, investment, work and other economic behaviors has been opened, the questions are endless. In economic terms, this involves exploring how preferences are formed by social processes. This is in contrast to the core assumption in most positive and normative economic theory that individual preferences are primitive and not to be questioned.

4.3 Efficiency costs

Directly complementary to research on the causes of inequality traps, is the theoretical and empirical exploration of their costs. As suggested above, we believe there may be a significant class of cases in which inequality traps lead to suboptimal outcomes with respect to either or both of the equitable development objective and efficiency. Work on this area

¹⁸ See, e.g., Alesina and Angeletos [3] and Bénabou and Tirole [17].

would build on existing literatures, notably on the efficiency costs of wealth inequalities, and on political processes favoring particular groups.

Thus the literature on rent-seeking societies (started by Krueger [45]) is relevant, to the extent that the economic institutions of rent-seeking are produced by unequal political and social processes. As in many areas, documentation of the costs associated with specific mechanisms of influence will be of particular interest. For example, in their studies of the Mexican financial system Haber and Kantor [37] and Haber and Maurer [38], not only document the processes whereby capture shaped institutional design, ownership and incentive structures, but explore the costs in terms of suboptimal lending and the greater vulnerability of the financial system, that contributed to the very large costs associated with the 1994–1995 Tequila crisis.

Rigorous assessment of costs again raises empirical challenges. For example, sensible estimation of counterfactual incomes along the distribution if the trap were absent might prove difficult. However, creative identification strategies should allow estimates both of losses incurred by the subordinate group (such as lower caste children who drop out of school), and of aggregate efficiency effects. As an example Banerjee et al. [13], explore why sugar cooperatives in Maharashtra experienced lower yields and slower growth in the fertile eastern region of the state compared with the arid western region between 1971 and 1993. They find that greater inequality in the east confers control rights on the wealthier members of the coops, providing incentives for them to set low prices for their members and divert rents to themselves. In the more egalitarian western region, sugar prices are set at higher levels and provide stronger production incentives.

4.4 Institutional change and policy design

The fourth and final research area we highlight in this article is the analysis of potential transitions between equilibria with and without an inequality trap, and the associated *processes of institutional change*. In relation to the account in Section 2, a society can potentially shift out of an inequality trap path for three categories of reason: external shocks, changes in policy, or changes in efforts. An external shock, such as market opening or shifts in global factor prices, could change the distribution of wealth and the bargaining power of different groups. Policy is partly endogenous, of course, and we are interested in the exogenous component – the “policy space” that may occur when a reformist faction comes into power, albeit within a broader context of unequal institutions. Effort is related to socio-cultural processes, and also has both an endogenous and exogenous components; we are here interested in the scope for purposive social processes to alter the pattern of effort, in areas ranging from school attendance to demands for government service delivery. There is a rich research agenda in all three areas and we focus on two categories: first, the exploration of transitions away from inequality traps, that could involve change in any or all of these areas; second, for cases where the political configuration enables action to escape an inequality trap, the analysis of policy choices that will further this end.

To explore the first question, an appropriate empirical approach might be to study societies and communities that have historically evolved from low-level equilibria, seeking to understand how they broke their own inequality traps. The WDR 2006 attempts a variant of this approach by presenting very brief summaries of historical transitions in Spain and Scandinavia – societies that were at some point in the past characterized by (relatively) low means and high inequality (in incomes and opportunities, it is argued), but which have transitioned to higher-mean, lower-inequality equilibria.

It was probably our ignorance, but we did not find many insightful accounts of institutional development in poorer countries which appear to have made similar transitions, such as Korea or Taiwan.¹⁹ Careful studies of how institutions changed and evolved in developing countries that experienced such transitions from widespread poverty and inequality to more prosperous and egalitarian societies would clearly be of great value. Equally important are detailed studies of specific institutional reforms at a local level, where there might be more hope of ascertaining impact a little more clearly. Recent analyses of reservations in Indian *panchayats* (local government assemblies), through which a fraction of seats is reserved for women or *dalits*, may be the best examples of this kind of research.²⁰ Another example is a study of city-level participatory budgeting in Brazil, that combines quantitative analysis of impacts (using a regression discontinuity approach as an identification strategy) with a range of in-depth “qualitative” approaches to illuminate the processes at work.²¹ There must be many other examples of equalizing institutional transitions across the developing world, a greater understanding of which would be of great value.

