Moriki Hosoe · Biung-Ghi Ju ·
Akira Yakita · Kiseok Hong *Editors*

Contemporary Issues in Applied Economics

Ten Years of International Academic Exchanges Between JAAE and KAAE

韓國應用經濟學會

JAAE Since 2006

Springer

Contemporary Issues in Applied Economics

Moriki Hosoe · Biung-Ghi Ju ·
Akira Yakita · Kiseok Hong
Editors

Contemporary Issues in Applied Economics

Ten Years of International Academic Exchanges Between JAAE and KAAE

Springer

Editors
Moriki Hosoe
Faculty of Economics
Kyushu University (Emeritus Professor)
Fukuoka, Japan

Akira Yakita
School of Economics
Nanzan University
Nagoya, Japan

Biung-Ghi Ju
Seoul National University
Seoul, Soul-t'ukpyolsi, Korea (Republic of)

Kiseok Hong
Department of Economics
Ewha Womans University
Seoul, Soul-t'ukpyolsi, Korea (Republic of)

ISBN 978-981-13-7035-9 ISBN 978-981-13-7036-6 (eBook)
https://doi.org/10.1007/978-981-13-7036-6

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

This book is the collected papers in commemoration of the tenth anniversary of academic exchange between the Japan Association for Applied Economics (JAAE) and the Korean Association for Applied Economics (KAAE). With the name of applied economics, both academic societies can be called sister ones. The international exchange of the two academic societies has reached a milestone of 10 years, since several researchers from KAAE joined at the JAAE meeting in the autumn at Chuo University. We were able to deepen academic exchanges by mutual presentation and discussion. We would like to express my gratitude to Prof. Hang Keun Ryu of Chung Ang University and Prof. Tamotsu Nakamura of Kobe University for their great contribution to the development of our international exchange. A feature of this book is to analyze various contemporary issues in applied economics, distinguished by advanced theoretical research and empirical analysis based on Japan and Korea. In particular, this book consists of four parts. In Part I, we investigate the fields of economic development, growth, and welfare, using tools of applied analysis. In Part II, the fields of inequality, redistribution, and intergenerational transfers are studied mainly on an empirical basis. In Part III, we focus on the fields of public policy and political economics. In last part the fields of resource and environmental economics are investigated on the basis of the data of Japan and Korea. We hope that two associations will promote the further cooperation of academic exchange on the basis of this project. Lastly, we wish to express our special thanks to Prof. Tohru Naito of Doshisha University for his assistance in editing this book.

Seoul, Korea (Republic of) Biung-Ghi Ju
Nagoya, Japan Akira Yakita
Seoul, Korea (Republic of) Kiseok Hong
Fukuoka, Japan Moriki Hosoe
January 2019

The original version of the book was revised: Belated corrections have been incorporated. The correction to the book is available at https://doi.org/10.1007/978-981-13-7036-6_22

Contents

Part I
Economic Development, Growth, and Social Welfare

Chapter 1
Improvement in Living Standards in the 1960–1970s and Economic Development in Korea

Sok Chul Hong

1.1 Introduction

Korea achieved a remarkable economic growth in the 1960–1970s after its economic foundation was ruined by Korean War (1950–53). This rapid growth is generally called as the 'Miracle on the Han River'. During the miraculous period, Korea achieved not only economic growth, but also a remarkable improvement in the standards of living including control of diseases, expansion of educational opportunities, increase in consumption, reduction of inequalities and so on. The improvement in living standards in the 1960–1970s has been a significant foundation that has led to a successful economic development today.

It is intriguing to know that living conditions or living standards in the United Kingdom and United States in the 19th century, which succeeded in industrialization and urbanization during the period, were as low as those in Korea right after the Korean War. They could not achieve the advancement in living standards until the early 20th century. Then, they could build welfare nations by succeeding in disease control due to advance in medical knowledge and investment in public health, and by making substantial investments in public sectors such as education and welfare.

The improvement in living standards in Korea during the 1960–1970s did not result from introducing new technologies, policies and system, which were not previously observed in other countries. But it was the most significant that Korea adopted policies and system, considering Korea-specific situation, that were successfully established in developed countries in the first half of the 20th century. In addition, the improvement could be achieved within a short period largely because of strong driving force by Korea government, social consensus and national participation, and implementation of effective policies and system.

S. C. Hong (✉)
Seoul National University, Seoul, Republic of Korea
e-mail: sokchul.hong@snu.ac.kr

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_1

The purpose of this study is to examine how Korea achieved the remarkable improvement in living standards in the 1960–1970s and discussing its background from various aspects. In Sect. 1.2, we briefly review the significance of living standards for better economic growth. We also discuss how researchers have developed the ways of measuring the standards of living by taking historical examples of improvement in living standards in the United Kingdom and United States. In Sects. 1.3–1.5, we quantitatively and qualitatively examine the process in which higher standards of living had been achieved in Korea in the 1960–1970s. We particularly study long-term improvements in disease control, educational attainment, female labor force participation, national and household expenditure, income inequalities, disparities in living conditions between urban and rural areas and so on. We also discuss various policies and system in the 1960–1970s that resulted in the improvements. In Sect. 1.6, we conduct cross-country analysis to see how the improvement in living standards in the 1960–1970s has affected today's economic status. The estimation result suggests the significance of enhanced living standards in economic growth. Finally, we discuss the implications for today's developing countries from Korean experiences in the 1960–1970s.

1.2 Significance of Living Standards and Historical Experiences in Developed Countries

Although[1] direct goal of economic activities is to earn wage and income, the ultimate goal is to have a satisfactory and opulent life through consumption and investment on the basis of earned income. However, increase in income and consumption does not always led to such a satisfactory life. The adequate supply of public goods for better education, environment, public health and welfare is the prerequisite condition for better life. The supply of public goods would be the main role of the governments. Accordingly, many countries have tried to seek efficient ways to measure the average and variation of living standards among populations and in societies. They have established policies and constructed national system to improve population's living standards by analyzing the long-term changes of various measures of living standards and their factors.

The standard of living discussed in the study is another description of the quality of life. How to measure the living standards began to be studied from the early 19th century. But living standards were qualitatively measured in early studies by describing people's living conditions mainly due to lack of data and analyzing tools. One of the earliest study was done by a 19th-century economist, Friedrich Engels. He examined the living standards among laborers in London by describing their income, expenditure, diet, life expectancy and so on (Engels 1845).

[1] The discussion in this section on a history of measuring the standards of living and British industrial revolution is largely based on the study by Hong (2013a, b).

Researchers began to systematically study the measures of living standards from the early 20th century because large-scale surveys and data collection were available in the period. The measure that economists first focused on is a material variable such as real wage. But there were measuring errors in estimating real wage because data on price were less reliable. In addition, the difference in the distribution of occupations and the level of industrialization across countries did not allow researchers to use real wage for comparing the standards of living across countries. Simon Kuznets largely resolved such problems by introducing the concept of national income in the 1940s. GDP per capita, which has been developed from the earlier concept of national income, is a representative measure of living standards today (Kuznets et al. 1941).

However, the use of material variables, wage and income, have been criticized because they less reflect non-material aspects in the quality of life. Such criticism has increased the efforts to develop new measure of living standards that captures not only material conditions but also non-material conditions. Most of all, a long debate on the standards of living in the period of British industrial revolution has contributed in extending the scope of living-standard research and developing more advanced measures of living standards (Hong 2013a, b).

As shown in Fig. 1.1, real wage rapidly increased after the 1820s when British was at the center of the industrial revolution. Accordingly the trend of real wage, which is a measure of material conditions, suggests that the standard of living was improving in Britain during the early 19th century. However, more recent studies have showed that non-material living conditions declined during the period when real wage are thought to have increased. For example, life expectancy, which is considered as a representative measure of non-material living conditions, did not improve during the period of industrialization. The life expectancy is estimated around 41 for the 1760s when the British industrial revolution began, but it was similar in the 1850s, when the revolution was completed. Poor sanitation and disease environments caused by industrialization and urbanization were the key factors that did not make any progress in life expectancy even if Britain experienced a substantial increase in income.

An increasing number of studies have tried to develop the comprehensive measures of living standards. Studies on average adult height among populations are worthy to note (Steckel and Roderick 1997). In particular, the use of adult height is useful when other data on material and non-material variables are less available. Adult height largely depends on net nutritional status in infancy, childhood and adolescence. Net nutritional status is determined by nutritional supply like diet or caloric intake and nutritional demand (or energy consumption) due to infections and labor. Generally speaking, nutritional supply is related with average income and agricultural productivity. Energy consumption is affected by non-material conditions such as sanitation, disease environment and labor intensity. Therefore, average adult height can be a comprehensive measure of living standards capturing material and non-material aspects of living standards.

The left-lower panel in Fig. 1.1 shows that average adult height substantially declined in Britain after the 1820s. This strongly suggests that industrialization reduced the standards of living causing poor sanitation, disease environment and

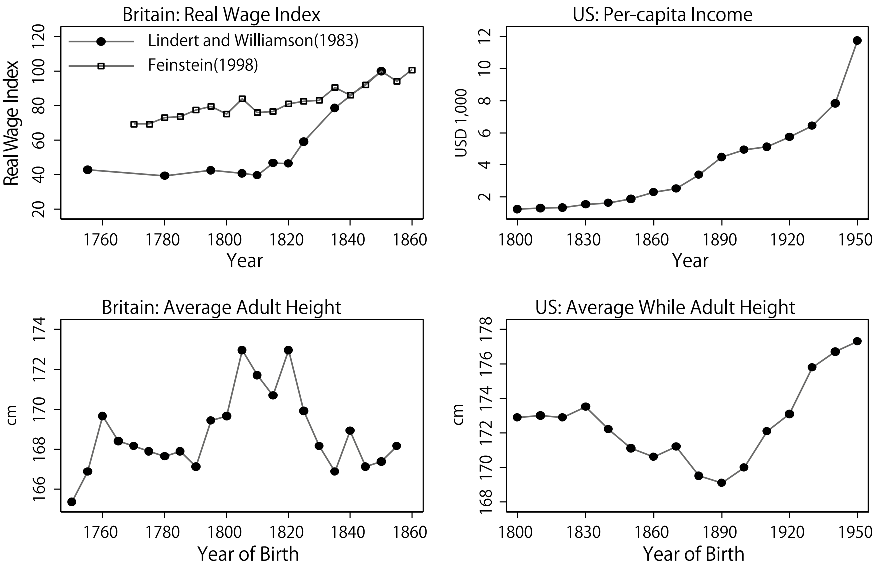


Fig. 1.1 Trends of income and adult average height during industrialization in Britain and the United States. *Sources* British real wage index—Lindert and Jeffrey (1983), Feinstein (1998); British adult average height—Harris (1998); US national income per capita—Carter et al. (2006); US adult average height—Fogel (1986), Carter et al. (2006). *Note* The base year for British real wage index is 1880. British adult average height was estimated using adult males aged at 24–29. US national income per capita was calculated in 2000 constant dollar. US adult average height is estimated for native white adult males

child labor although it increased income. The similar pattern is also found from the historical experience of the United States. As seen in the right panel of Fig. 1.1, the United States also experienced a rapid increase in income throughout the mid-19th century, i.e., the period of industrialization. But average adult height, which is a comprehensive measure of living standards, substantially declined during the same period. It has been studied that this decline resulted from lowered productivity in agriculture and reduced per-capita caloric intake due to large immigration from European countries. But it is more extensively accepted that this was caused by wide spread of infectious diseases due to industrialization and urbanization (Floud et al. 2011).

Although Britain and the United States experienced continued increase in income during the 19th century, non-material living standards did not improve until the early 20th century. Most substantial improvements were found in that life expectancy began to be over 50 and considerable investments were made in public sections such as public health, education and welfare. The advance in life expectancy resulted from the success in disease control. The introduction of germ theory and its spread helped to reveal the pathogen of various infectious diseases, and this provided effective ways of preventing and treating infectious diseases. Other factors for successful disease control and life-expectancy increase include the diffusion of medical knowledge and technologies, national campaign for disease eradication, building water and sewer system, and vaccination. Moreover, public investment was not limited only to public health. It is well known that large investments were also made for education and welfare. This made a considerable contribution to improve the quality of life (Lindert 2004).

On the other hand, the efforts for comprehensively measuring the standards of living have been continued throughout the second half of the 20th century (Nordhaus and James 1972; Usher 1980; Fogel 1986; Sen 1999; Schultz 2005). The human development index (HDI), which is estimated and reported by United Nations, is a representative result of the efforts. HDI emphasizes three components that can affect material and non-material living conditions: income, health and education. The index is estimated by putting the equal weight for the three components. As discussed above, income, health and education are the key factors that largely determined the improvement in living standards in Britain and the United States. Thus, today's emphasis of the three aspects in UN HDI suggests that income, health and education still play key roles in determining living standards today. Those factors are also important in that they can improvement next generation's living standards by enhancing human capital and reducing intergenerational transmission of inequalities. Moreover, HDI is useful in comparing living standards across time and countries.

Improvement in living standards in Britain and the United States in the early 20th century was an unprecedented experience that could upgrade human being's quality of life. Their successful experience also contributed to changing living standards in European and North American countries in the first half of the 20th century. Those in Asia, Africa and South America had a chance for improving their living standards in the mid-20th century. Korea is one of the countries which made the chance successful.

1.3 Disease Control and Advance in Health Status Among Populations

The most remarkable advance that Korea experienced is the improvement in health status among populations. Life expectancy at birth is an index to capture population's average health status in a society. A low life expectancy generally suggests that infant or child mortality rate is high. But a low infant or child mortality rate (even at zero rate) does not guarantee the highest life expectancy. In the case, life expectancy relies on how long old people live. Accordingly, life expectancy measures how much populations enjoy healthy life from birth to death.

Korea is one of the few countries which achieved a substantial increase in life expectancy within a short period. For better understanding the achievement, Fig. 1.2 shows the trend of Korean life expectancy at birth from the 1940s to recent years, and the US trend since the 1890s.

The trends show that US life expectancy at birth did not exceed 50 prior to 1900. It is well known that the same pattern is found in other countries which were industrial-

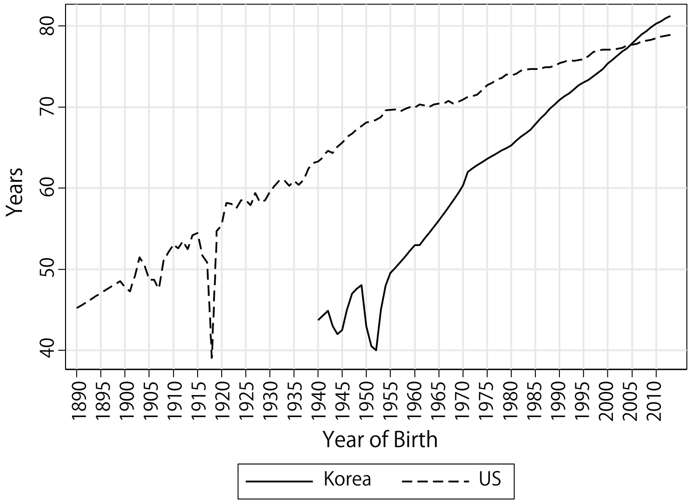


Fig. 1.2 Trends of life expectancy at birth in Korea and the United States. *Sources* Korea—World Bank for before 1961 and Institute for Health Metrics and Evaluation for after 1961; The United States—Carter et al. (2006)

ized in the 19th century. As discussed in the previous section, this is largely because of severe problems of sanitation and environment due to urbanization and the lack of medical knowledge and advance.

Life expectancy in Korea before the mid-1950s was as low as that in the United States in the late 19th century. Korean life expectancy in 1940, when Korea was still in the Japanese colonial era, was around 44; it was only about 43 in 1945 when Korea became independent. Although life expectancy slight increased after the independence, it stayed around early 40s throughout the Korea War. This suggests that the level of sanitation, disease environment and medical technologies was too poor to escape from disease and death as western countries experienced in the 19th century.

Korean life expectancy first exceeded 50 in 1956, and after then Korea experienced a remarkable increase. It is around 1905 that US life expectancy exceeded 50. Thus, Korea in the mid-1950s was about 50 years behind the United States in terms of life expectancy. But it took 15 years that Korean life expectancy reached 60 in 1970. The corresponding year when it reached 60 in the United States was 1930—25 years later from 1950. The gap between Korea and the United States declined from 50 years to 40 years. Then, the gap decreased to 35 years when Korean and US life expectancy reached 70 in 1990 and 1955, respectively.

The experience of developed countries shows that the control of diseases must take precedence in order for life expectancy to reach from 50 to 70. For example, the United States had made a large scale of investment for various public health programs from the turn of the 20th century to the 1930s when its life expectancy reached 60. In particular, it has been studied that campaign for better sanitation, eradication of infectious diseases and construction of infrastructure in public health sectors had played key roles in improving life expectancy (Centers for Disease Control 1999; Cutler and Miller 2005).[2]

As observed in the United States, the control of diseases is the main factor that made life expectancy reach 60 in Korea. The improvement of disease environments particularly played the most role in reducing mortality rate among infants and children because they were the most vulnerable to infectious diseases and environments. This benefit led to the increase in life expectancy. To see how much disease environment was improved in Korea in the 1970s, we review the incidence rate of Class 1 infectious diseases and the infant mortality rate in Fig. 1.3.

Class 1 infectious diseases are defined as those that are spread by water and foot, can easily spread to epidemics and so require urgent preventive actions. They include cholera, typhoid fever, paratyphoid, bacillary dysentery, enterohemorrhagic escherichia coli infections, HAV infections and so on. A large proportion of deaths in the United States resulted from these infectious diseases until the late 19th century, and they are still the most fatal diseases in developing countries even today. Figure 1.3 shows that Class 1 infectious diseases were very prevalent in Korea until the 1970s. For example, the incidence rate of typhoid fever in Korea in 1970 was about 13 per 100,000 populations. The level was as high as the rate of the United States in the 1900s (National Infectious Diseases Surveillance Data; Carter et al. 2006). Figure 1.3 show a remarkable decline in the incidence rate since 1970. The average rate of Class 1 infectious diseases was 16.2 per 100,000 in the 1960s, but it substantially declined throughout the 1970s and to 1.1 by 1980.

The success in controlling diseases is observed across most kinds of infectious diseases. Class 2 infectious diseases include diphtheria, whooping cough, tetanus, measles and so on. They are preventable by vaccination, and so are manageable through national vaccination programs. The average incidence rate of Class 2 infectious diseases in the 1960s was 93.4 per 100,000, but the rate declined to 20.2 in 1980(National Infectious Diseases Surveillance Data). Class 3 infectious diseases contain malaria, tuberculosis, leprosy, scarlet fever and so on, which generally have sporadic outbreaks and so need continued monitoring and preventing efforts. The incidence rate of the infectious diseases declined from 402 per 100,000 in the 1960s to 213 in 1980(National Infectious Diseases Surveillance Data).

[2]As the germ theory was introduced in the later 19th century, various campaigns for improving personal hygiene were conducted. They were very effective in decreasing morbidity rates from infectious diseases (Centers for Disease Control 1999). Typhoid fever is one of the infectious diseases that were more fatal in the United States until the early 20th century. It is well studied that typhoid fever began to be controlled as water and sewer system had been built in large cities since the late 19th century (Ferrie and Werner 2008). Moreover, the United States succeeded in eradicating hookworm in the 1910s and malaria in the 1920, which were prevalent in Southern states with hot and wet climatic conditions. The eradication contributed to not only improving disease environment but also enhancing human capital accumulation among populations (Bleakley 2007; Hong 2013a, b). In summary, campaign for better hygiene, construction of public health infrastructure and control of diseases were the key factors that substantially reduced mortality rates from infectious diseases and considerably improved life expectancy.

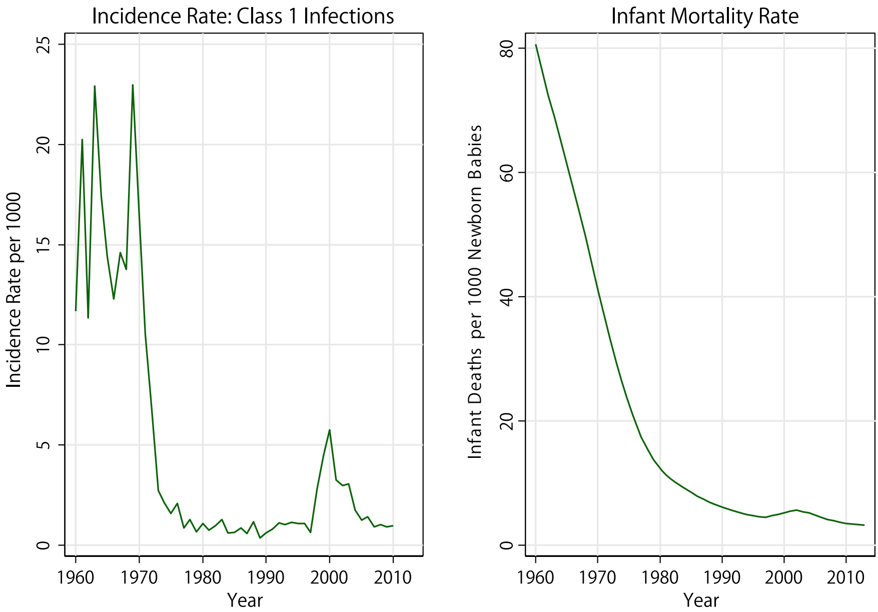


Fig. 1.3 Trends of incidence rate of Class 1 infectious diseases and infant mortality rate. *Sources* Incidence Rate of Class 1 Infectious Diseases—Annual Surveillance Report on Infectious Diseases; Infant Mortality Rate—KOSIS

Infants and children had the most benefit of disease controls in the 1960–1970s. The right panel in Fig. 1.3 shows the trend of infant mortality rate since 1960, which is defined as the number of deaths within one year from birth out of total number of newborn babies. The infant mortality rate in 1960 was 80.6 per 1,000. The rate is corresponding to the infant mortality rate of the United States in 1920(85.5 per 1,000)(Carter et al. 2006). This suggests that the level of sanitation and disease environment in Korea in the 1960s was similar with that in the United States in the 1920s. The infant mortality rate in Korea rapidly declined to 41.2 in 1970 and to 12.4 in 1980. The infant mortality rate in the United States reached 40 per 1,000 in the 1940s and 10s in the 1970s. The US rate had declined from 80 to 10 per 1,000 over 50 years. It is remarkable for Korea to achieve the same only for 20 years. Because infant mortality is closely related with sanitation and infectious diseases, it is highly suggested that there was a substantial investment for the control of diseases in the 1970s.