The second question concerns situations in which political equilibria shift sufficiently to bring in governments – or reform groups within governments – that want to pursue policies that will further the objective of equitable development. Now the question becomes a more practical one, of how best to break through inequality traps, given some political room for policy choice. However, this is not a purely technical decision, since “progressive” governments will often have to take account of political economy and socio-cultural considerations, so as to avoid well-intentioned policy designs being (re)captured by dominant groups during implementation, or failing for some other reason.

The question of the efficacy of policy interventions takes us to an area where there is a long tradition of work. In the last few years, there has been a particular emphasis on the use of evaluation techniques that allow robust identification of the impact of specific interventions (with randomized trials as the gold standard). A focus on inequality traps brings some specific angles to this category of work, that have been relatively underemphasized in the past. These flow directly from the concept of an inequality trap and the discussion of causative processes outlined above. When a trap exists, we will be interested in evaluation of an intervention in terms of the extent to which it causes changes in the dynamics of advantage for a group *through* action affecting the mechanisms creating inequality-preserving (formal or informal) institutions. In principle, this could involve any part of the initial distribution of advantage; we illustrate with cases affecting the bottom and the top of the distribution.

Interventions focusing on the disadvantaged include a wide array of “anti-poverty” measures. Examples range from conditional cash transfers that are targeted to poor families and conditional on specific behaviors, such as school attendance (like *Oportunidades* in Mexico, or *Bolsa Familia* in Brazil); to early childhood development interventions, like center-based care or family visits by social workers. They might also include micro-credit schemes based on group lending, so as to harness peer-monitoring as an instrument to alleviate both adverse selection and moral hazard problems. There are interventions to reduce teacher-absenteeism and to generally increase learning rates at public schools; health interventions based on recruiting community nurses to increase basic knowledge of

¹⁹ Heller [39] is an example of an account of the social and political transition in Kerala, India.

²⁰ See, for instance, Chattopadhyay and Duflo [27] and Besley et al. [20].

²¹ See Baiocchi et al. [10].

nutrition and hygiene among poor populations; community-based sports and arts programs to reduce the attractiveness of violent gang membership; “mobile courts” designed to reduce the transaction costs involved in resolving simple legal disputes; and so on. A somewhat different category of intervention concerns the role of information, including the media, in putting pressure on governments to pursue more pro-poor action, for example with respect to drought relief in India [18] and adoption of community health workers in the North East of Brazil [59].

With respect to the top of the distribution of advantage, we are interested in the policy implications of the concentration of power in dominant groups. Past work suggests at least two lines of research that are likely to be fruitful in these areas. First, there is work on the conditions under which incumbent influence can be reduced. An example concerns the role of both conscious choice and the influence of external shocks (including from international competition) in the design of financial systems that are both efficient and resilient to capture.²² Similar issues concern the design of the regulatory framework for the privatization and running of utilities with natural monopoly power. Second, there is work concerned with how policy design can affect the interaction between local inequalities and a range of interventions, from management of the commons to local “elite capture” of delivery of services and transfers.²³

The problem facing a progressive government, or an institution that has evolved to a situation where there is genuine political will to combat an inequality trap, combines technical, political and socio-cultural questions. Which of these policies work well? Which do not? Which have the highest benefit to cost ratio? Does the fact that a particular intervention works in India mean that it will also work in Zambia or Bolivia? More generally, how do interventions interact with the national and local political and social context? These questions have implications for methodology, and we emphasize two here. First, there will be great value to bringing together three kinds of research work which ought to be – but seldom are – intimately related: ex post impact evaluation; ex ante impact evaluation; and cost–benefit analysis. And second, *within* each of these areas, there will be a need to bring both quantitative and qualitative techniques to bear on the group-based, relational and institutional issues that are central to inequality traps – in a particular application of the recent emphasis on bringing together “quantitative” and “qualitative” techniques [43, 52].

Ex post impact evaluation is currently immensely popular in economics. The “discovery” of randomized experiments in the discipline has been accompanied by a very large number of studies that seek to rigorously evaluate the impact of particular programs (such as a cash transfer) on a set of measurable outcomes (such as school enrollment, family consumption and mother’s labor supply). While this is a very welcome development, these studies are seldom accompanied by three complementary kinds of analysis which are crucial for real-world policy usefulness: long-term follow-up; the modeling of behavior, and cost–benefit analysis. Long-term follow-up is particularly important for policies that seek to expand the opportunity sets of today’s children, and whose full impacts may not be felt until these children are adults, and in fact until they have children of their own.