The effects for the control of disease in Korea can be represented by three aspects: campaign for hygiene and sanitation, active public programs for promoting population health, and investment for public health infrastructure.

First, the campaign for hygiene and sanitation was much emphasized right after the Korea War and in the 1960s. Population was not aware of appropriate medical and public health knowledge for disease prevention and treatment because most

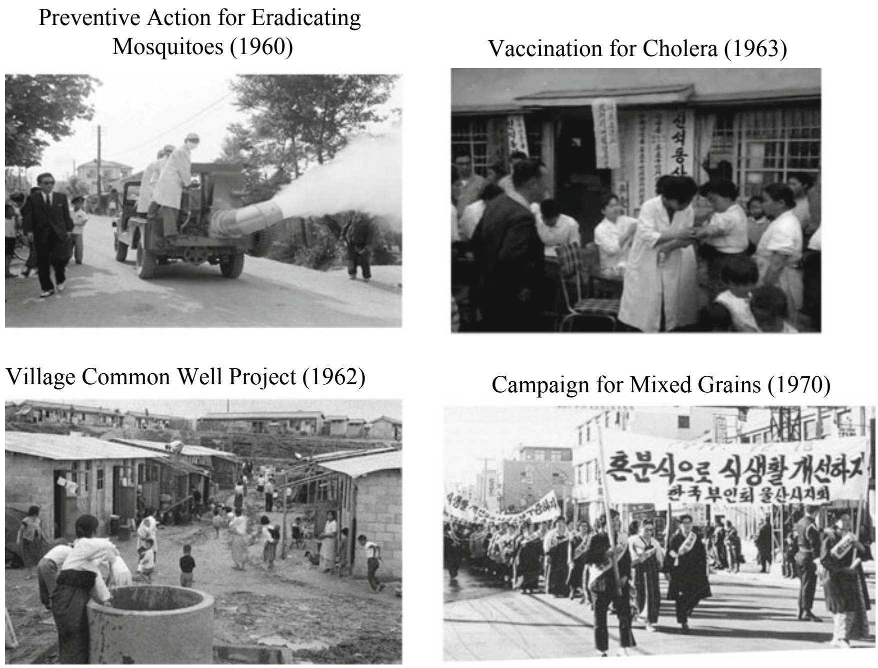


Fig. 1.4 Activities for enhancing population health in 1950–1970s. *Source* Korean National Archives Online DB

educational infrastructure was destroyed during the War. Therefore, it was the most cost-effective policy to conduct campaigns to inform population about the significance and ways of personal hygiene and sanitation. The information was provided to populations through poster, slogan and video records. Especially, the campaign was concentrated during summer, when sanitation became deteriorated.

National video records produced in the 1950–1970s show that such campaigns were wide spread (see Fig. 1.4). For example, various video records for campaign were produced in 1959, 1960, 1967 and 1969 to warn populations against Japanese encephalitis and to suggest that eradicating mosquitoes is the most effect ways for prevention. In 1963, a record was produced and spread across the nation to warn the risk of cholera and inform preventive measures such as vaccination and boiling drinking water. In addition, Korean government frequently conducted campaigns to promote personal hygiene particularly during summer. The content of the campaigns includes washing hands with soap, keeping drain and toilet clean, careful food intake, drinking clean water, which are essential to prevent infectious and contagious diseases.

To make the campaign effective, it is important to induce people to understand the significance of campaigns and to voluntarily participate them. Thus, it is necessary to conduct long-term campaigns and continued advertisement rather than short-term campaigns. In fact, health authorities carried forward various campaigns for hygiene and sanitation persistently throughout the 1970s. As its result, the incidence rate of major infectious diseases substantially declined throughout the periods.[3]

Second, various health-promoting national projects also played important roles in controlling for infectious diseases in the 1960–1970s. Three projects are worthy to discuss: nourishment improvement project, vaccination project, and intestinal parasite eradication project.

Nourishment improvement is essential for preventing diseases among infants, children and adolescence because it can improve their physical growth and immune system. In addition, it has been well studied that nutritional status in early life can determine the development of cognitive ability and lifetime health. As shown in Fig. 1.5, nutritional status in Korea had been rapidly improved after the Korean War. Average daily caloric intake per capita is estimated 2141 kcal in 1961, but it reached 3,000 kcal fifteen years later. Its direct impact can be found from the trend of average height among adolescences. For example, average height at age 17 among males who were born in 1960 was 167 cm in 1960. In next 20 years, the average height at age 17 increased to 172 cm by 5 cm among those born in the 1980s.[4]

Two aspects account for the increase of average caloric intake in the 1960–1970s. The government legislated for forbidding the production and distribution of distribute adulterated food in 1962. Based on the legislation, called as Food Sanitation Act, the government also introduced a license system for nutritionist and appointed universities for raising nutritionists. This provided an important basis for nourishment improvement. In addition, the government enacted 'National Health and Nutrition Improvement Order', and conducted national project for nourishment improvement in 1969. National nutrition survey was implemented in 1969. Its result was used to figure out the national status of nutrition and related problems. Campaigns for consuming mixed grains and changing diet pattern were launched in the 1970s. In particularly, because nutritional status was lower in rural areas than in urban areas, the national project for nourishment improvement was focused on rural areas by appointing demonstration villages, and education for nutrition and cooking.

National vaccination project was initiated by enacting 'Infectious Disease Prevention Act' in 1954. The act first selected smallpox, diphtheria, typhoid, typhus, paratyphoid, and tuberculosis as the diseases for regular schedule vaccination. The types of infectious diseases for vaccination were expanded throughout the 1960–1970s: BCD for children in 1963, measles in 1965, Japanese encephalitis in 1971, cholera 1976, and rubella in 1980. This vaccination project significantly contributed to improving infant and child mortality rate by preventing various infectious diseases.

---

[3]From the 1980s, Korea had more concerns on public-health issues of chronic diseases rather than infectious diseases. Korea government responded the change by introducing new policies and programs such as national insurance program, campaign for improving life style like quitting smoking.

[4]As discussed in the previous section, an individual's height depends on net nutritional status, i.e., nutritional supply minus demand. Thus, the increase of average height at age 17 would imply not only the increase of caloric intake, but also the reduced risk of exposure to infectious diseases.

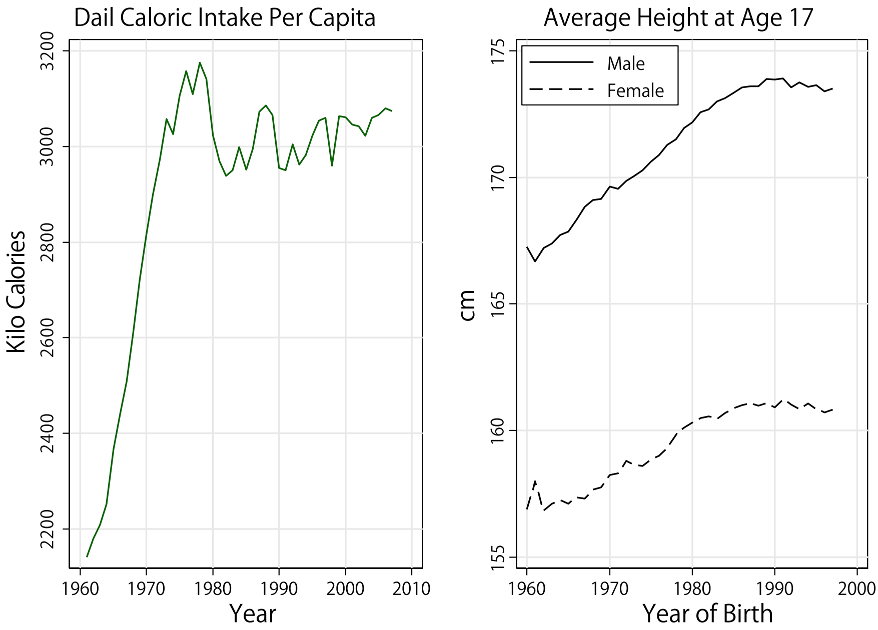


Fig. 1.5 Trends of average caloric intake and average height at Age 17. *Source* KOSIS

Intestinal parasite eradication project was also significant in improving nutritional status among children and adolescence. Although intestinal parasites are less fatal than other infectious diseases, they are very prevalent in the areas with poor sanitation and environment. It has been well studied that infection of intestinal parasite can lower cognitive ability and such impacts can last for a long period. For example, the United States succeeded in eradicating hookworm in the 1910s, and it has been analyzed that the cohorts born after the eradication had much benefit in accumulating better human capital (Bleakley 2007).

Korea began stool examination and parasite treatment programs from 1969 for all students in elementary, middle and high schools. The result of stool examination conducted in early years showed that about 77% of examined students were infected with parasites. This suggests that sanitary status was very poor in the 1960–1970s. The national project continued until the mid-1990s. As its result, the infection rate declined to 0.2% in 1995, when the project was discontinued (Kim et al. 2014).

Other national projects conducted in the 1960–1970s include tuberculosis eradication project, maternal and child health care project and so on.[5] Those projects

[5]In the 1960s, the national tuberculosis eradication project established regional TB examination offices, provided facilities for decoding X-ray results, and supported national survey of tuberculosis and BCD production. The project conducted in the 1970s includes establishing Korean Institute of Tuberculosis, improving X-ray test equipment, building up the international-level training program for tuberculosis management (Korean National Tuberculosis Association). On the other hand, the maternal and child health care project was conducted as a family-planning project at the begin-

are similar in that they were nation-widely conducted and were continued until clear performance was obtained. For example, since Korean National Tuberculosis Association was established in 1953, national tuberculosis emaciation project has persist until today. Owing to the project, the incidence rate of tuberculosis declined from 531 per 100,000 in 1970 to 136 in 1990.

Although the public health projects above were generally aimed to prevent and eradicate infectious and contagious diseases in a short period, it should be emphasized that their impact could last for a longer period. In particular, existing studies suggest that the control of diseases among infants, children and adolescence can enhance cognitive ability, increase the level of human capital, and improve labor productivity and health in adult.[6]

Third, the development of public health infrastructure was as significant as campaign or national projects in the process of disease control. After the Korean War, Korea did not have nation-wide organizations or institutes to deal with severe public health issues. To overcome the problem, the government first established 15 health care centers at the major cities and 471 community health care offices across the nation in 1953. Most of all, The Health Care Center Act was enacted in 1956. This provided a legislative foundation for establishing additional health care centers and expanding health-promoting activities in spite of lack of public health personnel and equipment.

The establishment of National Medical Center provided the momentum that led to more systematic action for population health status. The center was established by a presidential decree in 1958 with the support of health care personnel and equipment from three Scandinavian countries including Norway, Denmark and Sweden. Its management authority was transferred to Korean government in 1968, and the government enacted the special account act for the persistent budget of the center. After then, the center has played the core role for public health care service and policy development.

In addition to the establishment of organization and legislative foundation, the development of effective infrastructure for public health was begun from the 1960s. One representative project was to build up water supply system. Because the project needed much budget, the initial investment for building up waterworks much depended on financial aids from foreign countries. The development of waterworks was focused on major cities in the 1960s and was rapidly expanded to small cities and towns. Its outcome was remarkable as shown in Fig. 1.6. The rate of water supply—the ratio of populations to whom water pipes were connected—was below 20% in 1960, but it rapidly increased to 33% in 1970 and 55% in 1980.

The supply of sewer was a bit delayed than that of waterworks. Throughout the 1970s, industrialization was intensified and became one of the main factors that caused serious water pollution. Reflecting increasing concern about water pollution,

ning years. It was the mid-1980s that the project began to provide health-promoting programs for expecting mothers and infants (National Archives of Korea Online DB).

[6] Almond and Currie (2011) well summarize existing studies that have examined the significance of early-life conditions and the long-term impact of investment in early life on later outcomes.

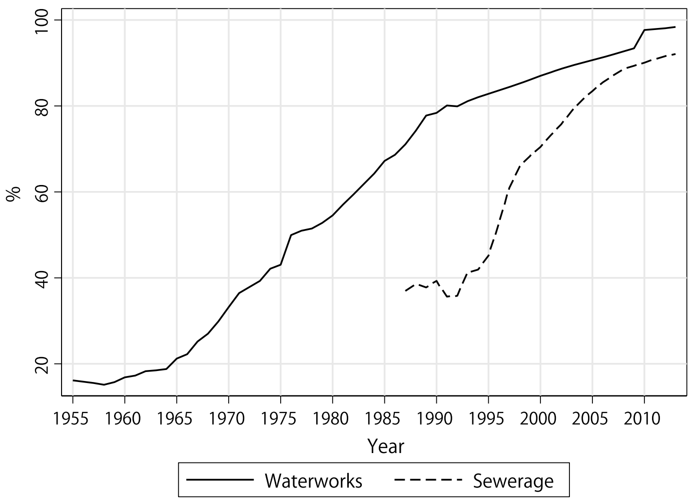


Fig. 1.6 Trends of waterworks and sewerage supply rate. *Source* Statistics of Waterworks, Statistics of Sewerage. *Note* The supply rate is calculated as the ratio of populations who were eligible for the benefit of waterworks or sewerage

the government established the first sewage disposal treatment plant for the Han River in 1976. Although the statistics prior to the mid-1980s are not available, the efforts of developing sewer made a significant achievement. In the 1990s, the rate of sewer supply was around 40%, which is only the half of the rate of water supply in the same period. According to existing studies that examined the impact of water and sewer supplies established in the early 20th century, the most benefit is found in that the supply of water and sewer significantly reduced the incidence rate from infectious diseases, particularly water-borne diseases like typhoid (Ferrie and Werner 2008). It is thought that the development of water and sewer system in Korea has had the similar impact on the control of disease.

So far we have discussed how Korea could succeed in eradicating infectious diseases and enhance the status of population health throughout the 1960–1970s. As the main policy background of the improvement, we reviewed the significance of campaign for personal hygiene and sanitation, active national projects of public health, and development of public health infrastructure. The interesting point is that the policies and programs conducted in Korea were those successfully done in developed countries such as Britain and the United States in the early 20th century. The difference between Korea and other developed countries is that Korea achieved what developed countries did for about half century only within about two decades. Two things should be emphasized as the background of the difference.

First, Korea well utilized the advantage of last mover. Western developed countries developed effective ways to prevent and eradicate infectious diseases undergoing lots of trial and error over many years. But Korea could rapidly adopt those policies and programs developed by developed countries. This saved the cost of time that would be spent for trial and error, and also increased the effectiveness of policies and programs. Second, the government's driving force and nation-wide participation were important as well. In particular, the government provided legislative foundation by enacting various public-health-related acts. In addition, it was the government that educated and continuously advertised the significance and ways of preventing and eradicating infectious diseases. Such activities were important under the situation that information and public finance was inadequate. Moreover, populations actively participated the policies and programs proposed by the government. This maximized the effectiveness of the policies, and so would be the most important background of the success.

After 1980, Korea fulfilled the control of disease. Infant mortality rate and incidence rate of infectious diseases declined to the level of developed countries in the mid-20th century. As seen in Fig. 1.2 above, life expectancy at birth continued to increase and reached 70 in 1985. This is the same with that of the United States in the 1950s. Since the 1980s, Korea has made great investment in not only the control of infectious diseases, but also the development of medical technologies, the supply of national insurance program, the project for improving health behavior and life style. This strategy in Korea was very similar with what the United States took after the 1950s. As the result of the strategy, Korean life expectancy exceeded that of the United States in the 2000, and Korea became one of the countries with highest life expectancy only 50 years after the Korea War.

1.4 Educational Investment and Expansion of Human Capital

Another key measure of living standards widely used is the level of educational attainment and the expansion of opportunity for education. A higher level of education and schooling can enhance labor productivity and human capital that are essential for growth and development. In addition, through education, people obtain various information for prevention of disease, better health and so on. Moreover, education can induce more people to participate in voting and so can promote the advance of politics. Therefore, education is significant for improving the quality of life in various aspects. Furthermore, the expansion of educational opportunity plays a role of accelerating social mobility and so reducing social inequalities. This benefit of education well accounts for why industrialized countries in the past had made efforts to improve educational environment.

Because there was few proper measure of educational attainment in the past, illiteracy rate was frequently used for the measure. In particular, the rate could represent the level of education for the period when schools or educational institutes were inadequate. It is estimated that Korean illiteracy rate was very high during the Japanese colonial era. As Fig. 1.7 shows, the rate was about 80% in 1945. Although the oppor-

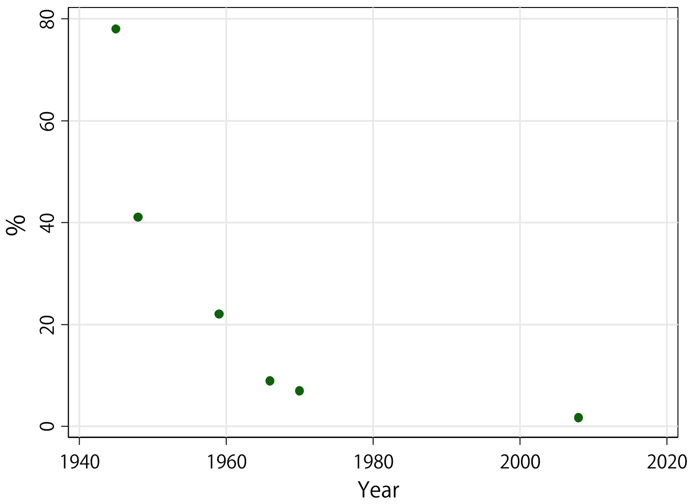


Fig. 1.7 Trend of illiteracy rate. *Source* Kim (2009)

tunity of schooling was increasing during the colonial era, the use of Korean was much discouraged compared with the use of Japanese or Chinese characters.

In the survey conducted right after the independence, illiteracy rate was reported to substantially decline to 41%. The rate surveyed in 1959 after the Korean War continued to decline to 22%. The main reason for rapid decline in illiteracy rate was nation-wide campaign for eradicating illiteracy. The campaign was continuously conducted throughout the 1950s. According to Korean National Archives records, the Korean ministry of education proposed 'Complete Eradication of Illiteracy Plan' in 1953. A slogan posted during the campaign says 'As illiterate person decreases more, Korea will be more democratized', which is a symbolic message that shows the goal of the campaign. From 1954, the government initiated the 5-year project for eradicating illiteracy. In particular, because illiteracy was highly prevalent in rural areas, the project was intensively conducted for 70–90 days in agricultural off-season to educate illiterate persons in rural areas. Such campaigns continued over the 1960s, and so the illiteracy rate declined to below 10% by 1970. On the other hand, as illiteracy rate decreased, the government changed the method of voting from symbol-based voting to character-based voting. The rate of voting participation also substantially increased. Therefore, the campaign for eradicating illiteracy is thought to play a great role in providing a foundation for promoting democracy in Korea (Korean National Archives Online DB).

In addition to the campaign for literacy, the successful settlement of public schooling was crucial as the basis of education in Korea. The start of public schooling was the legislation of 'Education Act' in 1949. The act specified that all the populations have the right to take 6-year elementary education, which has been the legislative foundation for compulsory elementary education in Korea. Although the compulsory

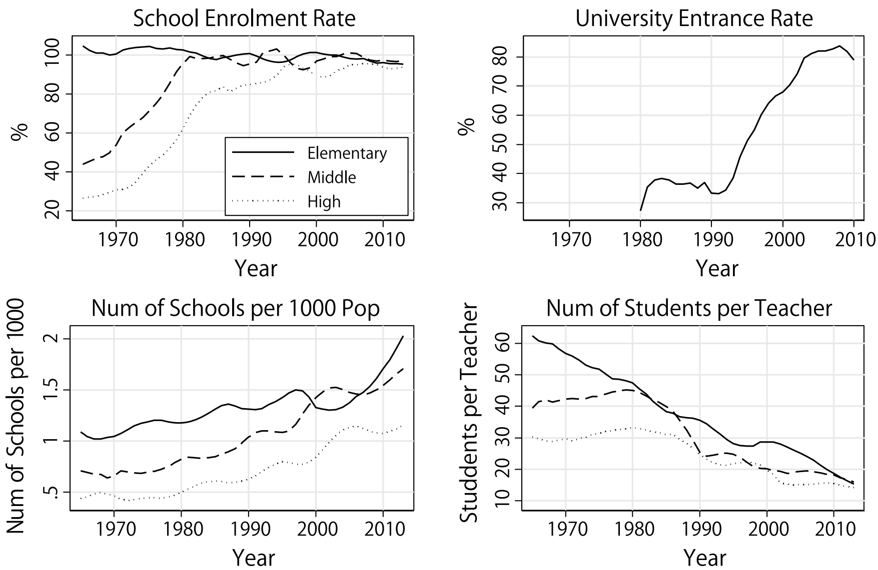


Fig. 1.8 Trends of educational indicators. *Source* Number of schools, students and teachers—Korean Educational Statistics Service; Population by Age—KOSIS. *Note* The enrollment rates of elementary, middle and high schools and the university entrance rate are calculated as the number of students who enrolled or entered each rank of schools divided by the number of populations at corresponding ages. We define school ages as follow: 7–12 for elementary school, 12–15 for middle school, 16–18 for high school, and 19–24 for university

elementary education was initiated in 1950, it failed due to the Korea War. It was 1952 when the compulsory education was re-initiated. In particular, the government proposed the '6-Year Plan for Completing the Compulsory Education' in 1954. This was effective for building out new schools and educational equipment and procuring necessary budget, increasing the elementary school enrollment rate up to 96%.

Figure 1.8 shows the trend of school enrollment rate for three ranks of school (elementary, middle, and high school) since 1965. The rate was estimated as the ratio of students enrolled out of total populations at the rank-specific school ages. The figure shows that the school enrollment rate in the elementary school already reached 100% in 1965. This is the result of persistent efforts for compulsory schooling since the 1950s.

The policy for public schooling was not focused only on increasing the number of students. Much effort was made to expand inadequate teachers and facilities in the 1960s. The policies such as 'the First Plan of Expanding Facilities for Compulsory Schooling' proposed in 1962 and its second plan initiated in 1967 were effective. Eventually, the compulsory elementary schooling was completed in the 1970s.

On the other hand, Fig. 1.8 shows that the middle-school enrollment rate reached 45% in 1970 and approached to almost 100% in 1980. The compulsory schooling for middle school was already achieved before it was enacted in 1984. The high-school enrollment rate was about 30% in 1970, rapidly increased to 60% in 1980 and 80% in 1990, and exceeded 90% after 2000. The figure also presents that the number of schools per population at the school ages continued to increase since the 1960s, and the number of students per teacher declined. This implied that there had been a substantial improvement not only in the quantity of schooling, but also its quality.

One of key features that characterize the development of education in Korea is that education has been utilized as a way of economic growth. Skill education, particularly in manufacturing industry, agriculture, fisheries, and so on, was necessary to accelerate industrialization and to overcome poor educational environments persisted after the Korean War. Although elementary education was focused on basis education, secondary and tertiary education attached importance to skill education. This provided a large number of skilled workers who were essential for industrialization.

The government supported skill education in various ways. It first set up a policy foundation by initiating 'the 5-Year Plan for Skill Education' in 1957. The 'Industrial Education Promotion Act' was enacted in 1963, and it provided a legislative foundation that connected between skill education and industrialization. In addition, a series of policies that link between sciences and skill education were initiated after the mid-1960s. They emphasized skill training and practice in skill education, and intensified the industrial cooperation.

Those policies after the 'Industrial Education Promotion Act' were also useful for improving the quality of skill education. The ratio of students enrolled in vocational high schools (i.e., schools for skill education) was 41.3% in 1963, and it increased to 52.5% in 1969 exceeding the ratio of students in academic high schools. In addition, the government's financial support also increased much so that the proportion of skill-education budgets out of the ministry of education's total budget increased from 0.8% in 1965 to 3% in 1969. Moreover, the government enacted the 'National Technique Qualification Act' in 1974 to enhance the social status of skilled, scientific and engineering personnel. This promoted spirit among students in vocational high schools, publicized the significance of skill education and contributed to the advance of skill and technology in Korea.

On the other hand, the expansion of educational opportunity in the 1960–1970s played a role in reducing the gender gap of educational attainment. According to Fig. 1.9, the average years of schooling among females aged 25–34 was 7 years, which was only about 74% of the average years (9.4 years) among males at the same sages. But the gap had been substantially narrowed to 84% in 1980 and 92% in 1990.

In addition, the expansion of female education triggered the increase of female labor force participation. Figure 1.9 shows that female labor force participation rate gradually increased from 37.2% in 1965 to about 50% in the mid-1990s. But the rate of male slightly declined from late 70% in 1965 to early 70% in recent years. As a result, the gender gap of labor force participation rate decreased from 42% points in 1965 to 30% points in the mid-1990s.

The increase of female labor force participation suggests that female labor force moved from household work to industrial sectors, supporting economic growth. The expansion of educational opportunity among females was discussed above as its key background, but the decline of fertility rate played another key role in increasing

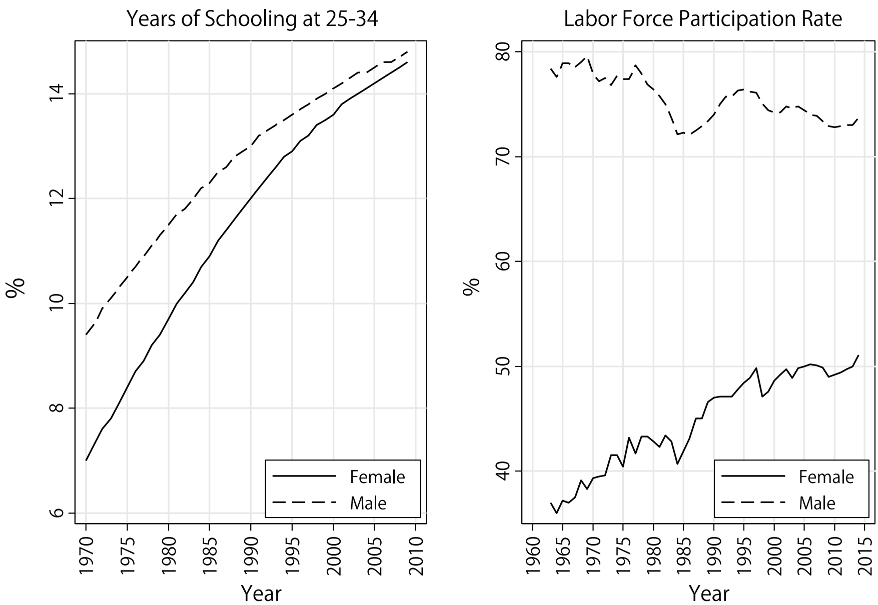


Fig. 1.9 Trends of crude birth rate. *Source* KOSIS. *Note* Labor force participation rate was estimated for populations aged at 15 or more

labor force participation. Figure 1.10 presents the trend of crude birth rate, which is defined as the number of newborn babies per 1,000 female populations. The rate was as high as 6 in the 1960s, but it substantially declined to the level below 2 by 1990. It is thought that the trend reduced the burdens of household work and infant care, and encouraged females to more actively participate in labor force.[7]

The policy of birth control in the 1960–1970s was another important policy that helped economic growth. Korea experienced a baby boom with high fertility rates after the Korean War, but mortality rate rapidly declined much during the same period as the control of disease became successful. This caused a rapid population increase. This made a burden for the economy, and the concern about birth control was raised. The government began to actively run the birth control project from 1963, developing the policies, organizations and personnel for the project. One of the key policy was the '10-Year Plan for Birth Control' initiated by the ministry of health and social affairs and the economic planning board in 1963.[8] Its main program was to encourage contraception for females at ages 20–44. Various policies for birth

[7]On the contrary, the decline of fertility rate can be caused by increased labor force participation among females.

[8]The birth-control project in the 1960–1970s has been considered as one of the most successful national projects. Most of all, the project has been thought to induce nation-wide voluntary participations. Various organizations such as community health-care centers, Korean Association of Family Plan, women's societies of the new village movement play leading roles to induce national participation.

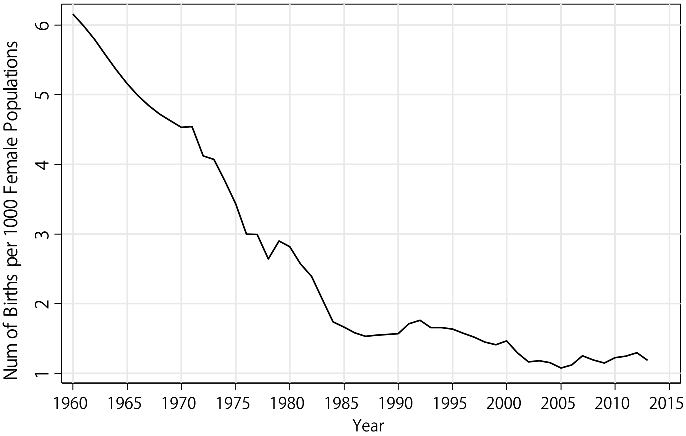


Fig. 1.10 Trends of crude birth rate. *Source* KOSIS. *Note* Crude birth rate is calculated as the number of newborn babies divided by female populations in 1000

control were introduced in the 1970s as well. For example, the government provided a tax benefit for families with 1–3 children in the mid-1970s. In addition, to weaken traditional son preference, the government amended the Family Act by allowing the right of inheritance to females in 1977.

In summary, the great investment in education and human capital in the 1960–1970s became the foundation for economic growth in Korea. This successfully completed the compulsory elementary schooling, raised lots of skilled works through skill education, and promoted rapid industrialization in the 1960–1970s. The expansion of educational opportunities is found not only in the quantity of education but also in its quality such as improvement in school facilities and teachers. Most of all, the reduction of the gender gap in educational opportunities was meaningful. This encouraged females to more participate in labor force. The project of birth control successfully lowered fertility rate, and also affected female labor force participation.

1.5 Increase of Consumption and Improvement in Inequalities

The rapid growth in the 1960–1970s brought about the increase consumption as much as the increase of income. Per-capita consumption expenditure is estimated as KRW800,000 in terms of 2010 constant value in 1965, and then it rapidly increased to KRW 1,190,000 in 1970 and to KRW 2,660,000 in 1980. It is thought that the increase of consumption clearly enhanced the quality of life. But it would be more important to know that there was a considerable change in consumption pattern.

Table 1.1 Ratio of household consumption expenditure by purposes (%) in 1970–2010

Consumption category	1970	1980	1990	2000	2010
Food, non-alcoholic beverages	39.7	34.8	22.7	13.5	12.8
Alcoholic beverages, tobacco	6.5	6.5	3.5	2.7	2.2
Clothing and shoes	10.3	9.1	7.2	6.4	6.2
Rent and utilities	9.2	9.6	13	20.2	18
Household facilities and management	3	3.2	5.4	4	3
Health care	2.6	2.7	3.7	2.9	4.3
Transportation	7.3	8.5	11.5	12	12.1
Communication	0.3	0.6	1.5	5.3	4.1
Leisure, entertainment	3.1	3.6	7.3	7.6	8.3
Education	2.7	4.3	5.5	5.4	6.9
Eat out, lodge	9.9	9.7	7.8	7.8	8
Others	5.4	7.2	10.9	12.3	14.1

Source KOSIS

Table 1.1 presents the trend of private consumption by its purpose after 1970. The most notable feature in the trend is the rapid decline of expenditure for clothing and food. It is observed that the proportion of consumption of food, clothing and shelter declined from 65.7% in 1970 to 39.2% in 2010.[9]

While the private expenditure for clothing, food and shelter declined, the proportion of expenditure for health care, transportation, communication, leisure and education greatly increased twice in 1970–2010. This suggests that the household demand for non-necessary goods rose much. In other words, the consumption pattern has changed toward improving the quality life as well as increasing the scale of consumption.

It is worthy to note that the expenditure for education was the category that had the most increase between 1970 and 1980. This implies that private demand for education and educational investment was very high even though the public expenditure for education was at a high level during the period.

On the other hand, the amount of public finance rapidly increased in the 1970s as much as did that of private consumption. According to Table 1.2, the ratio of government budge relative to GDP was about 18.8% in 1970 and increased to 22.4% by 1980. Because the rapid economic growth in the 1970s increased real national income by 2.2 times, it is suggested that the increase of public finance in terms of absolute value was substantial. The ratio did not change much until the 2010s.

It should be also paid attention that there has been a considerable change in public-expenditure pattern since 1970. The government budget was most expended for national defense mainly because of the confrontation between South and North Korea since the Korean War. The ratio of public expenditure for national defense

[9]The category of food, clothing and shelter in Table 1.1 contains the expenditure for food, non-food beverages, alcoholic beverages, tobacco, clothing, shoes, rent, and various utilities.

Table 1.2 Ratio of public expenditure by purposes (%) in 1970–2010

Expenditure category	1970	1980	1990	2000	2010
Proportion of public finance out of GDP	18.8	22.4	19.0	22.3	23.4
Distribution of public expenditure					
General public admiration	19.9	17.6	19.1	17.5	15.7
National defense	38.4	35.6	20.0	17.6	12.6
Public order and safety	8.3	7.6	8.5	9.7	7.2
Economic affairs	5.3	6.8	8.7	8.2	12.4
Environmental protection	0.8	0.7	1.9	1.9	2.4
Housing and regional development	1.2	2.2	0.6	1.0	0.7
Public health	1.3	5.0	10.7	16.7	23.5
Entertainment, culture and religion	0.7	0.8	1.1	1.2	1.3
Education	22.3	21.8	25.3	22.5	20.9
Social protection	1.7	1.8	4.1	3.8	3.2

Source KOSIS

reached 38.4% in 1970, and the amount of its budget continued to increase since then. However, the ratio of expenditure for national defense has declined over decades. One the other hand, the public expenditure for public health and health care considerable increased in 1970–1980. Its ratio out of total budget was only 1.3% in 1970, but increased to 5% by 1980. In fact, it continued to increase and so reached about 23.5% in 2010. It is found from Table 1.1 that the ratio of health care expenditure out of private consumption also increased from 2.6% in 1970 to 4.3% in 2010. This strongly suggests that the role of public section was very significant in enhancing the status of population health. As we discussed in Sect. 1.3, the statistics above well support that the government and communities played key roles in the control of disease in the 1960–1970s.

Another feature found in the trend of public finance pattern is that public expenditure for education was substantial from the beginning period of industrialization. The ratio of educational budget was 22.3% in 1970, which is the second greatest next to that for national defense. Even in 2010, the ratio of educational budget was 20.9%, which is the second largest next to that of health care. This suggests that Korea had a large demand of educational investment from the early stage of industrialization.

The last category of public expenditure in Table 1.2 shows the ratio of budget expended for social projection and its change over decades. The category of social protection contains the expenditure for social insurance, public aids, social welfare services, and so on. Although its proration out of national budget was small, it gradually increased by 1990. The purpose of social-protection expenditure is to support poor, disabled, and disadvantaged groups. Thus, it is suggested that, as the budget for social projection increased, the disparities and inequalities across groups declined.

How much the inequalities across groups had been eliminated during the 1960–1970s when Korea experienced the rapid economic growth. One of the most popular

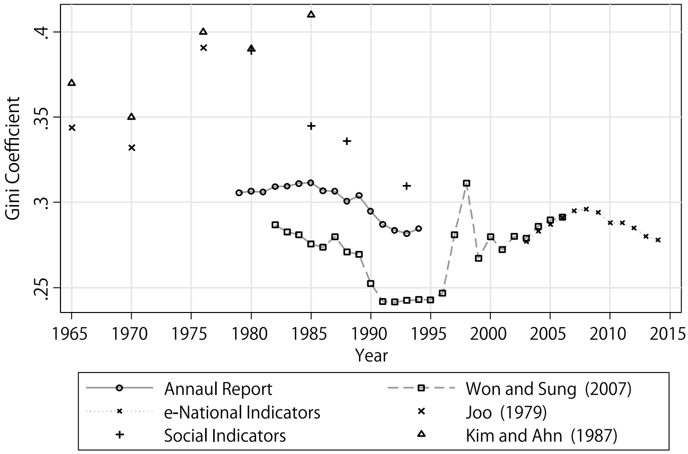


Fig. 1.11 Trend of income inequality measured by Gini coefficient. *Note* The Gini coefficient of Annual Report (Annual Report on the Family Income and Expenditure Survey was estimated using non-farm households including single-person households. The coefficient for Won and Sung (2007) and e-National Indicator uses non-farm household data excluding single-person households

way to see the change is to examine the trend of income inequality on the basis of Gini coefficient. Household survey, which allows a systematic analysis of inequality of household income, was not conducted until the late 1970s. In the earlier years, Gini coefficients were sporadically estimated and reported by Department of Statistics and researchers. Figure 1.11 shows the trend of Gini coefficient on the basis of various available estimates. The coefficient was between 0.35 and 0.4 until the late 1970s on average, but it is estimated that the inequality declined from the 1970s to the year just before the Asian financial crisis in 1997. The Gini coefficient soared due to the crisis in 1997, and has stayed around between 0.25 and 0.3. In conclusion, the income inequality in Korea measured by Gini coefficient had improved throughout the 1970s.

Actually, the level of Gini coefficient in the 1970s between 0.35 and 0.4 is frequently found for even developed countries today. This implies that, although average income was low in Korea in the 1960–1970s, the level of inequality was not that severe. It is generally observed that income inequality became worse when industrialization deepened, as the United States experienced in the 19th century. But income inequality had been lowered and stable during the period of rapid growth. This suggests that the policies for social welfare would be expanded.

The base for social welfare policies was established by the legislation of the 'Social Security Act' in 1963. The act clearly stated the basic principle for introducing the social security system and its significance. After the act was legislated, the social security system was gradationally expanded by providing various programs for socially disadvantaged classes throughout the 1960–1970s. The system had been completed by a series of legislation including the Livelihood Protection Act in 1963, the Health Insurance Act in 1963, the Industrial Accident Insurance Act in 1964, the National Pensions Act in 1973, the Medicaid Act in 1978, and so on. Of course, the coverage and target population of social security was small when it was first introduced. But throughout the 1980–1990s, each act was amended to increase its coverage and target population. We can say that such efforts for improving social welfare in the 1960–1970s laid the foundation stone that led Korea to a welfare state today.

The performance of lowered inequality in the 1960–1970s is also found from lowered gap between urban and rural areas. When Korea was at the early stage of industrialization in the mid-1960s, average income of rural workers was about 60% of the average of workers in cities. In addition, sanitary status, nutritional status, and living conditions in rural areas were much poor compared with those of urban areas. Migration out of rural areas was accelerated. The concern was increasing that the urban-rural gap would deepen as the nation became more industrialized.

However, as seen in Fig. 1.12, the income gap between urban and rural areas rapidly declined until the mid-1980s. This had been most influenced by the Saemaeul Movement (so-called the new village movement), which was a series of national projects conducted in the 1970s to modernize rural conditions.

The Saemaeul Movement is known to have originated from a project of 'Village Condition Improvement', which was first mentioned at a local governors meeting by President Park in April, 1970 (Korean National Archives Online DB). The Village Condition Improvement project conducted in 1970–1971 includes village common works such as main road expansion, establishment of common wells and common wash places. The government supported participating villages by providing 336 sacks of cement to each of 33,267 village across the nation. In addition, the government evaluated each village's performance and selected 16,600 villages for extra supports

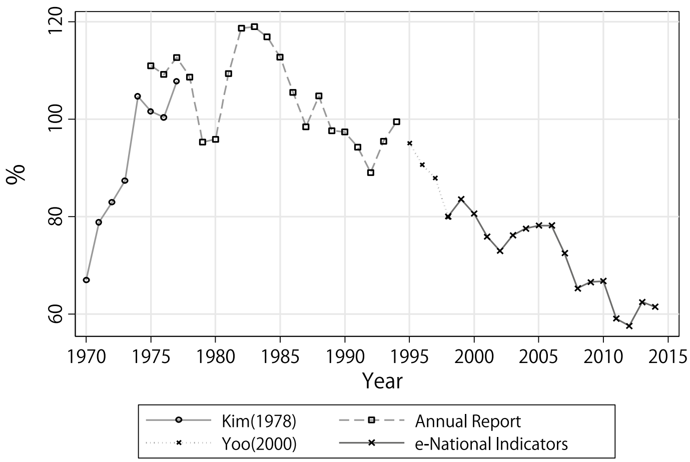


Fig. 1.12 Average income of rural workers relative to that of urban workers

of cement and iron bars. This means that the government had a strategy of incensive system by giving differential supports considering villages' performance. The incentive system significantly induced rural villages to actively participate in the Saemaeul Movement. Under the incentive system, the government selected villages that achieved good performance, had leaders well trained, and had adequate own budget for projects. Thus, the incentive system also promoted the spirit of cooperation and self-help among residents (Kim and Kim 2012).

The process of development of the Saemaeul Movement is classified into several stages, as presented in Table 1.3. First of all, the stage of foundation construction was conducted in 1970–1973. In addition to main road construction explained above, the movement in this stage includes the projects for improving rural environment such as modernizing roof and wall, those for building up rural infrastructure such as farm land arrangement and seed improvement, and those for increasing average income. Moreover, the reform of consciousness among rural residents was also contained as major projects, encouraging cooperation, thrift and saving.

The stage of project diffusion was taken in 1974–1976. As done in the previous stage, various projects for improving rural living conditions, income and spirit were conducted during this stage. House improvement and waterworks installation upgraded rural living conditions. Small stream maintenance, encouragement of complex farming, non-farm activities significantly increased average income among rural households. In addition, the programs for enhancing the spirit of the Saemaeul Movement were expanded as a nation-wide campaign. During the period, three key spirits of the Saemaeul Movement was created and diffused: diligence, self-help and cooperation.

Finally, the years of 1977–1979 are classified as the deepening phase of the Saemaeul Movement. The projects conducted in this period were larger scaled than those in previous stages, and they were diffused from rural areas and rural populations to urban areas and their residences. For example, the construction of bridges was implemented as a large-scale project for improving rural living conditions. The industrial complex was built up in rural areas to boost regional and rural economy. The three key spirits of the Saemaeul Movement were spread to urban areas by encouraging voluntary service activities for improving rural conditions and sisterhood among rural and urban villages.

In summary, the Saemaeul Movement was led by the government throughout the 1970s by conducting various projects for improving rural living conditions, income and spirit. It is very clear that the most representative result was the reduced gap between rural and urban areas, as shown in Fig. 1.12. It seems to be also obvious that the Movement also contributed to the control of disease though improving rural sanitary conditions and the expansion of educational opportunities and household income.

Table 1.3 The project of the Saemaeul Movement by three developmental stages

Stage	Period	Improvement purpose	Project
Foundation construction	1970–73	Rural living conditions	Expansion of village main road, improvement of roof and wall, construction of village common facilities
		Income increase	Establishment of agricultural roads, farm land arrangement, seed improvement, encouragement of Gye (traditional private fund) and exchange work
		Reform of spirit	Eradication of decadent trend, creation of cooperative environments, encouragement of thrift and saving
Project diffusion	1974–76	Rural living conditions	Improvement of houses, installation of waterworks, construction of community hall and common village facilities
		Income increase	Maintenance of footpath of rice field and small streams, encouragement of combined agriculture, operation of common workplace, discovery of non-farming income sources
		Reform of spirit	Diffusion of the spirit of the Saemaeul Movement (diligence, self-help and cooperation) from farmers to all the social classes including urban workers and students
Deepening phase	1977–79	Rural living conditions	Improvement of settlement structure, construction of large-scaled bridges and roads
		Income increase	Encouragement of stock raising and special crop cultivation, construction of Saemaeul plants, construction of industrial complex in rural areas
		Reform of spirit	Campaigns for promoting community spirit: sisterhood between rural and urban communities voluntary services in rural areas
			Campaigns for restoring cooperative community culture: help needy neighbors, improving alleys, keeping public order, conservation of nature
			Projects for promoting the stability of labor and management and productivity: establishment of industrial schools, expansion of welfare facilities in factories

Source Korean National Archives Online DB

1.6 The Impact of Improvement in Living Standards in the 1960–1970s

So far we have examined the background and development process of improvement in living standards. Most of all, the government showed a strong leadership by systematically introducing and implementing related policies and programs. This would be one of the key to the success, which we reviewed using various indexes of living standards. In this section, we compare the status of living standards across countries over decades, and analyze how the improvement in living standards in the 1960–1970s has affected outcomes today from a long-term perspective.

Comprehensive measures that can capture various aspects of living standards are required to compare living standards across countries. As discussed above, the Human Development Index (HDI) estimated by the United Nations is one of popular measures that satisfy such needs. The HDI utilizes three components that can affect the quality of life such as income, health and education. We also utilized those components to examine the improvement of living standards in Korea in the previous sections. The formula for the HDI is given below. The maximum of each component is the target value most desirable to human begins, and its minimum is assumed as the poorest level of living standards. The formula below produces an index that measures the level of each component, which ranges between zero and one. The HDI is estimated by weighted average of three indexes with equal weights.

$$\text{Component Index for the HDI} = \frac{x - \min(x)}{\max(x) - \min(x)},$$

x denotes the component of income, health or education.