²² See Rajan and Zingales [49] for a synthesis, and Claessens and Perotti [29] for a review of the political economy literature and discussion of policy implications.

²³ See Baland and Platteau [11, 12] on inequality and management of the commons, and Bardhan and Mookherjee [14, 15] on the general question of capture and decentralization.

The modeling of behavior in response to a policy, which is sometimes described as ex-ante policy evaluation, can help address questions of external validity. A differences-in-differences estimate of the impact of a particular intervention, based on randomly assigned treatment and control groups in a set of villages in Western Kenya is a fantastically important result... largely for villagers in Western Kenya. Unless the treatment in question is medical, and addresses conditions in which human beings are not thought to differ a great deal from place to place, the results of an experiment in one place are not terribly meaningful for potential adopters scrutinizing the results elsewhere. When outcomes are more complex than simple medical issues (such as whether an iron supplement increases productivity, or whether a de-worming pill works), and when impacts are mediated by local institutions – even the peculiarities of local markets – then experimental impact estimates are very difficult to generalize. They are valid locally, but not externally. This is where an understanding of *how* the observed impacts were attained is a particularly useful addition to a convincing measure of *what* impacts were achieved. Such an understanding typically requires modeling agents' behavior in response to the policy.²⁴

Another important but often-neglected aspect is the quantification of the costs and benefits of individual policies. It is plausible that most impact measures of a cash transfer would rise as the amount transferred increases. The reason no one advocates increasing these amounts *ad infinitum* is that policies have costs as well as benefits. For a policymaker, the correct estimate of the impact of a policy (e.g., enrollment among girls increases by 5%) is only the first step towards a decision. It is clearly necessary, but equally clearly insufficient. A good policymaker would also need to know how much it costs to increase enrollment by 5%, and what the benefits are. Although cost–benefit analysis was once very popular among development economists, it is less fashionable today.²⁵ Although the reasons – which involve how easy it is to attack the various assumptions one must inevitably make when estimating costs and benefits, sometimes far into the future – are understandable, the outcome is, nevertheless, regrettable.

5 Conclusions

In this paper, we offered some formal definitions of the concepts of equity, equality of opportunity and efficiency which were used in the World Development Report 2006, clarifying how they relate to John Roemer's original concepts. We defined an *equitable development policy* as the policy that maximizes the present value of the stream of advantage for the least-privileged group in society, subject to two important constraints: that extreme deprivation is eliminated, and that policies belong to a permissible set (that excludes not only technically infeasible but also socially unacceptable interventions). Such a policy objective does represent something of an evolution in development thinking, as it moves poverty reduction from an objective to a constraint. This obviously does not reduce its priority or urgency, since the constraint must be complied with, but it extends the policy aim beyond it.

²⁴ Todd and Wolpin [61] provide a good example of using a structural model of household behavior to analyze household responses to a conditional cash transfer. They then use data from the Progresa randomized experiment in order to test their model. When the model is able to predict experimental results reasonably well, it can also be used to investigate other settings in which the experimental results may be valid. The combination of experimental methods with behavioral modeling can thus provide cross-validation that greatly enhances the usefulness of both.

²⁵ Except for a few exceptions, such as Angrist et al. [4] and Thirumurthy et al. [60].

In response to Roemer [54], we clarified that the main message of the report was *not* that pursuing an equal opportunity policy – or indeed our equitable development policy – would attain maximum economic efficiency (in the Kaldor–Hicks sense). It was that many developing countries are likely to be trapped in low-level equilibria which are both unequal and inefficient, so that well-designed movements towards equitable development might well increase (although not maximize) efficiency.

We described a particular class of low-level equilibria as *inequality traps*: situations in which a disadvantaged group faces a long-run opportunity set (defined by the group's advantage distribution) that is worse than another ("dominant") group's, when in a feasible alternative equilibrium this is not the case. In very general terms, the trap arises through the persistence of social, economic and political inequalities that link a person's circumstances and efforts to those of previous generations, and make policy choices themselves reflect unequal distributions of power. Some examples were discussed, and two implications identified: inequality traps contribute to the persistence of unequal opportunities and may have efficiency costs in aggregate terms.