The original HDI utilizes life expectancy as the variable of health, and school enrollment rate and literacy rate for the variable of education. However, those variables of health and education are inadequate for the 1960–1970s so that we may not be able to compare the HDI across the countries in the past. To overcome the problem, we estimated alternative development index by replacing the variables of health and education. We use the enrollment rate of middle school students as a measure of education, and infant mortality rate instead of life expectancy.[10] The enrollment rate of middle school students has a clear variation across countries, and is thought to be closely related with the level of human capital. Infant mortality rate is thought to more represent the level of disease control than does life expectancy.

The development index estimated by the method above is presented in Fig. 1.13. The left panel of the figure shows the scatter plot between 1971 and 2013 development indexes. The clear positive correlation implies that the countries with higher living standards in 1971 also enjoy higher living standards today. It is worthy to note that

[10]When we use infant mortality rate, we use ‘x – min(x)’ as the value of numerator in the index formula above rather than ‘max(x) – x’ because living standards increase as infant mortality rate decreases. On the other hand, we use the values of maximum and minimum within sample data, respectively, for max(x) and min(x).

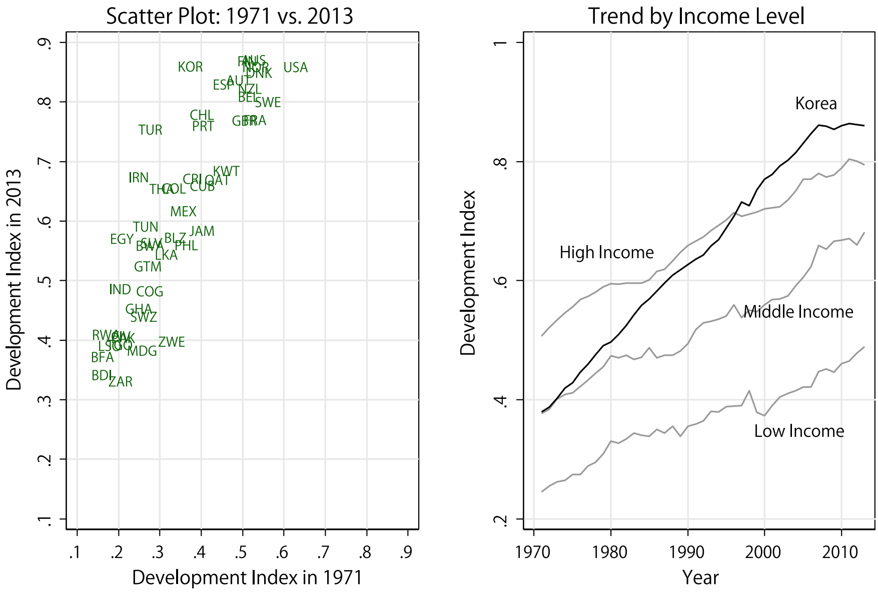


Fig. 1.13 Long-term trend of development index. *Note* The development index used in the figure was estimated over 1971–2013 by national income per capita as income component, infant mortality rate as health component, and enrollment rate of secondary school as education component. In the right panel above, we clustered countries into three groups in terms of per-capita income in 1970, and present the trend of average development index across the groups

the index for Korea is one of the top countries in 2013, but it was just at a middle rank 40 years ago among the countries analyzed.

The right panel in Fig. 1.13 shows the trend of average index over years by three income groups of countries in terms of per-capita income in 1970. It is notable that the cross-country gap observed in 1970 was not narrowed until today. Korea was exceptional in this general pattern. Korea was classified into low-income group in 1970, but its development index was high ranked among low-income countries. This suggests that its level of non-material living standards such as health and education was higher than that of material conditions in the 1970s. This is a result from investment in the control of disease and the expansion of educational opportunities. The more important fact found in the figure is that Korea has become one of the countries with best living standards or highest development index only 40 years after the investment was made in the 1960–1970s.

To study the long-term impact of improvement in living standards, we examine the association between living standards in the past and current income or living standards among previously low-income countries in the remaining of this section. From the result of this analysis, we will get evidence that shows that current high income and economic foundation was significantly influenced by the improvement in living standards in the 1960–1970.

Fig. 1.14 Correlations between income in 1970 and income and UN HDI in 2013 among low-income countries. *Sources* Per-capita national income for 1970 and 2013—Gapminder, UN HDI (Human Development Index) for 2013—UNDP *Note* The analysis above is based on 58 low-income countries whose per-capita income was low than USD 400 (in 2011 constant dollar) in 1970. All the income variables are measured in 2011 constant dollar

For the quantitative analysis, we selected about 50 countries, including Korea, which are classified as low-income countries in terms of per-capita income in 1970. Figure 1.14 shows the cross-country correlation between 1970 and 2013 per-capita income, and between 1970 income and 2013 development index as a measure of living standards. Two figures show that current income and living standards are little correlated with income 40 years ago. The initial endowment has little influence on current outcomes because we compare them only among low-income countries in 1970.

Figure 1.15 shows how income and UN HDI in 2013 are correlated with the percentage change in the 1970s of three components used for the development index, i.e., income, health and education. Although per-capita income in 1970 was little correlated with current outcomes, the change of income in the 1970s, i.e., economic growth rate, has positive correlations with per-capita income and HDI measured in 2013. In addition, as infant mortality rate more declined in the 1970s (i.e., as the control of disease was more successful), the country has higher income and HDI today. But we do not find a clear correlation between improvement in educational attainment in the 1970s and today's outcomes.

The scatter plots in Fig. 1.15 shows that today's income and living standards may be closely related with not only income growth in the 1970s, but also improvement in infant mortality rate as a result of disease control. Although the relation with

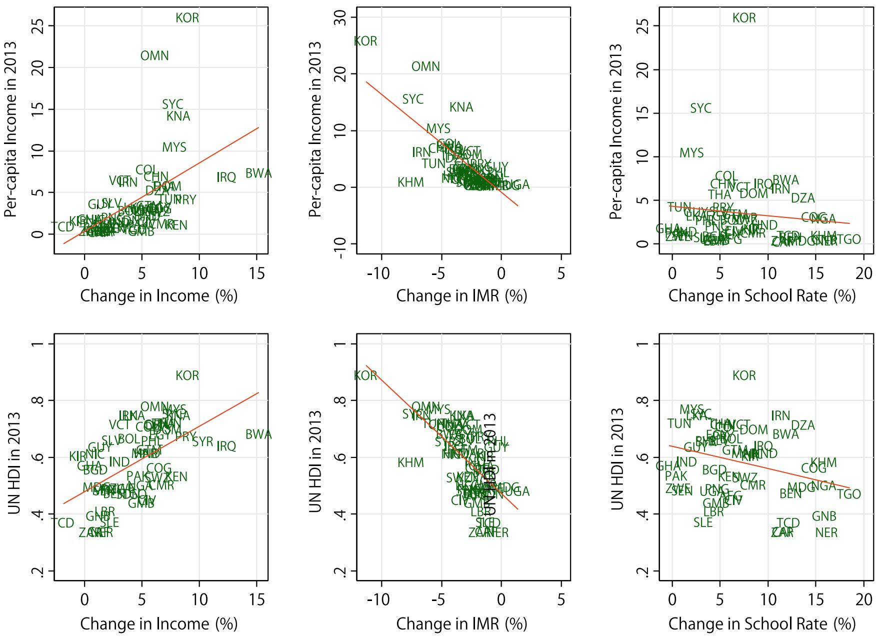


Fig. 1.15 Changes of income, health and education in the 1970s and correlations with income and UN HDI in 2013. *Sources* Per-capita income for 2013—Gapminder, UN HDI for 2013—UNDP, Per-capita income for the 1970s—Gapminder, Infant mortality rate and educational level for the 1970s—World Bank *Note* The analysis above is based on 58 low-income countries whose per-capita income was low than USD 400 (in 2011 constant dollar) in 1970. The rate of change of each component is estimated as annual percentage change over the 1970s. The solid lines on the graphs are liner trends lines

educational improvement is not clearly found from the figures, our discussion and evidence shown in previous sections suggest that educational improvement in the 1970s would have positive effects on income and living standards today.

Although the number of sample used here is small, we conducted some simple regressions to see whether the correlations observed in the scatter plots are statistically significant and which component has been more important. We use per-capita income in 2013 as the dependent variable in the regressions. Key control variables are the rate of change in income, infant mortality rate and middle-school enrolment rate in the 1970s. Additionally, we control for continent fixed effects to capture time-invariant characteristics specific to each continent, and the rate of change in the three components in the 1980s.

The regression results are presented in Table 1.4. The estimated coefficients in models (1)–(3) suggest that as the countries—low income countries in 1970—had more rapid income growth, more decline in infant mortality rate and higher educational attainment in the 1970s, they enjoyed higher income and living standards in 2013. These effects are estimated statistically significant. In model (1.4), we put

Table 1.4 Estimated effect of change in living standards in the 1970s on income and UN HDI in 2013

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	Per-capita income in 2013 (1,000 US$, PPP)						2013 UN HDI
Continent fixed effects	NO	NO	NO	NO	YES	YES	YES
Change rate in the 1980s	NO	NO	NO	NO	NO	YES	YES
Change rate in the 1970s (%, Annual)							
Income	0.8220***			0.1228	0.173	0.1032	0.0098**
	−0.1922			−0.1296	−0.1246	−0.1513	−0.0034
Infant mortality rate		−1.7247***		−1.8147***	−1.7102***	−1.5673***	−0.0261***
		−0.214		−0.204	−0.2094	−0.26	−0.0057
School enrollment rate			0.1895**	0.1303***	0.1649***	0.1354***	0.0004
			−0.061	−0.0338	−0.0346	−0.0368	−0.0009
Sample size	53	55	53	49	49	45	50
Adjusted R-squared	0.25	0.542	0.142	0.763	0.787	0.804	0.765

Note The analysis above is based on 58 low-income countries whose per-capita income was low than USD 400 (in 2011 constant dollar) in 1970. The data used in the regression analysis is the same with those used for Fig. 1.15. Models (1)–(6) use per-capita income in 2013 as dependent variable. The key controls in the regressions are the annual change rate of three components (income, health and education) in the 1970s. Each regression uses OLS estimations. Models (5) and (6) control for continent fixed effects to capture time-invariant continent-specific characteristics. Model (6) additionally controls for the change of three components in the 1980s. In model (7), we replace the dependent variable with the UN Human Development Index in 2013. Other specifications in model (7) are the same with those of model (6). The standard errors are reported in parentheses. *, **, *** denotes that the estimated coefficients are statistically significant at the 10%, 5%, and 1% level, respectively

the control variables of three components together. It is found that the coefficient of income growth is estimated insignificant. But the effects of improvement in infant mortality rate and education attainment are still estimated significant. The key result found in Table 1.4 is not that income growth in the 1970s was insignificant for today's income, but that the component of health and education was significant as much as income growth. In model (5), we add continent fixed effects to control for average characteristics of each continent. In model (6), we control for the rate of change in three components in the 1980s. But additional controls do not change the estimation result much. The results of models (5)–(6) are much similar with that of model (4), suggesting the significance of health and education. Finally, in model (7), we replace the dependent variable with the HDI in 2013 instead of per-capita income. It is estimated that as income grew more rapidly and infant mortality rate declined more in the 1970s, a higher level of living standards or HDI was achieved today. This also implies that the control of disease has been as important as income growth in achieving better living standards from a long-term perspective.

Rather than causality, the regressions above likely estimate the correlation between improvement in living standards in the 1970 and outcomes today on the basis of limited sample and data. Nevertheless, we can get a lesson from the analysis that the correlation is very significant and that the improvement in health and education is important as much as income growth for economic growth and further improvement of living standards, especially among under-developed countries. Various experiences of Korea that we discussed in previous sections would be excellent cases and evidence that support the implication found in this section.

1.7 Concluding Remarks

So far we have examined how Korea could improve the standards of living throughout the 1960–1970s in various aspects, including disease, educational opportunity, gender gap, female labor force participation, consumption and public expenditure, rural-urban disparity, and so on. It was effective that the government had adopted various policies and programs through which developed countries succeeded in eradicating diseases, expanding educational opportunities and industrialization in the early 20th century. Korea could achieve the remarkable development of living standards—so called the Miracle on the Han River—not by creating new strategies, but by successfully conducting adopted strategies within several decades.

Therefore, we need to find the fundamental background of the success in Korea not from what policies were implemented, but from how policies were implemented. First, the government's strong driving force should be emphasized. The driving force was very effective in constructing public infrastructure and social system which were lost by the Korean War. But the role of the government in the 1960–1970 was not always based on enforcement and intervention. The government provided legislative foundation and incentive system to induce more reasonable decision and behavior from populations. Second, social consensus and national participation was another key factor that led to the success. Although the government played a key role in initiating new policies and programs, they could be successful only when populations actively participate them. Finally, the government was successful in selecting effective policies from a long-term perspective and allocating limited budget effectively. In recent years, many researchers emphasize that the investment of enhancing health and education is significant for economic growth and very cost-effective particularly in the long-term perspective. Therefore, the increase of investment in health and education by private and public sectors in the 1960–1970s has been the most key base for growth and development today.

References

[In Korean]

Hong, Sok Chul. 2013a. Review of studies on living standards in the perspective of economic history. *Review of Economic History* 55: 137–173.
Joo, Hak-Joong. 1979. Estimate and determinant of income distribution by class. In *Income distribution in Korea and its determinants*, vol. 1. Korean Development Institute.
Kim, Dae-mo, and Kook-shin Ahn. 1987. *Status and determinants of income distribution in Korea, and population's mindset toward the issue of distribution*. Ministry of Education.
Kim, In-hwan. 1978. *Green revolution in Korea*. Rural Development Administration.
Kim, Joon-kyung, and Kwang-sung Kim. 2012. *2012 project modularizing the Korea experience of economic development: Economic and social change in Korea and the role of the Saemaeul Movement*. KDI School of Public Policy and Management.
Kim, Soon-im. 2009. Summary of national basic literacy in 2008. *New Korean Language Life* 19 (2): 17–32.
Kim, Tae-jong, Jong-il Chae, and Hyun-kab Chang. 2014. *2013 project modularizing the Korea experience of economic development: Intestinal parasite eradication project (1969–1995)*. KDI School of Public Policy and Management.
Won, Jong-hak, and Myung-jae Sung. 2007. *Reasons of increasing gap in income distribution and direction of policy response*. Korean Institute of Public Finance.
Yoo, Byung-seo. 2000. Policy issues for income parity between farm and nonfarm sectors. *Korea Journal of Agricultural Management and Policy* 27: 89–102.

[In English]

Almond, Douglas, and Janet Currie. 2011. Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives* 25 (3): 153–172.
Bleakley, Hoyt. 2007. Disease and development: Evidence from Hookworm eradication in the American South. *Quarterly Journal of Economics* 122 (1): 73–117.
Carter, Susan, Scott Gartner, Michael Haines, Alan Olmstead, Richard Sutch, and Gavin Wright (eds.). 2006. *Historical statistics of the United States: Earliest times to the present*. New York: Cambridge University Press.
Centers for Disease Control. 1999. CDC on infectious diseases in the United States, 1900–99. *Population and Development Review* 25: 635–640.
Cutler, David, and Grant Miller. 2005. The role of public health improvements in health advances: The twentieth-Century United States. *Demography* 42 (1): 1–22.
Engels, Friedrich. 1845. *The condition of the working-class in England*, ed. David McLellan (1969). London: Panther.
Feinstein, Charles. 1998. Pessimism perpetuated: Real wages and the standard of living in Britain during and after the industrial revolution. *Journal of Economic History* 58: 625–658.
Ferrie, Joseph, and Werner Troesken. 2008. Water and Chicago's mortality transition, 1850–1925. *Explorations in Economic History* 45: 1–16.
Floud, Roderick, Robert W. Fogel, Bernard Harris, and Sok Chul Hong. 2011. *The changing body: Health, nutrition, and human development in the Western World since 1700*. Cambridge University Press.
Fogel, Robert W. 1986. Nutrition and decline in mortality since 1700: Some preliminary findings. In *Long-term factors in American economic growth*, ed. Stanley Engerman, and Robert Gallman. Chicago: University of Chicago Press.

Harris, Bernard. 1998. Gender, height and mortality in nineteenth- and twentieth-century Britain: Some preliminary reflections. In *The biological standard of living in comparative perspective*, ed. John Komlos, and Joerg Baten, 413–448. Stuttgart: Franz Steiner Verlag.

Hong, Sok Chul. 2013b. Malaria: An early indicator of later disease and work level. *Journal of Health Economics* 32 (3): 612–632.

Kuznets, Simon, Lillian Epstein, and Elizabeth Jenks. 1941. *National income and its composition, 1919–1938*, vol. I. National Bureau of Economic Research, Inc.

Lindert, Peter. 2004. *Growing public: Social spending and economic growth since the eighteenth century*. Cambridge: Cambridge University Press.

Lindert, Peter, and Jeffrey Williamson. 1983. English workers living standard during the industrial revolution. *Economic History Review* 36 (1): 1–25.

Nordhaus, William, and James Tobin. 1972. Is growth obsolete? In *Economic growth: Fiftieth anniversary colloquium V*, 1–80. New York: National Bureau of Economic Research.

Schultz, T. Paul. 2005. Productive benefits of health: Evidence from low-income countries. In *Health and economic growth*, ed. G. López-Casasnovas, B. Rivera, and L. Currais, 257–286. Cambridge, MA: MIT Press.

Sen, Amartya. 1999. *Development as freedom*. Oxford University Press.

Steckel, Richard, and Roderick Floud. 1997. *Health and welfare during industrialization*. Chicago: The University of Chicago Press.

Usher, Dan. 1980. *The measurement of economic growth*. Oxford: Blackwell.

[Database]

Annual Report on the Family Income and Expenditure Survey, Statistics Korea.

Annual Surveillance Report on Infectious Diseases, Centers for Disease Control and Prevention.

Korean National Archives Online DB. http://www.archives.go.kr.

Korean Social Indicators, Statistics Korea.

Korean Statistical Information Service [KOSIS]. http://kosis.kr.

Statistics of Sewerage, Ministry of Environment.

Statistics of Waterworks, Ministry of Environment.

Chapter 2
Fertility Dynamics with Family Bargaining

Akira Yakita

2.1 Introduction

Statistical discrimination might lead to shorter and less frequent labor market participation of women than would otherwise be measured.[1] Such is apparently the case especially in Japan, as emphasized by Kawaguchi (2008) and Yamaguchi (2017). It is statistically plausible that women are likely to withdraw from the labor market when they marry or have children. Therefore, women tend to be given only light and easy work and thereby earn a lower wage income. When capital accumulation is low, physical labor is considerably important for goods production. Men have more physical strength than women. Therefore, the male wage rate is high. Although biological differences exist between women and men, it has been recognized that statistical gender discrimination might not be a necessary consequence of biological gender differences.

Discrimination is asserted only as a consequence of a self-fulfilling prophecy, as argued by Kawaguchi (2008) and Yamaguchi (2017). Consequently, unlike men, women must bear disproportionately heavy burdens of housework, child-bearing, and child-rearing at home even while working in the labor market. These circumstances motivate women and men to negotiate their family decisions such as family labor supply, childbearing, and child-rearing. This paper presents analyses of the effects of family bargaining on fertility dynamics along with economic development. It has been acknowledged that the persistently declining fertility rate has turned upward at

[1] Schwab (1986) defines statistical discrimination as an employer's action according to a stereotype such as women quit more frequently than men. Although early arguments presented in the literature (e.g., Arrow 1972) implicitly assume that statistical discrimination must be efficient in producing output, Schwab (1986) shows that they might be inefficient. This paper does not address issues of efficiency.

A. Yakita (✉)
Nanzan University, Nagoya, Japan
e-mail: yakita@nanzan-u.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_2

higher income levels in economically developed countries since around 2000 (Apps and Rees 2004; Day 2012).

In a pioneering paper, Galor and Weil (1996) report an analysis of fertility and childrearing portioning between genders in a neoclassical growth model, assuming biological gender differences. Day (2016) develops the model to present the possibility of a fertility rebound in a Galor and Weil (1996) type model by incorporating market childcare services.[2] In contrast to these works in unitary models, Kemnitz and Thum (2015) present an analysis of the effects of gender wage gaps on fertility in a collective model of family decisions. However, they do not consider the dynamics. Yakita (2018a) reports results of an analysis of fertility dynamics along with economic development in a unitary model without family bargaining, assuming the production of market childcare services following Day (2012). As described herein, the analyses of fertility dynamics described by Yakita (2018a) are extended with consideration of family bargaining in a collective model.

Setting up of the model follows the pattern of a collective model of Komura (2013), who assumes heterogeneous preferences for having children. The strong bargaining power of a spouse with a stronger preference for rearing children is expected to raise the fertility rate. If the bargaining power depends on the relative wage rate and if the wage rate depends on capital accumulation, then economic development powered by capital accumulation influences the bargaining power distribution, which consequently affects fertility dynamics.

The main results are the following. As the relative female wage rate rises along with economic development, the fertility rate rebounds in spite of the increased opportunity cost of child-rearing time at home if women have strong preferences for having children. Efficient provision of market childcare services reinforces the upward dynamics of fertility. It also helps alleviate gender discrimination.

The structure of the paper is the following. The next section sets up a growth model. Section 2.3 presents the model dynamics. The last section concludes the paper.

2.2 Model

This paper presents consideration of a collective model of family used by Komura (2013), who emphasizes gender-based difference in preferences for having children. Couples might choose between rearing their children at home and purchasing market childcare services. An underlying assumption for this paper is that, because of existing statistical discrimination, only women rear their children at home. Men work fulltime and inelastically in the labor market, although women choose to allocate their time between child rearing at home and market labor.

[2] Apps and Rees (2004) and Hirazawa and Yakita (2009) also assume the availability of childcare outside the home.

For analytical purposes, child rearing at home and market childcare services are assumed to be perfect substitutes. This assumption is useful to clarify the effects of market childcare availability on family decision making. Market childcare service production uses only female labor with material goods inputs, as described by Day (2016) and Yakita (2018a). Market childcare services are provided as club goods such as kindergartens or childcare centers, where a childcare worker can care for more than one child at a time.

Goods production technology can be represented as a production function with capital, mental labor, and physical labor with constant returns to scale, as described by Galor and Weil (1996).

2.2.1 Households

In this model[3], a family comprises a woman and a man, i.e., a couple. Individuals live for two periods: first a young working-period and then an old retired-period. The length of lifetime is certain. The length of each period is normalized to one. To avoid the matching issue in marriage, equal members of women and men are assumed. In addition, a couple is assumed to rear the same number of girls and boys. Each couple has a common utility over consumption and having children.

The utility function of a couple in period t is[4]

$$u_t = [\gamma^f \theta + \gamma^m (1 - \theta)] \ln n_t + \rho \ln c^2_{t+1}, \tag{2.1}$$

where c_{t+1} represents second-period family consumption and n_t denotes the number of girl-boy pairs.[5] $\rho \in (0, 1)$ stands for the time preference factor. $\gamma^i > 0$ is the utility weight of person i on having children. The relative magnitude of the utility weight of women might be greater than, equal to, or less than that of men.[6] Also, $\theta = \theta(w^f/w^m)$ denotes the bargaining power of a women in a family, where w^f_t and w^m_t respectively represent the wage rates of women and men.[7] An important assumption is that $\theta(0) = 0$, $\theta(1) = 1/2$ and $\theta'(w^f/w^m) > 0$ for $0 < w^f/w^m \leqslant 1$.[8]

[3]This section is attributable to Kemnitz and Thum (2015).

[4]Komura (2013) assumes a similar family utility function.

[5]For analytical simplicity, first-period consumption of the family is omitted here. The period of childhood is not addressed explicitly in this paper.

[6]Actually, the ideal numbers of children of women and men vary among countries. The OECD Family Database (http://www.oecd.org/social/family/database.htm. Accessed 12 January 2017) reported that the average personal ideal numbers of children for women and men 15 years old and older in the mid-2000s were, respectively, 2.58 and 2.50 in France, 2.48 and 2.48 in Japan, 2.42 and 2.45 in UK, 1.96 and 2.17 in Germany, and 2.12 and 2.04 in Italy.

[7]For the problem to be economically meaningful, $[\gamma^f \theta + \gamma^m (1 - \theta)] < \rho$ is assumed.

[8]Bargaining power can instead be assumed to depend on the relative wage income. Kemnitz and Thum (2015) use the relative wage income.

Letting P_t be the price of external childcare, the first-period budget constraint of the couple can be written as

$$w_t^m + (1 - \tilde{l}_t)w_t^f = P_t x_t + s_t, \tag{2.2}$$

where $\tilde{l}_t$ denotes the (internal) child-rearing time of a wife at home, x_t represents the amount of external childcare purchased, and s_t denotes savings for their retirement. The second-period budget constraint is

$$r_{t+1}s_t = c_{t+1}^2, \tag{2.3}$$

where r_{t+1} stands for the gross interest rate in period $t+1$. Letting $1/\varphi$ be a required time input to rear a pair of children, the total time input necessary to rear n_t pairs of children is given as $n_t/\varphi = \tilde{l}_t + x_t$, where $0 \leqslant \tilde{l}_t \leqslant 1$ and $x_t \geqslant 0$. The assumption of perfect substitutability is not made merely for simplicity but also for clarifying the role of the availability of external childcare services outside the home. This formulation enables us to consider a situation in which parents can choose either internal or external childcare services.[9]

Deconstruction of the utility maximization problem confronted by a couple must be done to two steps: cost minimization of child bearing and utility maximization. The cost of rearing n_t pairs of children is given as $C = w_t^f \tilde{l}_t + P_t x_t = (w_t^f - P_t)\tilde{l}_t + P_t(n_t/\varphi)$. By minimizing the cost for rearing n_t pairs of children subject to time constraint $0 \leqslant \tilde{l}_t \leqslant 1$, one can obtain the following cost function:

$$C(n_t) = \begin{cases} (i)\ \frac{P_t n_t}{\varphi} & when\ \ w_t^f > P_t \\ (ii)\ \frac{n_t}{\varphi} w_t^f & when\ \ w_t^f \leqslant P_t\ and\ \frac{n_t}{\varphi} < 1 \\ (iii)\ w_t^f + P_t(\frac{n_t}{\varphi} - 1) & when\ \ w_t^f \leqslant P_t\ and\ \frac{n_t}{\varphi} \geq 1 \end{cases} \tag{2.4}$$

In case (i), the couple chooses $\tilde{l}_t = 0$ and purchases $x_t = n_t/\varphi$ of external childcare to have n_t pairs of children. The wife supplies fulltime labor to the market. In case (ii), the wife allocates the time endowment between internal childcare $0 < \tilde{l}_t = n_t/\varphi < 1$ and market labor $0 < 1 - \tilde{l}_t < 1$. In case (iii), the couple chooses to rear n_t pairs of children at home; the wife does not supply labor to the market, i.e., $\tilde{l}_t = 1$. In cases (ii) and (iii), they are unwilling to purchase external childcare, i.e., $x_t = 0$.

As the second step, the couple chooses the number of children n_t and consumption c_{t+1} to maximize family utility subject to the following intertemporal budget constraint:

$$w_t^m + w_t^f = \frac{c_{t+1}}{r_{t+1}} + C(n_t). \tag{2.5}$$

The first-order conditions for utility maximization are

[9]Kemnitz and Thum (2015) assume the same formulation. Under a logarithmic utility function, however, it is infeasible for utility maximization to have $\tilde{l}_t = 0$ *and* $x_t = 0$ simultaneously.

$$\rho/c_{t+1} = \lambda_t/r_{t+1} \; and \; \bar{\gamma}/n_t = \lambda_t C'(n_t) \tag{2.6}$$

and budget constraint (2.5), where $\bar{\gamma} \equiv \gamma^f \theta + \gamma^m(1-\theta)$. Using cost function (2.4), and from (2.5) and (2.6), one obtains the following solutions. In case (i) in which $w_t^f > P_t$, one obtains

$$n_t = \frac{\bar{\gamma}\varphi}{\bar{\gamma}+\rho}\frac{w_t^f + w_t^m}{P_t} \; and \; s_t = \frac{\rho}{\bar{\gamma}+\rho}(w_t^f + w_t^m) \tag{2.7}$$

For case (ii) in which $w_t^f \leqslant P_t$ and $n_t/\varphi < 1$, we have

$$n_t = \frac{\bar{\gamma}\varphi}{\bar{\gamma}+\rho}\frac{w_t^f + w_t^m}{w_t^f} \text{ and } s_t = \frac{\rho}{\bar{\gamma}+\rho}(w_t^f + w_t^m). \tag{2.8}$$

For optimal plan (2.8) to be consistent with condition $n_t/\varphi < 1$, one must have $w_t^f/w_t^m > \bar{\gamma}/\rho$. For case (iii), in which $w_t^f \leqslant P_t$ and $n_t/\varphi \geqslant 1$, if the optimal plans are interior solutions to the problem, then we obtain $n_t = \frac{\bar{\gamma}\varphi}{\bar{\gamma}+\rho}\frac{P_t + w_t^m}{P_t}$. However, for n_t in the optimization condition to be consistent with condition $n_t/\varphi \geqslant 1$, one must have $P_t/w_t^m \leqslant \bar{\gamma}/\rho$. Therefore, $w_t^f/w_t^m \leqslant \bar{\gamma}/\rho$ because $w_t^f \leqslant P_t$. Couples are not willing to purchase external childcare in case (iii). Therefore, a corner solution is obtained[10]:

$$n_t = \varphi \text{ and } s_t = w_t^m. \tag{2.9}$$

2.2.2 Goods Production Sector

The consumption goods production technology is assumed to be given by the following constant-returns-to-scale production function.[11]

$$Y_t = F(K_t, L_t) + bL_t^m, \tag{2.10}$$

where Y_t represents aggregate output, K_t signifies aggregate physical capital, L_t denotes aggregate labor, and L_t^m expresses aggregate male labor in period t. $F(K, L)$ is a linear-homogeneous function in physical capital K and labor L; and $b > 0$ is the constant marginal productivity of male physical labor. Capital accumulation presumably increases labor productivity: $F_{LK} > 0$. Aggregate labor L_t is the sum of non-physical labor of female and male workers, $L_t = L_t^m + L_t^{fY}$, where L_t^{fY} denotes the female non-physical labor used in the goods production sector.

[10]Here, $w_t^m/P_t < (\bar{\gamma}+\rho)/\bar{\gamma}$ is assumed.

[11]The production function is assumed by Galor and Weil (1996) and by Kimura and Yasui (2010).

Letting N_t be the number of couples in period t, then the production function can be rewritten in percouple terms as

$$Y_t/N_t = F(k_t, 1 + l_t^{fY}) + b, \tag{2.11}$$

where $k_t = K_t/N_t$, $L_t^m/N_t = 1$ and $l_t^{fY} = L_t^{fY}/N_t$. The zero-profit conditions are

$$F_K(k_t, 1 + l_t^{fY}) = r_{t+1}, \tag{2.12}$$

$$F_L(k_t, 1 + l_t^{fY}) = w_t^f, \text{ and } F_L(k_t, 1 + l_t^{fY}) + b = w_t^m. \tag{2.13}$$

The gender wage gap is given as $(w_t^m - w_t^f)/w_t^f = b/F_L(k_t, 1 + l_t^{fY})$. It is noteworthy that even if women do not supply market labor as in case (iii), i.e., even when $l_t^{fY} = 0$, then the *potential* female wage rate can be given as the marginal product.

2.2.3 Childcare Production Sector

The production of childcare presumably requires goods input $B > 0$ per unit of childcare output and also that it requires female labor. The goods inputs include equipment and facilities for child rearing such as childcare centers and kindergartens. Therefore, it can be assumed to depend on the level of childcare service output. As the number of children cared for in the sector increases, the degree to which the input of goods must care for them also increases. The childcare production technology can be written as $X_t = \mu L_t^{fX}$, where X_t represents the aggregate product of childcare, L_t^{fX} stands for the labor employed, and μ stands for the labor productivity in the sector. Presumably, $\mu > 1$ because each childcare worker is expected to care for more than one child at a time.[12] This assumption means that the childcare services are a kind of club good. Goods input B is necessary to maintain the club goods property for the childcare production output size.

Profit of the childcare production sector is given as

$$\pi_t^X = P_t X_t - w_t^f L_t^{fX} - BX_t, \tag{2.14}$$

where BX_t stands for the total goods cost. The zero-profit condition in this sector can be written as

$$P_t\mu - w_t^f - B\mu = 0. \tag{2.15}$$

[12]Japanese national standards for the number of children per childcare worker at childcare centers are about three for children under age one and about six for children of 1 and 2 years old.

Therefore, $P_t \gtreqless w_t^f$ as $\frac{\mu B}{\mu-1} \gtreqless w_t^f$. When the female wage rate is sufficiently low relative to the per-unit goods cost, the price of external childcare is higher than the female wage rate, and *vice versa*. It is noteworthy that the services will not be produced in cases (ii) and (iii) because external childcare services are not demanded.

2.3 Dynamics

2.3.1 Market Equilibrium

The dynamics of the system is examined for each of three cases, as determined by the capital market equilibrium

$$K_{t+1} = s_t N_t, \tag{2.16}$$

where s_t is given as (2.7), (2.8), and (2.9) for each case.

The other important market is the labor market. Assuming that men are employed full time, the equilibrium condition is explainable by female labor. In case (i), because both spouses of each couple supply full-time labor to the market and purchase external childcare, one has $L_t^f = L_t^m = N_t$ and $\tilde{l}_t = 0$, where L_t^f denotes the female labor force. Couples do not purchase external childcare in cases (ii) and (iii). Therefore, $L_t^{fX} = 0$. Wives supply market part-time labor in case (ii), although they do not supply market labor in case (iii). Therefore,

$$L_t^f = N_t = L_t^{fY} + L_t^{fX}, \text{ and } L_t^{fY}, L_t^{fX} > 0 \text{ in } case\ (i), \tag{2.17a}$$

$$L_t^f = (1 - \tilde{l}_t)N_t = L_t^{fY} \text{ in } case\ (ii), \text{ and} \tag{2.17b}$$

$$L_t^f = 0 \text{ in } case\ (iii). \tag{2.17c}$$

The aggregate labor used in goods production is $L_t = L_t^m + L_t^{fY}$. The amount of labor employed in external childcare production is L_t^{fX}.

Finally, the equilibrium condition in the childcare market is $X_t = n_t N_t$ in case (i).[13] Childcare services production is linear in female labor. Therefore, the supply of childcare is determined to be equal to the demand, as expressed by the right-hand of the equilibrium condition.

An examination of the dynamics of each case is presented below.

Case (i) From (2.7), (2.13), (2.15), and the definition of the bargaining power,

[13]In case (i), from (2.7), and (2.13), one has $n_t = n(l_t^{fY}; k_t)$. The equilibrium condition in the external child-care market can be rewritten as $n_t = \mu(1 - l_t^{fY})$ using the labor market equilibrium condition (2.17a). Therefore, one can obtain $l_t^{fY} = l^{fY}(k_t)$. Consequently, variables in period t, w_t^f, w_t^m, r_{t+1}, P_t, l_t^{fY}, l_t^{fX}, Y_t, X_t, n_t, and s_t, are determined for a given k_t.

$$n_t = \frac{\bar{\gamma}\varphi}{\bar{\gamma}+\rho}\frac{2F_L(k_t, 2-n_t/\mu\varphi)+b}{F_L(k_t, 2-n_t/\mu\varphi)/\mu+B}, \tag{2.18}$$

from which one obtains $n_t = n(k_t)$. Differentiating (2.18) with respect to k_t yields

$$\frac{dn_t}{dk_t} = \frac{[2-\frac{\bar{\gamma}+\rho}{\bar{\gamma}\varphi\mu}n_t + (\gamma^f-\gamma^m)\frac{b+2F_L}{(b+F_L)^2}\frac{\rho b\theta'}{\rho+\bar{\gamma}}]F_{LK}}{\frac{\bar{\gamma}+\rho}{\bar{\gamma}\varphi}(\frac{F_L}{\mu}+B)+\frac{F_{LL}}{\varphi\mu}[2-\frac{\bar{\gamma}+\rho}{\bar{\gamma}\varphi\mu}n_t+(\gamma^f-\gamma^m)\frac{b+2F_L}{(b+F_L)^2}\frac{\rho b\theta'}{\rho+\bar{\gamma}}]}. \tag{2.19}$$

The sign of dn_t/dk_t depends on the signs of $F_L + F_{LL}(2-\frac{n_t}{\varphi\mu})$, $2-\frac{(\bar{\gamma}+\rho)n_t}{\bar{\gamma}\varphi\mu}$, and $(\gamma^f-\gamma^m)\frac{b+2F_L}{(b+F_L)^2}\frac{\rho b\theta'}{\rho+\bar{\gamma}}$.[14] One can demonstrate that $F_L + F_{LL}(2-\frac{n_t}{\varphi\mu}) > 0$ holds if $F(K, L)$ is the CES function. Using (2.15) and $w_t^m = w_t^f + b$, one can obtain $\partial n_t/\partial w_t^f = P_t[2-\frac{(\bar{\gamma}+\rho)n_t}{\bar{\gamma}\varphi\mu}] > 0$ because children are normal goods. It seems plausible in this case that parents consider price changes when purchasing external childcare services. Unless γ^f is too small to make the numerator of the right-hand side of (2.19) negative, one can plausibly obtain $dn_t/dk_t > 0$, as reported by Yakita (2018a). If $\gamma^f > \gamma^m$ ($\gamma^f < \gamma^m$), then $dn_t/dk_t > 0$ is greater (smaller) than in the case of $\gamma^f = \gamma^m$.[15]

The equilibrium condition in the capital market can be written as

$$n(k_t)k_{t+1} = \frac{\rho}{\bar{\gamma}+\rho}[2F_L(k_t, 2-\frac{n_t}{\varphi\mu})+b], \tag{2.20}$$

from which one obtains

$$\frac{dk_{t+1}}{dk_t} = \frac{\rho}{\bar{\gamma}\varphi\mu}[1-(\gamma^f-\gamma^m)\frac{(F_L+\mu B)b\theta'}{(F_L+b)^2}](F_{LK}-F_{LL}\frac{1}{\varphi\mu}\frac{dn_t}{dk_t}), \tag{2.21}$$

and

$$\frac{dw_t^f}{dk_t} = (F_{LK}-F_{LL}\frac{1}{\varphi\mu}\frac{dn_t}{dk_t}) > 0. \tag{2.22}$$

Although $dk_{t+1}/dk_t > 0$ if $\gamma^f \leqslant \gamma^m$, the sign of the right-hand side of (2.21) cannot be determined *a priori* if $\gamma^f > \gamma^m$. For analytical convenience, the effect of the change in the bargaining power on the fertility rate is assumed to be not so great as to hamper capital accumulation, i.e., $dk_{t+1}/dk_t > 0$ even when $\gamma^f > \gamma^m$.

Lemma 2.1 *When women supply all their time to the labor market and couples purchase market childcare, wage rates and the number of children increase as capital accumulates if the women's utility weight of having children is greater than that of men.*

[14] Presumably, the denominator of the right-hand side of (2.19) is positive for the continuity of n_t with respect to k_t.

[15] One might have $dn_t/dk_t < 0$ if γ^f is sufficiently small relative to γ^m, but the denominator of the right-hand side of (2.19) remains positive. Nevertheless, such as case is implausible.

Case (ii) From (2.13), n_t in (2.8) can be rewritten as

$$n_t = \frac{\bar{\gamma}\varphi}{\bar{\gamma}+\rho}[2 + \frac{b}{F_L(k_t, 2 - n_t/\varphi)}], \tag{2.23}$$

from which one can obtain $n_t = n(k_t)$. One can then readily demonstrate that

$$\frac{dn_t}{dk_t} = \frac{-[1-(\gamma^f-\gamma^m)\frac{F_L\rho\theta'(2F_L+b)}{\bar{\gamma}(\rho+\bar{\gamma})(F_L+b)^2}]\varphi F_{LK}}{\frac{\bar{\gamma}+\rho}{\bar{\gamma}}\frac{(F_L)^2}{b} - [1-(\gamma^f-\gamma^m)\frac{F_L\rho\theta'(2F_L+b)}{\bar{\gamma}(\rho+\bar{\gamma})(F_L+b)^2}]F_{LL}}. \tag{2.24}$$

In case (ii), wives allocate their time between internal child-rearing ($\tilde{l}_t = n_t/\varphi < 1$) and market (part-time) labor $(1-\tilde{l}_t)$. Because couples do not demand external child-care, one obtains $l_t^{fY} = 1 - \tilde{l}_t = 1 - n(k_t)/\varphi$, where capital accumulation demands more female labor in goods production, i.e., $dl_t^{fY}/dk_t > 0$.

The equilibrium condition in the capital market is given as

$$k_{t+1} = \frac{\rho}{\bar{\gamma}\varphi}F_L[k_t, 1 + l^{fY}(k_t)], \tag{2.25}$$

from which one can obtain

$$\frac{dk_{t+1}}{dk_t} = \frac{\rho}{\bar{\gamma}\varphi}[1 - \frac{\gamma^f-\gamma^m}{\bar{\gamma}}\frac{bF_L\theta'}{(F_L+b)^2}]\frac{dw_t^f}{dk_t} \tag{2.26}$$

and

$$\frac{dw_t^f}{dk_t} = F_{LK} + F_{LL}\frac{dl_t^{fY}}{dk_t} = F_{LK} + F_{LL}(-\frac{dn_t}{dk_t})$$

$$= \frac{\frac{\bar{\gamma}+\rho}{\bar{\gamma}}\frac{(F_L)^2}{b}F_{LK}}{\frac{\bar{\gamma}+\rho}{\bar{\gamma}}\frac{(F_L)^2}{b} - [1-(\gamma^f-\gamma^m)\frac{F_L\rho\theta'(2F_L+b)}{\bar{\gamma}(\rho+\bar{\gamma})(F_L+b)^2}]F_{LL}} > 0. \tag{2.27}$$

The denominator of the right-hand side of (2.27) is assumed to be positive for the continuity of the wage rate with respect to k_t. For analytical purposes, $dk_{t+1}/dk_t > 0$ is assumed, as in the previous case.

If $\gamma^f \leqslant \gamma^m$, then we have $dn_t/dk_t < 0$ from (2.24). Conversely, if $\gamma^f > \gamma^m$, then

$$1 - \frac{\gamma^f-\gamma^m}{\bar{\gamma}}\frac{F_L\theta'}{(F_L+b)^2}b \geqslant 1 - \frac{\gamma^f-\gamma^m}{\bar{\gamma}}\frac{F_L\theta'}{(F_L+b)^2}\frac{\rho}{\rho+\bar{\gamma}}(2F_L+b) \tag{2.28}$$

because one can safely assume that $s_t(=\rho(2F_L+b)/(\rho+\bar{\gamma})) \geqslant b$. If the right-hand side of (2.28) is positive, then $dn_t/dk_t < 0$ and $dk_{t+1}/dk_t > 0$.[16] However, if

[16]The denominator on the right-hand side of (2.24) is assumed for the continuity of n_t with respect to k_t, as in the previous case.

savings are sufficiently great or if the wage rate is sufficiently high that the right-hand side is negative, then one might have $dn_t/dk_t > 0$ and $dk_{t+1}/dk_t > 0$. As capital accumulates, the wage rates become higher (see (2.27)). Thereby, $dn_t/dk_t > 0$ might hold.

Lemma 2.2 *When women partly supply their time to the labor market and women care for a child at home, (a) the wage rates increase but the number of children decreases as capital accumulates if the women's utility weight of having children is smaller than that of men; and (b) the wage rates and the fertility rate might increase as capital accumulates if the women's utility weight related to having children is greater than that of men.*

Case (iii) From (2.9) and (2.13),

$$k_{t+1} = \frac{1}{\varphi}[F_L(k_t, 1) + b], \tag{2.29}$$

from which one can obtain

$$\frac{dk_{t+1}}{dk_t} = \frac{1}{\varphi} F_{LK} > 0, \tag{2.30}$$

and[17]

$$\frac{dw_t^f}{dk_t} = F_{LK} > 0. \tag{2.31}$$

Lemma 2.3 *The time paths of variables are independent of the utility weights that spouses assign to having children.*

2.3.2 Dynamics of the Development Path

As explained in this section, the time paths of the wage rates and the fertility rate are examined, assuming that the initial per-couple capital stock is sufficiently small. For expositional purposes, it is assumed that the economy does not fall into a trap of long-term equilibrium in case (ii) or case (iii). Moreover, a gender wage gap is assumed to exist, satisfying $\frac{b\bar{\gamma}(\mu-1)}{\mu(\rho-\bar{\gamma})} < B$ for given utility weights γ^f and γ^m.[18] If the inequality of the assumption holds strictly, then some range of the female wage

[17] Although no women work in the market, the potential female wage rate is given as the marginal product of labor in the goods production sector.