Although there is a growing body of empirical evidence on links between distribution (of wealth, power and status) and various outcomes, we argued that there are four broad research areas where knowledge is scant, and scope for innovative work is high. The first area is the measurement of inequality of opportunity itself, and the diagnosis of inequality traps in terms of the mobility patterns of advantaged and disadvantaged groups.

The second area is the identification of the causal processes that create and reproduce inequality traps. In particular, we singled out political processes of institutional capture (whether at the national or local level, and whether institutions are legislative, judicial, or executive); and cultural processes affecting group-based interactions, such as beliefs of inferiority becoming self-fulfilling by affecting the behavior of members of disadvantaged groups. The third area is the quantification of the aggregate efficiency costs that may arise from the reduced opportunities available to the disadvantaged. To the extent that these forgone economic opportunities are not compensated by gains of comparable magnitudes to other, better-off groups, they will entail a loss in a Kaldor-Hicks sense. Better quantification of these losses, wherever they exist, add to the arguments for breaking inequality traps.

The fourth area combines the understanding of processes of institutional change, through which societies transit out of inequality traps to superior equilibria, with the assessment of specific policy interventions aimed at expanding the opportunity sets of the disadvantaged. This is a very broad area, which can – and should – combine a variety of methods: from long time-span historic accounts at the country level to impact assessments of (quasi-) exogenously assigned institutional reform at the local level. It must also be complemented by better evaluation of the actual policies that these evolving institutions (including governments) can adopt, in order to pursue their objectives, such as equitable development. We had three specific senses in which evaluations should be "better": they should be long-term, following recipients (and controls) for much longer periods than is typically done today; they should combine experimental (or quasi-experimental) methods with modeling of individual response behavior and processes of social change, with cross validation, in order to increase external validity; and they should seek to value costs and benefits of interventions.

As more and better data becomes available in developing countries, and young researchers apply themselves to these questions – and others that we have not thought about – we are certain that the next World Development Report that focuses on the links between equity and development, years from now, will have a much broader and deeper theoretical and empirical basis than the 2006 vintage.

Acknowledgment The authors are grateful for comments from Karla Hoff. The views presented in this paper are exclusively those of the authors, and should not be attributed to the World Bank or any other institution.