[18] If this assumption is not satisfied, then the fertility rate might not decrease with the female wage rate. However, economically developed and even economically developing countries have sustained decreased fertility during recent decades. Therefore, this assumption is apparently not only plausible, but realistic.

rate exists, within which women choose to rear children at home rather than incur the high costs of external childcare.

Presuming first that the initial female wage rate is sufficiently low to satisfy $w_t^f/w_t^m \leqslant \bar{\gamma}/\rho$ or, equivalently, $w_t^f \leqslant \bar{\gamma}b/(\rho - \bar{\gamma})$, then the economy is in case (iii) in which the fertility rate is constant i.e., $n_t = \varphi$. The female wage rate is too low for women to supply labor to the market. If the marginal utility of having children is still high at $\tilde{l}_t = 1$, i.e., if the women are in the corner solution, then the couple foregoes the utility that might derive from having more children.

From (2.30) and (2.31), it is readily apparent that capital accumulates and the female wage rate rises as time passes. Increases in the female wage rate narrow the gender wage gap, thereby increasing the bargaining power of women. If the female utility weight related to having children is lower than that of men, i.e., if $\gamma^f \leqslant \gamma^m$, then the reduced gender wage gap moves the economy from case (iii) to case (ii) in which $w_t^f/w_t^m > \bar{\gamma}/\rho$. However, it is not necessarily the case that $\gamma^f > \gamma^m$. Although the economy moves to case (ii) as time passes if $1 > (\gamma^f - \gamma^m)\theta'/\rho$, the economy stays in case (iii) if $1 \leqslant (\gamma^f - \gamma^m)\theta'/\rho$.[19] When $1 < (\gamma^f - \gamma^m)\theta'/\rho$, $\bar{\gamma}/\rho$ becomes greater than the gender wage gap w_t^f/w_t^m at $w_t^f/w_t^m = \bar{\gamma}/\rho$ but the female wage rate rises concomitantly with capital accumulation.[20]

When $1 > (\gamma^f - \gamma^m)\theta'/\rho$ at $w_t^f/w_t^m = \bar{\gamma}/\rho$, the female wage rate becomes sufficiently high to satisfy $w_t^f/w_t^m > \bar{\gamma}/\rho$ or, equivalently, $w_t^f > \bar{\gamma}b/(\rho - \bar{\gamma})$. Then, the economy goes into the phase of case (ii). At $w_t^f(k_t) = b\bar{\gamma}/(\rho - \bar{\gamma})$, one has $n_t = \varphi$ and $k_{t+1} = \rho b/[\varphi(\rho - \bar{\gamma})]$. Therefore, the paths of n_t and k_{t+1} are continuous in k_t at $w_t^f(k_t) = b\bar{\gamma}/(\rho - \bar{\gamma})$ if $1 > (\gamma^f - \gamma^m)\theta'/\rho$.

In the phase of case (ii), if the utility weight of women on having children is not so great, then the fertility rate decreases along with capital accumulation (see (2.24)). Having too many children raises the opportunity cost of rearing children at home. Therefore, couples reduce the number of children until the marginal utility of having children equals the opportunity cost. Women increase the labor supply to the market, consequently reducing the internal child-rearing time at home. The increased female wage income increases per-couple capital, which in turn increases the wage rate, as might be apparent from (2.26) and (2.27). However, unless the female wage rate becomes higher than the price of external childcare, couples are not willing to purchase external childcare.

By contrast, if the women's utility weight assigned to having children is sufficiently great, then one cannot rule out the possibility of a positive relation between fertility and capital accumulation. The fertility rate might rise as the economy accumulates capital and wage rates increase. In such a case, although the opportunity cost of child-rearing time at home increases, women prefer to have more children and therefore increase the internal child rearing time without purchasing market childcare services. The intuition underlying this result is the income effect. Therefore,

[19] In this case, the long-term per-couple capital stock is given as $k_{(iii)} = (1/\varphi)[F_L(k_{(iii)}, 1) + b]$, where $F_L(k_{(iii)}, 1) < b\bar{\gamma}/(\rho - \bar{\gamma})$ and the fertility rate is $n_{(iii)} = \varphi$.

[20] In this case, one cannot rule out the possibility that the economy might move to case (i) while skipping case (ii).

whether the fertility rate decreases in case (ii) depends on the relative utility weight that women assign to having children. Capital accumulation raises the wage rates of both women and men throughout the phase of case (ii).

After the female wage rate becomes sufficiently high to satisfy $w_t^f \geqslant \mu B/(\mu - 1)$, the economy moves into the phase of case (i), in which $w_t^f \geqslant P_t$.[21] Substituting $w_t^f(k_t) = \mu B/(\mu - 1)$ into (2.18), (2.20), (2.23), and (2.25), the time paths of n_t and k_{t+1} can be shown to be continuous in k_t at $w_t^f(k_t) = \mu B/(\mu - 1)$, at which per couple capital level, the values of the fertility rate and per couple capital stock are $n_t = \bar{\gamma}\varphi[2 + b(\mu - 1)/(\mu B)]/(\bar{\gamma} + \rho)$ and $k_{t+1} = \rho\mu B/[\bar{\gamma}\varphi(\mu - 1)]$ for both cases (ii) and (i). Therefore, the time paths are also continuous in k_t at $w_t^f(k_t) = \mu B/(\mu - 1)$. At this wage rate, women start to supply fulltime labor to the labor market while purchasing external childcare to rear their children. The female market labor is allocated between goods production and the external childcare production sector to equate the female wage rates in both sectors, although mothers rear their children at home immediately before the wage rate satisfies $w_t^f = P_t$.

The fertility rate increases with the per-couple capital stock, i.e., $dn_t/dk_t > 0$, although the magnitude of the effect of capital accumulation on fertility depends on the relative utility weights of women and men on having children (see (2.19)). The greater the women's utility weight is, the greater the effect of capital accumulation on fertility becomes.[22] The female wage rate increases with the per-couple capital stock (see (2.22)). Therefore, the fertility rate rises through the income effect as the female wage rate increases. As might be readily apparent from (2.15), the price of external childcare becomes lower relative to the female wage rate as the female wage rate increases. For that reason, couples can afford to purchase more external childcare as the female wage rate rises.

Because per-couple capital stock and the female wage rate remain constant in the long-term equilibrium, the fertility rate also becomes constant in the long term, as might be apparent from (2.20). If the stability condition is satisfied, then one has the long-term per-couple capital satisfying[23]

$$k = \frac{\rho}{\bar{\gamma}\varphi\mu}\{F_L[k, 1 + l^{fY}(k)] + \mu B\}, \tag{2.32}$$

and the long-term fertility rate satisfying

$$n = \frac{\bar{\gamma}\varphi}{\bar{\gamma} + \rho}\frac{2F_L(k, 2 - n/\mu\varphi) + b}{F_L(k, 2 - n/\mu\varphi)/\mu + B}. \tag{2.33}$$

[21]The possibility cannot be ruled out, *a priori*, that the economy is trapped in a lower long-term equilibrium in case (ii), in which the per-couple capital stock is given as $k_{(ii)} = (\rho/\bar{\gamma}\varphi)F_L[k_{(ii)}, 1 + l^{fY}(k_{(ii)})]$, where $F_L[k_{(ii)}, 1 + l^{fY}(k_{(ii)})] < \mu B/(\mu - 1)$. In this case, the economy does not have a fertility rebound.

[22]The possibility that the effect is negative when the women's utility weight is too small relative to that of men cannot be ruled out theoretically. Nevertheless, it is apparently implausible.

[23]The stability condition is given as $dk_{t+1}/dk_t < 1$.

The gender wage gap remains even in the long term, but the gap is smaller when the long-term percouple capital stock is greater.

From (2.32) and (2.33), one obtains[24],[25]

$$\frac{dn}{d\gamma^f} \gtreqless 0, \quad \frac{dk}{d\gamma^f} < 0, \text{ and } \frac{dw^f}{d\gamma^f} = F_{LK}\frac{dk}{d\gamma^f} - \frac{F_{LL}}{\bar{\gamma}\varphi}\frac{dn}{d\gamma^f} \gtreqless 0. \tag{2.34}$$

Although the long-term fertility rate might be higher than, equal to, or lower than φ, a higher women's utility weight assigned to having children is not necessarily associated with a higher fertility rate. A high utility weight of women lowers the per-couple capital stock. The intuition underlying these results is that having more children increases the cost of external child rearing, which in turn reduces capital formation and the savings of couples. The lower wage rate deriving from the reduced capital stock might discourage couples from having more children.

Theorem 2.1 *(a) If the preference of women for having children is weaker than that of men, then the fertility rate turns upward when couples purchase market childcare at a price that is lower than the female wage rate. By contrast, (b) if the preference of women for having children is sufficiently stronger than that of men, then a fertility rebound might occur even when mothers care for their children at home.*

The time path of the fertility rate is presented in Fig. 2.1, where the horizontal line shows the level of per-couple capital stock.[26] The path of the fertility rate shows a reverse-J-shape or U-shape in the female wage rate (and per-couple capital stock).[27] It is noteworthy that if the utility weight assigned by women to having children is sufficiently great, then the fertility rebound might occur even when couples do not purchase market childcare services (Fig. 2.1b). The wage income of couples increases as capital accumulates. The increased income might induce couples to increase the number of children and spend more of the mothers' time on child-rearing at home, thereby reducing the female market labor supply through the income effect. With the increased bargaining power of women, the income effect on having children overwhelms the negative effects of the increased opportunity cost. This case contrasts against the result obtained with a unitary model such as that presented by Yakita (2018a), in which the fertility rate declines steadily in the phase of case (ii).

---

[24]See Appendix.

[25]Presumably, (2.21) has a positive sign.

[26]For these figures, a gender wage gap satisfying $[b\bar{\gamma}(\mu - 1)]/[\mu(\rho - \bar{\gamma})] < B$, i.e., case (ii) is assumed to exist.

[27]Figure 1 presents the case in which mothers care for their children at home in case (ii), i.e., $\tilde{l}_t < 1$, even when the female wage rate approaches $w_t^f(k_t) = \mu B/(\mu - 1)$. If $\tilde{l}_t = 1$ at $w_t^f(k_t) = \mu B/(\mu - 1)$, then the time path might have a flat part correspondingly.

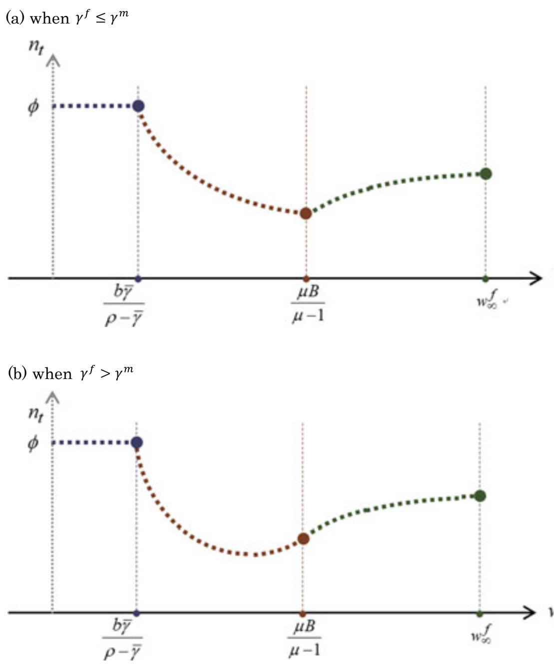


Fig. 2.1 Change in the fertility rate

2.4 Conclusion

A collective model was used to demonstrate that changes in relative bargaining power of women and men occurring along with economic development strongly influence fertility dynamics. Bargaining power is assumed to depend on the gender wage gap, which derives from labor productivity differences between women and men. A fertility rise can occur when couples start to purchase market childcare services and substitute it for internal child-rearing by mothers at home. This result has been reported in the literature (Kemnitz and Thum 2015; Yakita 2018a), but it is noteworthy that another possibility of fertility rebound can be presented.

The fertility rebound might occur even without couples' purchasing market childcare services if the women's utility weight related to having children is sufficiently high. The increased wage income induces couples to have more children and to reduce the mothers' labor supply to the market, which reflects the high preference of mothers for having children by increasing the bargaining power of women. Assuming the existence of a long-term equilibrium, one can also show that a higher preference of women for having children does not necessarily raise and possibly lowers long-term fertility as a consequence of family bargaining, *ceteris paribus*. These results are novel. Moreover, they contrast against the result reported by Kemnitz and Thum (2015) that the family bargaining systematically biases fertility downward.

Results of the present analyses also imply that the gender wage *income* gap stems only from biological differences between genders in the long term. Efficient provision of market childcare might eliminate, or at least alleviate, statistical discrimination between genders in the long term, with both parents increasingly bearing the burdens of housework and child-bearing as equally as they can.[28]

Neither human capital formation nor the parents' preference for the quality of children has been considered. Parents might spend their income on human capital investment in children as well as childcare services.[29] Simultaneous consideration of the respective effects of preferences for the quantity and quality on the fertility dynamics remains as a subject for future study.

Acknowledgements Financial support from the Japan Society for the Promotion of Science KAKENHI Grant No.16H03635 is gratefully acknowledged.

Appendix

Effects of the female utility weight related to long-term fertility and the wage rate

From (2.32) and (2.33),

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} dk/d\gamma^f \\ dn/d\gamma^f \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}, \tag{2.35}$$

where

$$\begin{aligned}
a_{11} &= 1 - \frac{\rho}{\varphi\mu\bar{\gamma}}[1 - \frac{(2F_L + b)b\theta'}{(F_L + b)^2}(\gamma^f - \gamma^m)]F_{LK} \\
a_{12} &= \frac{\rho}{\bar{\gamma}\rho\varphi}[1 - \frac{(2F_L + b)b\theta'}{(F_L + b)^2}(\gamma^f - \gamma^m)]\frac{F_{LL}}{\mu\varphi} \\
a_{21} &= -[2 - \frac{\bar{\gamma} + \rho}{\bar{\gamma}\varphi\mu} + \frac{(2F_L + b)\rho b\theta'}{(\bar{\gamma} + \rho)(F_L + b)}(\gamma^f - \gamma^m)]F_{LK}, \\
a_{22} &= \frac{\bar{\gamma} + \rho}{\bar{\gamma}\varphi\mu}(F_L + \mu B) + [2 - \frac{\bar{\gamma} + \rho}{\bar{\gamma}\varphi\mu} + \frac{(2F_L + b)\rho b\theta'}{(\bar{\gamma} + \rho)(F_L + b)}(\gamma^f - \gamma^m)]\frac{F_{LL}}{\varphi\mu} \\
g_1 &= -\frac{\rho}{\bar{\gamma}^2\varphi\mu}(F_L + \mu B)\theta < 0, \\
g_2 &= \frac{\rho\varphi\mu}{(\bar{\gamma} + \rho)^2}\frac{2F_L + b}{F_L + \mu B}\theta > 0.
\end{aligned}$$

[28] The Human Development Report by United Nations (http://hdr.undp.org/en/data. Accessed 21 December 2017) shows a negative relation between the Human Development Index and the Gender Inequality Index during 2000–2015.

[29] Yakita (2018b) presents an analysis of the consequences of family Nash bargaining on fertility and human capital accumulation by considering both the quantity and quality of children.

From (2.35) one can obtain

$$\frac{dn}{d\gamma^f} = H^{-1}(a_{21}g_1 - a_{11}g_2)$$

and

$$\frac{dk}{d\gamma^f} = H^{-1}(a_{22}g_1 - a_{12}g_2) \tag{2.36}$$

where $H = a_{11}a_{22} - a_{12}a_{21} > 0$ from the stability condition, $dk_{t+1}/dk_t < 1$. From the assumption of $0 < dk_{t+1}/dk_t$, it follows that $a_{12} > 0$. When $dn_t/dk_t > 0$, $a_{21} < 0$. a_{22} is the denominator on the right-hand side of (2.19), which can be assumed to be positive for the continuity of the solution with respect to $\gamma^f - \gamma^m$. Presumably, $\partial k_{t+1}/\partial k_t > 0$, which is plausible.

References

Apps, P., and R. Rees. 2004. Fertility, taxation and family policy. *Scandinavian Journal of Economics* 106 (4): 745–763.

Arrow, K. J., 1972. Some mathematical models of race discrimination in the labor market. In *Racial discrimination in economic life*, ed. Anthony Pascal. Lexington: Lexington Books (Chap. 6).

Day, C. 2012. Economic growth, gender wage gap and fertility rebound. *Economic Record* 88 (s1): 88–99.

Day, C. 2016. Fertility and economic growth: The role of workforce skill composition and child care prices. *Oxford Economic Papers* 68 (2): 546–565.

Galor, O., and D. Weil. 1996. The gender gap, fertility, and growth. *American Economic Review* 86 (3): 374–387.

Hirazawa, M., and A. Yakita. 2009. Fertility, child care outside the home, and pay-as-you-go social security. *Journal of Population Economics* 22 (3): 565–583.

Kawaguchi, A. 2008. *Jenda Keizai Kakusa (Economic gender gap)*. Tokyo: Keiso-Shobo. (in Japanese).

Kemnitz, A., and M. Thum. 2015. Gender power, fertility, and family policy. *Scandinavian Journal of Economics* 117 (1): 220–247.

Kimura, M., and D. Yasui. 2010. The Galor-Weil gender-gap model revisited: From home to market. *Journal of Economic Growth* 15 (4): 323–351.

Komura, M. 2013. Fertility and endogenous gender bargaining power. *Journal of Population Economics* 26 (3): 943–961.

Schwab, S. 1986. Is statistical discrimination efficient? *American Economic Review* 76 (1): 228–234.

Yakita, A. 2018a. Female labor supply, fertility rebounds, and economic development. *Review of Development Economics* 22 (4): 1667–1681.

Yakita, A. 2018b. Fertility and education decisions and child-care policy effects in a Nash-bargaining family model. *Journal of Population Economics* 31 (4): 1177–1201.

Yamaguchi, K. 2017. *Hatarakikata no Danjo Hubyodo: Riron to Jissho-Bunseki (Gender inequality in working: Theory and empirical analysis)*. Tokyo: Nihon Keizai-Shinbun Sha. (in Japanese).

Chapter 3
Skill-Biased Innovation, Growth, and Inequality

Yasuyuki Osumi

3.1 Introduction

Parallel to the widespread inequality phenomena in advanced countries, there has been a growing literature on skill-biased technology and inequality.[1] Among them, having analyzed the dynamics of technical unemployment in the framework of the induced factor bias of innovation, Stiglitz (2014) argued that the formulation of the induced skill-biased innovation is one of the promising researches for analyzing the various inequalities in the OECD countries. One of the implications of the induced innovation framework in line with Kennedy (1964) and Samuelson (1966) is that relatively increasing of the factor share can induce firms to introduce their own factor-augmenting technical progress in the maximization of the instantaneous cost reduction rate of change on the concavity of the innovation frontier. In this setting, an increase in the capital share of income relative to the labor share of income, for instance, leads to a bias toward a capital-augmenting technical progress. If the elasticity of substitution between capital and labor is smaller than unity, the steady state is stable, with reducing the capital share that leads to lower capital-augmenting innovation. However, if the elasticity of substitution is larger than unity, the long run equilibrium is unstable, and thus this capital-augmenting technical progress further increases the capital share without bound. Stiglitz (2014) conjectured that the same logic can apply to the case of the innovation of augmenting skilled labor and unskilled labor. Namely, an increase in the income of skilled labor relative to unskilled labor

[1] See Acemoglu (2002) and Hornstein et al. (2005). Recently, there are literature on capital-augmenting technical progress as an improvement in an Automation and Artificial Intelligence. See Acemoglu and Restrepo (2017a), Acemoglu and Restrepo (2017b), Kotlikoff and Sachs (2012), Graetz and Michaels (2015), and Korinek and Stiglitz (2017). In this paper, as an exogenous parameter, we analyze the impact of capital-augmenting technology on skill-biased innovation and income inequality.

Y. Osumi (✉)
University of Hyogo, Kobe, Japan
e-mail: osumi@econ.u-hyogo.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_3

can lead to skill-biased innovation. If the elasticity of substitution between skilled and unskilled labor is larger than unity, the skill-biased technology further increases the skilled labor income, with more inducing the skill-biased innovation.

However, to consider this conjecture formally on the induced skill-biased innovation in a growth economy, at least two modifications are needed. The one is a three-factor framework and the associated elasticity of substitution in the three-factor case. In a growth model, capital accumulation plays a role in the dynamics and associated income inequalities. Therefore, the analysis of induced skill-biased innovation needs a three-factor framework including capital stock. In this three-factor framework, the substitutability and complementarity among skilled, unskilled labor and capital play significant roles in the stability of the steady state and formulation of endogenous biased innovation and inequality. Of three-factor production structures, if the economy has the widely estimated relevant capital-skill complementarity structure that implies an elasticity of substitution between capital and skilled labor smaller than unity but that of substitution between capital and unskilled labor larger than unity,[2] we investigate what the stability condition of the steady state is and how the induced skill bias innovation and associated income inequality behave in the dynamic system.

The other one is the innovation possibility frontier.[3] The weakness of the innovation possibility frontier is that the frontier itself does not have the possibility of shifting.[4] This weakness makes the determination of the growth rate unrealistic and the constancy of biased technologies and of the factor income ratio in the steady state. Especially, the latter case is produced by the unique relationship between skill-augmenting technical progress and the factor income ratio, which is derived from the maximization of the instantaneous cost reduction rate of change on the concavity of the stationary innovation frontier. Therefore, in this stationary innovation frontier framework, skill bias innovation and income inequality do not change unless the frontier curve shifts or the growth economy is in a transitional state.

To deal with these two respects, based on a three-factor framework with a two-level CES production function that can capture relevant capital-skill complementarity technology,[5] we develop endogenous skill-biased innovation in the standard (Solow 1956) neoclassical growth model to analyze the dynamics of income inequalities. As noted, Stiglitz (2014) provided the conjecture of skill biased innovation, but did not extend the model analysis. Developing the possible shift in the innovation possibility frontier in line with Kennedy (1964), Samuelson (1966), and Drandakis and Phelps (1966) having a tradeoff relationship between skill-augmenting and unskilled-augmenting technical progress,[6] we investigate the characteristics of the induced skill-biased innovation and the dynamical properties of income inequality and labor shares in the three-factor growth model.

[2]See Krusell et al. (2000) and Duffy et al. (2004).

[3]See Acemoglu (2010), Korinek and Stiglitz (2017).

[4]See Nordhaus (1973) and Acemoglu (2015).

[5]For pioneer works, see Griliches (1969) and Sato (1967).

[6]Caselli and Coleman (2006) considered the same type of frontier in capital-skill complementarity. However, they did not analyze the induced biased innovation.

Besides, according to the formulation of Adachi et al. (2019), we develop the induced innovation frontier incorporating the externality of capital accumulation that can expand outward the innovation frontier.[7] R & D activity can be embodied in the newly capital stock, and thus we implicitly take account of this property as the possible shift in the innovation frontier. Developing these modifications, our model can then analyze growth and inequality in the induced innovation framework.

Our three main results are as follows. First, capital-skill complementarity technology and the shift in the innovation possibility frontier play significant roles in the stability of the steady state and dynamics of labor shares and wage inequality. The stability is fulfilled with some empirically relevant capital-skill complementarity and this steady state is oscillatory. This implies that even if there is the elasticity of substitution between skilled and unskilled labor larger than unity, which is the empirically relevant case,[8] the elasticity of substitution between capital and skilled labor smaller than unity makes the steady state stable. In this steady state, there is no biased innovation. However, the induced skill-biased innovation can occur in the transitional state.

Second, when the innovation frontier does not shift, the induced skill bias innovation and skilled/unskilled labor income ratio, which means labor income inequality, do not change in the steady state. However, when the innovation frontier shifts with capital accumulation, a population growth has a favorable influence on growth and the income distribution. Because population growth can expand the innovation possibility frontier outward due to its positive influence on capital accumulation. Thus, in some conditions, the increase in population growth can produce economic growth and decrease income inequality. However, the opposite case applies. Namely, a decline of population growth, which is the relevant case in most advanced countries, is likely to lead to a decline in the growth rate and an increase in income inequality.

Third, our analysis has an implication of the heterogeneity of population growth in both labor types on the induced bias of innovation, growth, and distribution. If population growth in unskilled labor is larger than that in skilled labor, which means a scarcity of skilled labor in the growth process, even in the steady state the skill biased innovation can be introduced. Thus, it can provide more growth. However, this is likely to produce more income inequality.

The structure of the paper is as follows. Section 3.2 presents the basic model. Section 3.3 analyzes the dynamic system and the implications for the induced bias innovation and income inequality. Section 3.4 concludes.

[7] Samuelson (1966) suggested that the capital share of income can be one of the shift parameter of the innovation possibility frontier.

[8] See Ciccone and Peri (2005). In the two-factor case, if the elasticity of substitution between skilled and unskilled labor is larger than unity, the steady state may always be unstable and thus the induced skill-biased innovation can make the income of skilled labor increasing over time.

3.2 Basic Model

3.2.1 Three-Factor Production Function

We consider a three-factor production function of twice differentiable and homogenous of degree one:

$$Y = F(A_1L_1, A_2L_2, BK) \tag{3.1}$$

where Y is output, L_1 is skilled labor, L_2 is unskilled labor, K is capital stock, A_1 is skilled labor efficiency, A_2 is unskilled labor efficiency, and B is capital efficiency.

We assume that the three-factor production function is a weakly separable sub-aggregate production function. For the production function, we specify a nested two-level constant elasticity of substitution (CES) production function having two elasticity parameters: the elasticity of substitution between capital and skilled labor σ_1, and the elasticity of substitution between capital and unskilled labor σ_2:

$$\begin{aligned} Y &= F(L_1, L_2, K) \\ &= \Big[[\delta_2\{\delta_1 L_1^{(1-\sigma 1)/\sigma 1} + (1-\delta_1)K\}^{(1-\sigma 1)/\sigma 1}]^{(1-\sigma 2)\sigma 1/(1-\sigma)\sigma 2} \\ &\quad + (1-\delta_2)L_2^{(1-\sigma 2)/\sigma 2} \Big]^{\sigma 2/(1-\sigma 2)} \end{aligned} \tag{3.2}$$

In this specification, $\sigma_2 > \sigma_1$ gives the capital-skill complementarity technology[9] that has been widely estimated (Krusell et al. 2000; Hornstein et al. 2005), and we deal with this ongoing technical progress. In particular, our analysis focuses on inequalities in empirically relevant capital-skill complementarity $\sigma_2 > 1 > \sigma_1$.[10]

Under the assumption of constant returns to scale, we can rewrite the production function as follows.

$$Y = BKf(A_1L_1/BK, A_2L_2/BK) \tag{3.3}$$

Since we consider the long-run perfect competitive economy that has full employment, we simply denote L_1, L_2 as follows:

$$L_1 = uL, L_2 = (1-u)L, \tag{3.4}$$

[9]Defining $c_{ij} \equiv F_{ij}F/F_iF_j$ as the partial elasticity of complementarity between i and j, capital-skill complementarity is indicated as an inequality whereby the elasticity of complementarity between capital and skilled labor is larger than that between capital and unskilled labor $c_{1K}(= F_{1K}F/F_1F_K) > c_{2K}(= F_{2K}F/F_2F_K)$. In our two-level CES production technology, $\sigma_2 > \sigma_1$ implies $c_{1K} > c_{2K}$ since $c_{1K} - c_{2K} = (\sigma_1^{-1} - \sigma_2^{-1})/(1 - F_2L_2/F)$. Thus, $\sigma_2 > \sigma_1$ implies capital-skill complementarity. Note that in our three-factor case, the elasticity of substitution is not always equal to the inverse of the elasticity of complementarity. Specifically, $c_{1K} = (\sigma_1^{-1} - \sigma_2^{-1})/(1 - F_2L_2/F) + \sigma_2^{-1} \neq \sigma_1^{-1}$ although $c_{2K} = \sigma_2^{-1}$.

[10]See Duffy et al. (2004) and Hornstein et al. (2005).

where $u(\equiv L_1/L)$is a fraction of skilled labor supply in total labor supply L. From (3.3), the output per capital $y(\equiv Y/K)$ is then described as

$$y = Bf[ucx/B, (1-u)x/B] \tag{3.5}$$

where $c(\equiv A_1/A_2)$ is the skilled/unskilled labor efficiency ratio and $x(\equiv A_2L/K)$is the effective labor capital ratio evaluated by unskilled labor efficiency. Over time, the movement of induced skill-biased innovation represents the dynamics of c and capital accumulation represents the dynamics of x. Later, we examine the dynamics of these two variables.

In the long-run economy, the skilled wage w_1 and the unskilled wage w_2 are competitively determined as their own marginal products. In this setting, the skilled wage rate w_1 and unskilled wage rate w_2 are given by $w_1 = \partial F/\partial L_1 = A_1 f_1$, $w_2 = \partial F/\partial L_2 = A_2 f_2$, respectively. Therefore, wage inequality $\omega(\equiv w_1/w_2)$ and labor income inequality $a/b(\equiv w_1L_1/w_2L_2)$ implying the labor share ratio are described respectively as follows.

$$\omega = f_1c/f_2, a/b = w_1L_1/w_2L_2 = f_1c/f_2\{u/(1-u)\} \tag{3.6}$$

where a is the skilled labor share and b is the unskilled labor share. We note that the movements in wage inequality and income inequality synchronize as long as the fraction of labor is constant, in other words, the labor mobility across sectors does not occur.

3.2.2 *Induced Innovation Frontier*

Consider the induced bias technologies in line with Kennedy (1964), Samuelson (1966) type. We focus on two augmenting technologies, namely skill-augmenting technology $\alpha(\equiv \dot{A}_1/A_1)$, and unskilled-augmenting technology $\beta(\equiv \dot{A}_2/A_2)$ where the dot denotes dx/dt. Their rates of technical changes are given by the following innovation possibility frontier $q(\alpha, \beta, g) = 0$, which is rewritten as

$$\beta = \beta(\alpha, g), \qquad \beta_\alpha < 0, \beta_{\alpha\alpha} < 0,\ \beta_g > 0,\ \beta_{\alpha g} > 0 \tag{3.7}$$

Here, $\beta_\alpha < 0, \beta_{\alpha\alpha} < 0$ exhibit the concavity of the innovation frontier that implies the resource constraints devoted to these factor bias technologies. Moreover, $g(\equiv \dot{K}/K)$ expresses capital accumulation and $\beta_g > 0$ represents a shift of the innovation possibility frontier.[11] Following Adachi et al. (2019), we assume that this comes from the R & D activity embodied in the newly capital stock, which can pro-

[11] As noted before, Samuelson (1966) suggested that the capital share of income is one of the shift parameter of the innovation possibility frontier that has a tradeoff between capital-augmenting and labor -augmenting technical progress.

duce more possible biased technologies. We formulate this as the external effect of capital accumulation on the innovation frontier. Thus, $\beta_g > 0$ expresses the expansion of the innovation frontier. Furthermore, we assume $\beta_{\alpha g} > 0$, implying that the equilibrium skill-augmenting technology is an increasing function of capital accumulation. Then, representative firms facing the innovation frontier are to maximize the instantaneous cost reduction rate of change $a\alpha + b\beta(\alpha, g)$ with respect to α, where a is the skilled labor share and b is the unskilled labor share. Solving this maximization problem yields $-\beta_\alpha(\alpha, g) = a/b$. This means that the tangency of the innovation possibility curve equivalent to the skilled/unskilled labor share ratio, implying that income inequality determines the equilibrium skill-biased innovation. This is explicitly shown as

$$-\beta_\alpha(\alpha, g) = \frac{f_1[ucx/B, (1-u)x/B]cu}{f_2[ucx/B, (1-u)x/B](1-u)} \tag{3.8}$$

where

$$a = w_1 L_1/Y = f_1 cux/Bf = \frac{f_1[ucx/B, (1-u)c/B]ucx/B}{f[ucx/B, (1-u)x/B]} \tag{3.9}$$

$$b = w_2 L_2/Y = f_2(1-u)x/Bf = \frac{f_2[ucx/B, (1-u)x/B](1-u)x/B}{f[ucx/B, (1-u)x/B]} \tag{3.10}$$

Equation (3.8) is solved for each equilibrium biased innovation $\alpha*$, and $\beta*$as follows.

$$\alpha* = \alpha(x, c, g, B, u) \tag{3.11a}$$

$$\beta^* = \beta[\alpha(x, c, g, B, u), g] = \beta(x, c, g, B, u) \tag{3.11b}$$

3.2.3 Dynamics

We consider a standard Solow type neoclassical growth model. Following this model, aggregate savings determine investment, and thus can provide capital accumulation. Then the rate of change of the effective labor capital ratio evaluated at unskilled labor efficiency $\dot{x}/x$, and that of the skilled/unskilled labor efficiency ratio $\dot{c}/c$ are given by as follows.

$$\dot{x}/x = \dot{A}_2/A_2 + \dot{L}/L - \dot{K}/K = \beta + n - sBf[ucx/B, (1-u)x/B] \tag{3.12}$$

$$\dot{c}/c = \dot{A}_1/A_1 - \dot{A}_2/A_2 = \alpha - \beta \tag{3.13}$$

where $\hat{x}(\equiv\dot{x}/x)$denotes the rate of change in time. Here, $s$represents the saving rate assumed to be simply constant, $n$is the rate of change in total labor supply implying population growth. Then, Eqs. (3.12), and (3.13) give the following dynamic system

$$\dot{x}/x = \beta(x, c; g, B, u) + n - sBf[cux/B, (1-u)x/B] \tag{3.14}$$

$$\dot{c}/c = \alpha(x, c; g, B, u) - \beta(x, c; g, B, u) \tag{3.15}$$

$$g = sBf[cux/B, (1-u)x/B] \tag{3.16}$$

The steady state is given by the solution to

$$\beta[\alpha(x*, c*; g*, B, u), g*] + n = sBf[uc * x * /B, (1-u)x * /B] \tag{3.17}$$

$$\alpha(x^*, c^*; g*, B, u) = \beta[\alpha(x^*, c^*; g^*, B, u), g^*] \tag{3.18}$$

$$g^* = sBf[c^*ux^*/B, (1-u)x^*/B] \tag{3.19}$$

We note that in the steady state there are the non-bias innovations, therefore, in our model, the skill-biased innovation appears only in the transitional case. However, even in the steady state, we can also have the skill-biased innovation in the case of the heterogeneity of population growth, particularly when the population growth of unskilled labor is larger than that of skilled labor. We later see this. In the next section, we analyze the properties of the dynamics and the effects on the induced bias and inequalities.

3.3 Analysis

We first consider the stability of the dynamics and then analyze the comparative statics of income inequality, the induced bias innovation and growth in the steady state. Finally, we investigate the implication of the heterogeneity of population growth.

3.3.1 Stability

Before analyzing the dynamic system, we start by examining the properties of the equilibrium biased technologies. In Eq. (18.6), total differentiating with respect to α, x, c, g, and B yields the following equation.

$$\begin{aligned}(\beta_{\alpha\alpha}\alpha/\beta_\alpha)\hat{\alpha} &= (f_{11}l_1/f_1 + f_{12}l_2/f_1 - f_{21}l_1/f_2 - f_{22}l_2/f_2)(\hat{x} - \hat{B}) \\ &\quad + (f_{11}l_1/f_1 - f_{21}l_1/f_1 + 1)\hat{c} - (\beta_{\alpha g}g/\beta_\alpha)\hat{g}\end{aligned} \tag{3.20}$$

where $\hat{x}(\equiv dx/x)$denotes the percentage change in x,and $l_1 \equiv A_1L_1/BK, l_2 \equiv A_2L_2/BK$. Based on our two-level CES production function, this equation is specified as follows:[12]

$$(\beta_{\alpha\alpha}\alpha/\beta_{\alpha})\hat{\alpha} = \frac{\kappa}{1-b}(\sigma_2^{-1} - \sigma_1^{-1})(\hat{x} - \hat{B}) + \frac{1}{1-b}(-a\sigma_2^{-1} - \kappa\sigma_1^{-1} + a + \kappa)\hat{c} - (\beta_{\alpha g}g/\beta_{\alpha})\hat{g} \tag{3.21}$$

where $\kappa(= 1 - a - b)$represents capital share. Therefore, assuming capital-skill complementarity technology ($\sigma_2 > \sigma_1$), we find the effect of these parameters on the equilibrium skill-biased innovation in the elasticity form as follows.

$$\alpha_x x/\alpha = \frac{\beta_\alpha}{\beta_{\alpha\alpha}\alpha}\frac{\kappa}{1-b}(\sigma_2^{-1} - \sigma_1^{-1}) < 0 \tag{3.22a}$$

$$\alpha_c c/\alpha = \frac{\beta_\alpha}{\beta_{\alpha\alpha}\alpha}\frac{1}{1-b}(-a\sigma_2^{-1} - \kappa\sigma_1^{-1} + a + \kappa) \tag{3.22b}$$

$$\alpha_g g/\alpha = \frac{-\beta_{\alpha g}}{\beta_{\alpha\alpha}\alpha}g > 0 \tag{3.22c}$$

$$\alpha_B B/\alpha = \frac{-\beta_\alpha}{\beta_{\alpha\alpha}\alpha}\frac{\kappa}{1-b}(\sigma_2^{-1} - \sigma_1^{-1}) > 0 \tag{3.22d}$$

Similarly, we have the effect of each parameter on the equilibrium unskilled-biased innovation in the elasticity form as follows.

$$\beta_x x/\beta = \beta_\alpha \alpha_x x/\beta > 0 \tag{3.23a}$$

$$\beta_c c/\beta = \beta_\alpha \alpha_c c/\beta \tag{3.23b}$$

$$\tilde{\beta}_g g/\beta = (\beta_\alpha \alpha_g + \beta_g)g/\beta \tag{3.23c}$$

$$\beta_B B/\beta = \beta_\alpha \alpha_B B/\beta < 0 \tag{3.23d}$$

The results of x, c, and B come from the movement on the tradeoff relationship between skill-augmenting and unskilled-augmenting technical progress of the innovation frontier. Thus, we have the converse effects on α and β. However, the results of g come from the shift of the innovation frontier. This provides the possibility of increasing both α and β. Later, we see that these differences provide different outcomes in our analysis of inequality.

Here, an increase in $x$leads to a decrease in α, but an increase in β. However, increases in g and B lead to an increase in α but a decrease in β. A similar result

[12]In our weak separable two-level CES production technology, $f_{ij}l_i/f_j (i, j = 1, 2)$ are specified as follows: $f_{11}l_1/f_1 = -(\kappa\sigma_1^{-1} + ab\sigma_2^{-1})/(1 - b) < 0, f_{12}l_2/f_1 = b\sigma_2^{-1} > 0$, $f_{21}l_1/f_2 = a\sigma_2^{-1} > 0$, and $f_{22}l_2/f_2 = -(1 - b)\sigma_2^{-1} < 0$. Thus, these specifications give Eq. 3.21.

applies to the effect of c although this effect is ambiguous. However, the increase in c can sufficiently lead to a decrease in α and therefore an increase in β if the stability condition in (3.27) is satisfied. We note that this stability condition is corresponding to an elasticity of substitution less than unity in the two-factor case. In the three-factor, two-level CES case, some capital-skill complementarity that are in $\sigma_2 > 1 > \sigma_1$ and $1 > \sigma_2 > \sigma_1$ provide the stability. In this case, the relatively increase in skill-biased technology decreases skill-biased innovation that can produce stability. However, the effect of capital accumulation g produces different outcomes. We have a positive effect of g on both α and β if $\beta_{\alpha g} > 0$and $\tilde{\beta}_g = \beta_\alpha \alpha_g + \beta_g > 0$ are provided. These expansion effects of capital accumulation play significant roles in the comparative statics of inequalities in the steady state.

Now, let us consider the stability condition of the dynamics. Linearizing in Eqs. (3.17), (3.18) that incorporates (3.19) at the steady state and rearranging, we have the following dynamic matrix equation:

$$\begin{pmatrix} \delta(\dot{x}/x)/\delta x \\ \delta(\dot{c}/c)/\delta c \end{pmatrix} = \begin{pmatrix} \frac{\beta_{\alpha\alpha}}{\beta+n}(\alpha_x x/\alpha) + A(a+b) & \frac{\beta_{\alpha\alpha}}{\beta+n}(\alpha_c c/\alpha) + Aa \\ \frac{\alpha}{\alpha+n}(1-\beta_\alpha)(\alpha_x x/\alpha) + B(a+b) & \frac{\alpha}{\alpha+n}(1-\beta_\alpha)(\alpha_c c/\alpha) + Ba \end{pmatrix} \begin{pmatrix} dx/x \\ dc/c \end{pmatrix} \tag{3.24}$$

where

$$A = \beta_\alpha \alpha_g + \beta_g - 1 = \tilde{\beta}_g - 1 \tag{3.25a}$$

$$B = (1-\beta_\alpha)\alpha_g - \beta_g = \alpha_g - \tilde{\beta}_g \tag{3.25b}$$

Describing the matrix of the partial derivatives of the differential equations as J, the stability of the steady state is then locally satisfied when the trace of the matrix J is negative and the determinant of the matrix J is positive. With some calculation, the trace and determinant of the matrix are respectively given by

$$\begin{aligned} trJ &= \frac{\alpha}{\alpha+n}\{\beta_\alpha \alpha_x x/\alpha + (1-\beta_\alpha)\alpha_c c/\alpha\} + A(a+b) + Ba \\ &= -(1-\beta_\alpha-\beta_g)\left[\frac{\alpha}{(\alpha+n)^2}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{1}{b(1-b)}(a\sigma_2^{-1} + b\kappa\sigma_1^{-1} - a - b\kappa) + b\right] < 0 \end{aligned} \tag{3.26}$$

$$\begin{aligned} \det J \equiv \Delta &= \frac{\alpha}{\alpha+n}\{\beta_\alpha B - (1-\beta_\alpha)A\}\{a\alpha_x x/\alpha - (a+b)\alpha_c c/\alpha\} \\ &= \frac{1}{\alpha+n}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}(1-\beta_\alpha-\beta_g)\frac{1}{1-b}(a\sigma_2^{-1} + b\kappa\sigma_1^{-1} - a - b\kappa) > 0 \end{aligned} \tag{3.27}$$

From these two inequalities, we have the following proposition about stability in the steady state.

Proposition 3.1 *With endogenous factor-biased innovation that has the possible shift because of capital accumulation in the three-factor, two-level CES production economy, the steady state is sufficiently stable if the expansion effect of the innovation*

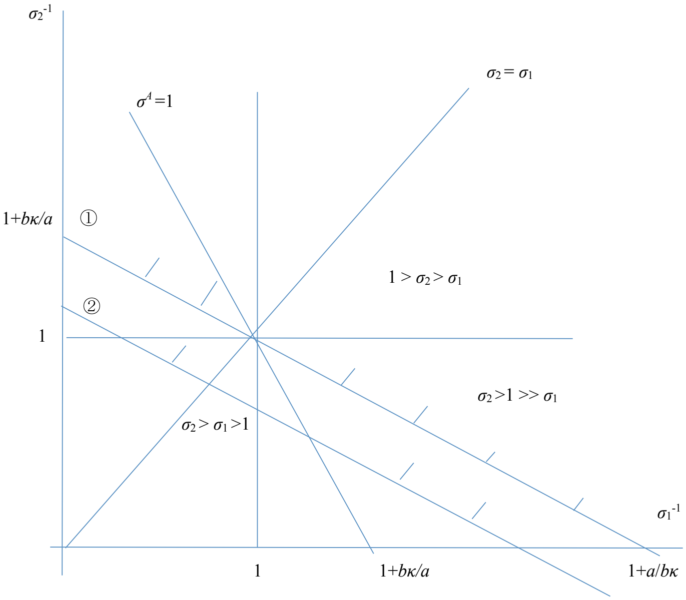


Fig. 3.1 Stability condition ①det $J>0$ if $a\sigma_2^{-1}+b\kappa\sigma_1^{-1}>a+b\kappa$ ②tr$J<0$ $\sigma^A=(a+b\kappa)/(\alpha\sigma_1^{-1}+b\kappa\sigma_2^{-1})$

frontier is weaker and there is some capital-skill complementarity. Specifically, the steady state is sufficiently stable and oscillatory if

$$1-\beta_\alpha-\beta_g>0,\quad a\sigma_2^{-1}+b\kappa\sigma_1^{-1}-a-b\kappa>0. \tag{3.28}$$

Figures 3.1 and 3.2 show the Proposition 3.1 in the case of $1-\beta_\alpha-\beta_g>0$. We have three remarks. First, the two stability conditions imply that dynamic stability is produced by the stable stationary innovation frontier and some factor complementarity production technology. The latter case is corresponding to an elasticity of substitution smaller than unity in the two-factor case. However, the former case is a new one. The implication is that the steady state becomes unstable even in some factor complementarity if $\beta_g$is so large such that the R & D associated with capital accumulation can produce more expanding the frontier. However, if this is not likely to be the case over time, it is plausible to focus on the case of $1-\beta_\alpha-\beta_g>0$. In this case, Fig. 3.1 shows the Proposition 3.1 in two kinds of elasticity of parameters space $\sigma_1^{-1}-\sigma_2^{-1}$.