References

1. Acemoglu, D., Robinson, J.: Economic origins of dictatorship and democracy. Cambridge University Press, Cambridge, MA (2005)
2. Acemoglu, D., Johnson, S., Robinson, J.: The colonial origins of comparative development: an empirical investigation. *Am. Econ. Rev.* **91**(5), 1369–1401 (2001)
3. Alesina, A., Angeletos, G.-M.: Fairness and redistribution. *Am. Econ. Rev.* **95**(4), 960–980 (2005)
4. Angrist, J., Bettinger, E., Bloom, E., King, E., Kremer, M.: Vouchers for private schooling in Colombia: evidence from a randomized natural experiment. *Am. Econ. Rev.* **92**(5), 1535–1558 (2002)
5. Appadurai, A.: The capacity to aspire: culture and the terms of recognition. In: Rao, V., Walton, M. (eds.) *Culture and Public Action*. Stanford University Press, Stanford, CA. (2004)
6. Araujo, M., Caridad, F., Francisco, L., Peter, Özler, B.: Local inequality and project choice: theory and evidence from Ecuador. World Bank Policy Research Working Paper Series #3997 (2006)
7. Arneson, R.: Equality of opportunity for welfare. *Philos. Stud.* **56**, 77–93 (1989)
8. Atkinson, A.B., Stiglitz, J.E.: Lectures on Public Economics. McGraw-Hill, London (1980)
9. Attanasio, O., Meghir, C., Santiago, A.: Education choices in Mexico: using a structural model and a randomized experiment to evaluate progressa. University College London, Processed (2001)
10. Baiocchi, G., Chaudhuri, S., Heller, P.: Evaluating Empowerment: Participatory Budgeting in Brazil. World Bank, Washington, DC Processed (2005)
11. Baland, J.M., Platteau, J.P.: Wealth inequality and efficiency in the commons. Part I. The unregulated case. *Oxf. Econ. Pap.* **49**(4), 451–482 (1997)
12. Baland, J.M., Platteau, J.P.: Wealth inequality and efficiency in the commons. Part II. The regulated case. *Oxf. Econ. Pap.* **50**(1), 1–22 (1998)
13. Banerjee, A., Moohkerjee, D., Munshi, K., Ray, D.: Inequality, control rights, and rent seeking: sugar cooperatives in Maharashtra. *J. Polit. Econ.* **109**(1), 138–190 (2001)
14. Bardhan, P., Mookherjee, D.: Capture and governance at the local and national levels. *AEA Papers and Proceedings* **90**(2), 135–139 (2000)
15. Bardhan, P., Mookherjee, D.: Decentralizing antipoverty program delivery in developing countries. *J. Public Econ.* **89**, 675–704 (2005)
16. Bénabou, R.: Unequal societies: income distribution and the social contract. *Am. Econ. Rev.* **90**(1), 96–129 (2000)
17. Bénabou, R., Tirole, J.: Belief in a just world and redistributive politics. *Q. J. Econ.* **121**(2), 699–746 (2006)
18. Besley, T., Burgess, R.: The political economy of government responsiveness: theory and evidence from India. *Q. J. Econ.* **117**(4), 1415–1451 (2002)
19. Besley, T., Coate, S.: Lobbying and welfare in a representative democracy. *Rev. Econ. Stud.* **68**, 67–82 (2001)
20. Besley, T., Pande, R., Rahman, L., Rao, V.: The politics of public good provision: evidence from Indian local governments. *Journal of the European Economic Association* **2**, 416–426 (2004)
21. Betts, J., Roemer, J.: Equalizing Opportunity through Educational Finance Reform. mimeo, Public Policy Institute of California, San Francisco, CA (1999)
22. Bourdieu, P.: *The Logic of Practice*. Stanford University Press, Stanford (1990)
23. Bourguignon, F., Verdier, T.: Oligarchy, democracy, inequality and growth. *J. Dev. Econ.* **62**(2), 285–313 (2000)
24. Bourguignon, F., Ferreira, F., Leite, P.: Conditional cash transfers, schooling, and child-labor: micro-simulating Brazil's *Bolsa Escola* program. *World Bank Econ. Rev.* **17**(2), 229–254 (2003)
25. Bourguignon, F., Ferreira, F., Menéndez, M.: Inequality of outcomes and inequality of opportunities in Brazil. World Bank Policy Research Working Paper Series 3174 (2003)
26. Buchanan, J.M.: A Hobbesian interpretation of the Rawlsian difference principle. *Kyklos* **29**(1), 5–25 (1976)
27. Chattopadhyay, R., Duflo, E.: Women as policy makers: evidence from a randomized policy experiment in India. *Econometrica* **72**(5), 1409–1443 (2004)
28. Checchi, D., Peragine, V.: Regional disparities and inequality of opportunity: the case of Italy. IZA Discussion Paper No.1874 (2005)
29. Claessens, S., Perotti, E.: The links between finance and inequality: channels and evidence. Background paper for the WDR 2006 (2005)

30. Cohen, G.A.: On the currency of Egalitarian justice. *Ethics* **99**, 906–944 (1989)
31. Dworkin, R.: What is equality? Part 2. Equality of resources. *Philos. Public Aff.* **10**(3), 185–246 (1981)
32. Engerman, S., Sokoloff, K.: Factor endowments, inequality and paths of development among new world economies. *Economia* **3**(1), 41–110 (2002)
33. Engerman, S., Sokoloff, K.: The evolution of suffrage institutions in the new world. *J. Econ. Hist.* **65**(4), 891–921 (2005)
34. Ferreira, F.H.G. Education for the masses? The interaction between wealth, educational and political inequalities. *Econ. Transit.* **9**(2), 533–552 (2001)
35. Galasso, E., Ravallion, M.: Decentralized targeting of an antipoverty program. *J. Public Econ.* **89**, 705–727 (2005)
36. Grossman, G., Helpman, E.: Protection for sale. *Am. Econ. Rev.* **84**(4), 833–850 (1994)
37. Haber, S., Kantor, S.: Getting Privatization Wrong: The Mexican Banking System, 1991–2003. Stanford University, Stanford, CA Processed (2004)
38. Haber, S., Maurer, N.: Related Lending and Economic Performance: Evidence from Mexico. Stanford University, Stanford, CA Processed (2004)
39. Heller, P.: The Labor of Development: Workers and the Transformation of Capitalism in Kerala, India. Cornell University Press, Ithaca, NY (1999)
40. Hoff, K.: Paths of Institutional Development: A institutional development: a view from economic history. *World Bank Res. Obs.* **18**(2), 205–226 (2003)
41. Hoff, K., Pandey, P.: Belief systems and durable inequalities: an experimental investigation of Indian caste. Washington, D.C.: World Bank Policy Research Working Paper Series 2875 (2004)
42. Hoff, K., Pandey, P.: Economic consequences of social identity: discrimination, social identity and durable inequalities. *AEA Papers and Proceedings*, May 2006, 206–211 (2006)
43. Kanbur, R. (ed.): Qualitative and quantitative poverty appraisal: Complementarities, tensions and the way forward. Delhi: Permanent Black (2003)
44. Khan, B. Zorina, S., Kenneth L.: Patent institutions, industrial organization and early technological change: Britain and the United States. In: Berg, M., Bruland, K. (eds.) *Technological Revolutions in Europe*. Elgar, Cheltenham, UK (1998)
45. Krueger, A.O.: The political economy of the rent-seeking society. *Am. Econ. Rev.* **64**(3): 291–303 (1974)
46. Nozick, R.: *Anarchy, State and Utopy*. New York: Basic (1974)
47. Piketty, T.: Social mobility and redistributive politics. *Q. J. Econ.* **CX**(3), 551–584 (1995)
48. Pistoiesi, N., Lefranc, A., Trannoy, A.: Inequality of opportunities vs. inequality of outcomes: are Western societies all alike?. Mimeo, Université de Cergy-Pontoise (2005)
49. Rajan, R.G., Zingales, L.: Saving Capitalism from the Capitalists: Unleashing the Power of Financial Markets to Create Wealth and Spread Opportunity. Crown Business, New York (2003)
50. Rajan, R.G., Zingales, L.: The persistence of underdevelopment: institutions, human capital, or constituencies? National Bureau of Economic Research Working Paper 12093 (2006)
51. Rao, V.: On ‘inequality traps’ and development policy. *Development Outreach*, February, 10–13 (2006)
52. Rao, V., Woolcock, M.: Integrating qualitative and quantitative approaches in program evaluation. In: Bourguignon, François, Pereira Da Silva, Luiz (eds.) *Tool Kit for Evaluating the Poverty and Distributional Impact of Economic Policies*. World Bank, Washington, DC (2003, Oxford University Press)
53. Roemer, J.E.: *Equality of Opportunity*. Harvard University Press, Cambridge, MA (1998)
54. Roemer, J.E.: The 2006 World Development Report: equity and development – a review. *Journal of Economic Inequality* **4**, 233–244 (2006).
55. Roemer, J.E., Aaberge, R., Colombino, U., Fritzell, J., Jenkins, S. P., Marx, I., Page, M., Pommer, E., Ruiz-Castillo J., Segundo, M.J.S., Traanes, T., Wagner, G., Zubiri, I.: To what extent do fiscal regimes equalize opportunities for income acquisition among citizens? *J. Public Econ.* **87**, 539–565 (2003)
56. Sen, A.: Equality of what? In: McMurrin S. (ed.) *Tanner Lectures on Human Values*. Cambridge University Press, Cambridge, MA (1980)
57. Sen, A.: *Commodities and Capabilities*. North Holland, Amsterdam (1985)
58. Steele, C.M.: Thin ice: ‘stereotype threat’ and black college students. *Atl. Mon.* **284**(2), 44–54 (1999)
59. Tendler, J.: Good government in the tropics. Johns Hopkins University Press, Baltimore (1997)
60. Thirumurthy, H., Zivin, J., Goldstein, M.: The economic impact of aids treatment: labor supply in Western Kenya. World Bank.(2006, Processed)
61. Todd, P.E., Wolpin, K.I.: Ex-ante evaluation of social programs. University of Pennsylvania (2005, Processed)
62. Van de Gaer, D., Schokkaert, E., Martínez, M.: Three meanings of intergenerational mobility. *Economica* **68**, 519–537 (2001)
63. World Bank: *World Development Report 2006: Equity and Development*. World Bank, Washington, DC (2005, Oxford University Press)