Second, in the three-factor case, the steady state is stable in capital-skill complementarity $1>\sigma_2>\sigma_1$ and also in some empirically relevant capital-skill complementarity $\sigma_2>1>\sigma_1$. The latter implies that even if there is the elasticity of substitution between skilled and unskilled labor larger than unity, implying $\sigma_2>1$, which

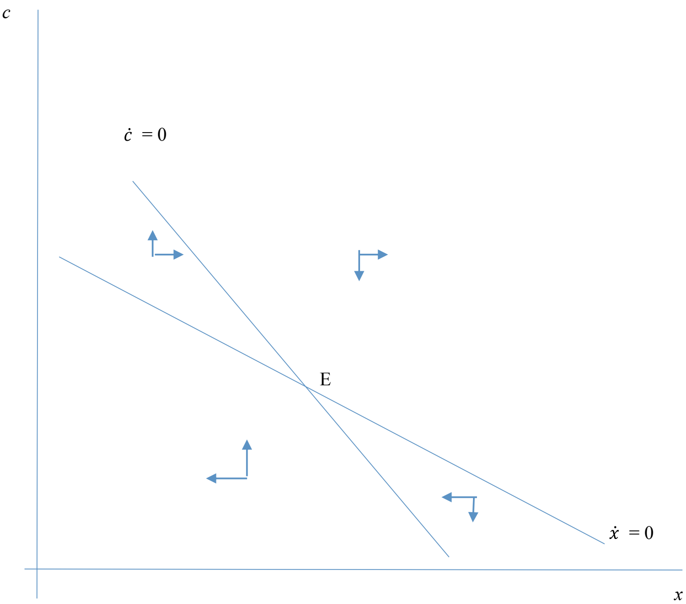


Fig. 3.2 Stability case $a\sigma_2^{-1} + b\kappa\sigma_1^{-1} > a + b\kappa$

is the empirically relevant case,[13] the smaller elasticity of substitution between capital and skilled labor makes the steady state stable. Figure 3.2 shows the stability case in which convergence is oscillatory. However, if $a\sigma_2^{-1} + b\kappa\sigma_1^{-1} - a - b\kappa < 0$, implying some factor substitutability, the steady state is unstable and the saddle point shown in Fig. 3.3. This is just like the elasticity of substitution lager than unity in the two-factor case. Namely, in this case, the skill-biased technology produces more skilled labor income and thus the more skill-biased innovation. Hence, this induced innovation produces more income inequality if $a\sigma_2^{-1} + b\kappa\sigma_1^{-1} - a - b\kappa < 0$, implying the larger substitutability between unskilled labor and capital in even some relevant capital-skill complementarity.

Finally, this stability condition is unrelated to the aggregate elasticity of substitution between capital and overall labor $\sigma^A (= (a + b\kappa)/(a\sigma_1^{-1} + b\kappa\sigma_2^{-1}))$.[14] Indeed, even if the aggregate elasticity of substitution is smaller than unity, instability is likely to be occur.[15] This is because the innovation frontier does not contain capital-augmenting technology and thus the maximization of the instantaneous cost reduction rate of change does not focus on the capital-biased innovation. Thus, if the innovation frontier has capital-augmenting technology and labor-augmenting

[13]See Ciccone and Peri (2005). As noted in footnote 8, if the elasticity of substitution between skilled and unskilled labor is larger than unity, the steady state may always be unstable in the two-factor case.

[14]σ^A is derived in the appendix.

[15]See Fig. 3.1.

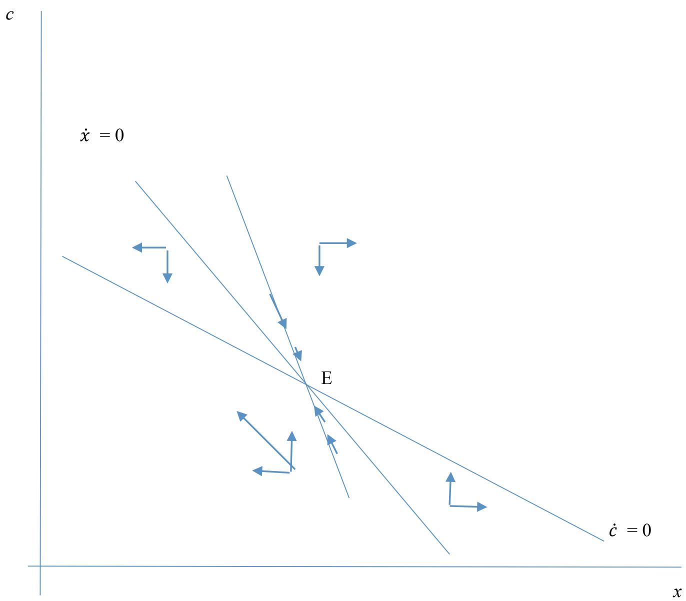


Fig. 3.3 Unstable case $a\sigma_2^{-1} + b\kappa\sigma_1^{-1} < a + b\kappa$

technology in the three-factor case, aggregate elasticity of substitution smaller than unity may play a crucial role in the stability condition.

3.3.2 Comparative Statics

Analyzing the comparative statics in the steady state, we see that the factor substitutability and the expansion effect of the innovation frontier play significant roles for the results of the analysis. First, we consider the case of the effective labor capital ratio and skilled/unskilled efficiency ratio. Total differentiation in the steady state produces the following matrix:

$$\begin{pmatrix} \frac{\beta_\alpha \alpha}{\beta+n}(\alpha_x x/\alpha) + A(a+b) & \frac{\beta_\alpha \alpha}{\beta+n}(\alpha_c c/\alpha) + Aa \\ \frac{\alpha}{\alpha+n}(1-\beta_\alpha)(\alpha_x x/\alpha) + B(a+b) & \frac{\alpha}{\alpha+n}(1-\beta_\alpha)(\alpha_c c/\alpha) + Ba \end{pmatrix} \begin{pmatrix} dx/x \\ dc/c \end{pmatrix}$$
$$= \begin{pmatrix} -A \\ -B \end{pmatrix} ds/s + \begin{pmatrix} -\frac{\beta_\alpha \alpha}{\beta+n}(\alpha_B B/\alpha) - Ac \\ -\frac{\alpha}{\alpha+n}(1-\beta_\alpha)(\alpha_B B/\alpha) - Bc \end{pmatrix} dB/B$$
$$+ \begin{pmatrix} -\frac{n}{\beta+n} \\ 0 \end{pmatrix} dn/n \qquad (3.29)$$

We assume the stability condition $\Delta > 0$ is satisfied. In this case, from this matrix, we obtain the following results for the capital-skill complementarity technology:

$$\hat{x}/\hat{s} = \frac{1}{\Delta}\frac{1}{\alpha+n}(1-\beta_\alpha-\beta_g)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{1}{1-b}(-a\sigma_2^{-1}-\kappa\sigma_1^{-1}+a+\kappa) < 0 \quad (3.30a)$$

$$\hat{x}/\hat{B} = \frac{1}{\Delta}\frac{1}{\alpha+n}(1-\beta_\alpha-\beta_g)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\kappa(1-\sigma_1^{-1}) < 0 \quad (3.30b)$$

$$\hat{c}/\hat{s} = \frac{1}{\Delta}\frac{1}{\alpha+n}(1-\beta_\alpha-\beta_g)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{-\kappa}{1-b}(\sigma_2^{-1}-\sigma_1^{-1}) > 0 \quad (3.30c)$$

$$\hat{c}/\hat{B} = \frac{1}{\Delta}\frac{1}{\alpha+n}(1-\beta_\alpha-\beta_g)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{-\kappa}{1-b}(\sigma_2^{-1}-\sigma_1^{-1}) > 0 \quad (3.30d)$$

$$\hat{x}/\hat{n} = \frac{1}{\Delta}\frac{-n}{\beta+n}\left[\frac{1}{\alpha+n}(1-\beta_\alpha)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{1}{1-b}(-a\sigma_2^{-1}-\kappa\sigma_1^{-1}+a+\kappa)+(\alpha_g-\tilde{\beta}_g)a\right] \quad (3.31)$$

$$\hat{c}/\hat{n} = \frac{1}{\Delta}\frac{n}{\beta+n}\left[\frac{1}{\alpha+n}(1-\beta_\alpha)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{\kappa}{1-b}(\sigma_2^{-1}-\sigma_1^{-1})+(\alpha_g-\tilde{\beta}_g)a\right] \quad (3.32)$$

Thus, both an increase in the saving rate and advancement of capital-augmenting technical progress lead to a decrease in the effective labor capital ratio evaluated at unskilled efficiency, but an increase in the skill/unskilled efficiency ratio. Alternatively, the effects of population growth on these two equilibrium variables are ambiguous. However, if the expansion effect of frontier is small or the effects of capital accumulation on both α and β are almost the same, then the decline of population growth has the same influence on x and c. Specifically, the decline in population growth decreases the effective labor capital ratio evaluated at unskilled efficiency, but increases the skill/unskilled efficiency ratio. Note that skill-biased innovation occurs only in the transition path toward the new steady state. Figure 3.4 illustrates this. Therefore, even in the case of population growth, there is non-biased technical progress in the steady state.

Next, let us consider the effects of the parameters on labor shares, inequality and biased technologies in the steady state. Table 3.1 summarizes the results. Total differentiation with respect to the skilled labor share a, unskilled labor share b, and aggregate labor share $s_L(=a+b)$ in our specified production function provides the following equations:

$$\begin{aligned}\hat{a} &= (f_{11}l_1/f_1+f_{12}l_2/f_1+1-f_1l_1/f-f_2l_2/f)(\hat{x}-\hat{B})+(1+f_{11}l_1/f_1-f_1l_1/f)\hat{c}\\ &= \frac{\kappa}{1-b}(b\sigma_2^{-1}-\sigma_1^{-1}+1-b)(\hat{x}-\hat{B})+\frac{1}{1-b}(-\kappa\sigma_1^{-1}-ab\sigma_2^{-1}+\kappa+ab)\hat{c}\end{aligned} \quad (3.33)$$

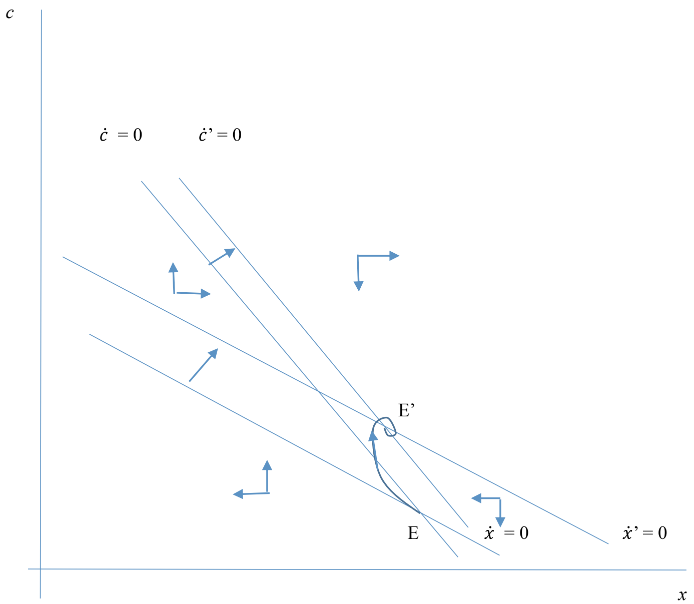


Fig. 3.4 Stable case $(a\sigma_2^{-1} + b\kappa\sigma_1^{-1} < a + b\kappa)$ $s \uparrow, B \uparrow \Rightarrow x \downarrow, c \downarrow \Rightarrow \uparrow$

Table 3.1 Effects of the parameters on labor shares, inequality, and biased innovation

	x	c	a	b	s_L	a/b	α	β	g
s	$-$	$+$	$-$	$-$	$-$	0	0	0	0
B	$-$	$+$	$-$	$-$	$-$	0	0	0	0
n	$+^d$	$-^e$	$+^d$	$+^d$	$+^d$	$-^f$	$+$	$+$	$+$

Note $\sigma_2 > 1 > \sigma_1$, $\Delta = a\sigma_2^{-1} + b\kappa\sigma_1^{-1} - a - b\kappa > 0$
$d > 0$, $e < 0$ if $\alpha_g \cong \tilde{\beta}_g$, $f < 0$ if $\alpha_g > \tilde{\beta}_g$

$$\hat{b} = (f_{21}l_1/f_2 + f_{22}l_2/f_1 + 1 - f_1l_1/f - f_2l_2/f)(\hat{x} - \hat{B}) + (f_{21}l_1/f_2 - f_1l_1/f)\hat{c}$$
$$= \kappa(1 - \sigma_2^{-1})(\hat{x} - \hat{B}) + a(\sigma_2{}^{-1} - 1)\hat{c} \tag{3.34}$$

$$\hat{s}_L = \frac{a}{a+b}\hat{a} + \frac{b}{a+b}\hat{b} \tag{3.35}$$

Taking account of (3.30), (3.31), (3.32) in the above equations, we have the following consequences in some empirically relevant capital-skill complementarity ($\sigma_2 > 1 > \sigma_1$). Concerning the saving rate and capital-augmenting technological progress, we have the following same results. This is because the frontier does not shift in the steady state:

$$\hat{a}/\hat{s} = \hat{b}/\hat{s} = \hat{s}_L/\hat{s} = \frac{1}{\Delta}\frac{1}{\alpha+n}(1-\beta_\alpha-\beta_g)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\kappa(1-\sigma_2^{-1})(1-\sigma_1^{-1}) < 0 \tag{3.36}$$

$$\hat{a}/\hat{B} = \hat{b}/\hat{B} = \hat{s}_L/\hat{B} = \frac{1}{\Delta}\frac{1}{\alpha+n}(1-\beta_\alpha-\beta_g)\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\kappa(1-\sigma_2^{-1})(1-\sigma_1^{-1}) < 0 \tag{3.37}$$

$$\hat{a}/\hat{s} - \hat{b}/\hat{s} = 0, \quad \hat{a}/\hat{B} - \hat{b}/\hat{B} = 0 \tag{3.38}$$

$$\hat{\alpha}/\hat{s} = \hat{\beta}/\hat{s} = 0, \quad \hat{\alpha}/\hat{B} = \hat{\beta}/\hat{B} = 0 \tag{3.39}$$

Namely, both an increase in the saving rates and the advancement of capital-augmenting technical progress decrease in the skilled labor shares, unskilled labor share, and aggregate labor share at the same rates. Thus, the capital shares go up, but labor income inequality a/b does not occur. These results come from the fixity of the innovation possibility frontier in the steady state. Because in the steady state, the non-shifting innovation frontier provides the constancy of the tangency of frontier curve, which is equivalent to the labor income ratio. Hence, the movement of skilled labor shares can synchronize with that of the unskilled labor share and the non-biased innovation occurs in the steady state.

However, the consequences of population growth provide different outcomes. Because the innovation has the possibility of expanding owing to capital accumulation. As a result, the shift in the innovation frontier has a positive influence on growth, and thus this influence has the possible effect on income inequality. The results are given as follows:

$$\hat{a}/\hat{n} = \frac{1}{\Delta}\frac{-n}{\beta+n}\left[\frac{1}{\alpha+n}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}(1-\beta_\alpha)\kappa(1-\sigma_2^{-1})(1-\sigma_1^{-1}) + (\alpha_g-\tilde{\beta}_g)\frac{b}{1-b}(a\sigma_2^{-1}+\kappa\sigma_1^{-1}-a-\kappa)\right] \tag{3.40}$$

$$\hat{b}/\hat{n} = \frac{1}{\Delta}\frac{-n}{\beta+n}(1-\sigma_2^{-1})\left[\frac{1}{\alpha+n}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}(1-\beta_\alpha)\kappa(1-\sigma_1^{-1}) + (\alpha_g-\tilde{\beta}_g)a\right] \tag{3.41}$$

$$\hat{s}_L/\hat{n} = \frac{1}{\Delta}\frac{-n}{\beta+n}\kappa\left[\frac{1}{\alpha+n}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}(1-\beta_\alpha)(1-\sigma_2^{-1})(1-\sigma_1^{-1}) + (\alpha_g-\tilde{\beta}_g)\frac{b}{1-b}\frac{a}{1-\kappa}(\sigma_1^{-1}-\sigma_2^{-1})\right] \tag{3.42}$$

$$\hat{a}/\hat{n} - \hat{b}/\hat{n} = \frac{-1}{1-\beta_\alpha-\beta_g}\frac{\beta_{\alpha\alpha}}{\beta_\alpha}(\alpha_g-\tilde{\beta}_g)n \tag{3.43}$$

$$\hat{\alpha}/\hat{n} = \hat{\beta}/\hat{n} = \frac{\beta_g}{1-\beta_\alpha-\beta_g}\frac{n}{\beta} > 0 \tag{3.44}$$

Specifically, in some relevant capital-skill complementarity, although the effects of the population growth on each labor share become ambiguous because of the expansion effect, labor income inequality declines if its influence on skill-biased innovation is greater than that on unskilled-biased innovation. Moreover, if the influence of the expansion frontier on each biased innovation is almost the same, the population growth leads to decreases in the skilled labor share, unskilled labor share, and aggregate labor share. Furthermore, we find a positive influence on both skill-biased and unskilled-biased innovations at the same rate, and thus, the positive effect on growth. Conversely, in this case, a decline of the population provides the opposite outcomes. Summarizing the outcomes, we have the following proposition.

Proposition 3.2 *With endogenous factor-biased innovation that has the possible shift because of capital accumulation in the three-factor case, with some relevant capital-skill complementarity technology, the following holds:*

1. *Because in the stable steady state, the change in the saving rate and advancement of capital-augmenting technology cannot produce a shift in the innovation possibility frontier, both skill-biased and unskilled-biased innovation do not change. Thus, labor income inequality does not occur, although the increase in the saving rate and advancement of capital-augmenting technical progress can lead to decreases in the skilled labor share, unskilled labor share, and aggregate labor share at the same rates.*
2. *However, even in the stable steady state, because population growth can have the possible shift of the innovation frontier, both skill-biased and unskilled-biased innovations can proportionally change. Therefore, the decline of population growth is likely to move the frontier inward, leading to decreases in the skilled labor share, unskilled labor share, and aggregate labor shares if the expansion effect of both skill-biased and unskilled-biased innovation is weaker. Furthermore, an increase in income inequality can occur if the expansion effect of skill-biased innovation is larger than that of unskilled-biased innovation.*

In the stability case of the steady state, our outcomes for labor shares in the steady state depend crucially on the property of some substitutability between capital and unskilled labor, in other words, some substitutability between skilled and unskilled labor among some relevant capital-skill complementarity technology $\sigma_2 > 1 > \sigma_1$. Conversely, even in the steady state, there is some complementarity between capital and unskilled labor among capital-skill complementarity $1 > \sigma_2 > \sigma_1$, and thus some outcomes can be reversed.

3.3.3 Implications of the Heterogeneity of Population Growth

Finally, we consider the case of the heterogeneity of population growth. One of the weakness of our previous analysis is that the steady state has no biased innovation, implying skill-biased innovation equivalent to unskilled-biased innovation. However, even in the steady state in our framework, the equilibrium skill-biased innovation can be introduced if the population growth of unskilled labor is larger than that of skilled labor. This is likely to occur if the newly arrived technologies provide the obsolescence of many types of skills implying relatively increasing number of the unskilled labor, in other words, the more scarcity of skilled labor in the growth process.

To analyze this case, we have some modifications of our framework. Let n_1 and n_2 denote the population growth of skilled labor and unskilled labor, respectively, and we assume that the population growth of unskilled labor is larger than that of skilled labor $n_2 > n_1$. In this setting, rewriting output per capital leads to

$$y = Bf(\tilde{c}\tilde{x}/B, \tilde{x}/B) \tag{3.45}$$

where $\tilde{c}(\equiv A_1L_1/A_2L_2)$ is the effective skilled /unskilled labor ratio and $\tilde{x}(\equiv A_2L_2/K)$ is the effective unskilled labor capital ratio. Then, the dynamics of $\tilde{x}$ and $\tilde{c}$ are given as follows:

$$\dot{\tilde{x}}/\tilde{x} = \beta + n_2 - sBf(\tilde{c}\tilde{x}/B, \tilde{x}/B), \quad \dot{\tilde{c}}/\tilde{c} = \alpha - \beta + n_1 - n_2 \tag{3.46}$$

Since the formulation of factor-biased innovation is the same, the solution to the maximization of the instantaneous cost reduction rate of change on the concavity of the innovation possibility frontier yields $-\beta_\alpha(\alpha, g) = a/b$ which is shown as $-\beta_\alpha(\alpha, g) = \frac{f_1(\tilde{c}\tilde{x}/B, \tilde{x}/B)\tilde{c}}{f_2(\tilde{c}\tilde{x}/B, \tilde{x}/B)}$. Thus, we have $\alpha* = \alpha(\tilde{x}, \tilde{c}, g, B), \beta* = \beta[\alpha(\tilde{x}, \tilde{c}, g, B), g] = \beta(\tilde{x}, \tilde{c}, g, B)$. Therefore, the dynamics are explicitly given by

$$\dot{\tilde{x}}/\tilde{x} = \beta(\tilde{x}, \tilde{c}, g, B) + n_2 - sBf(\tilde{c}\tilde{x}/B, \tilde{x}/B) \tag{3.47a}$$

$$\dot{\tilde{c}}/\tilde{c} = \alpha(\tilde{x}, \tilde{c}, g, B) - \beta(\tilde{x}, \tilde{c}, g, B) + n_1 - n_2 \tag{3.47b}$$

$$g = sBf(\tilde{c}\tilde{x}/B, \tilde{x}/B) \tag{3.47c}$$

Hence, the steady state is given by the solution to

$$\beta(\tilde{x}*, \tilde{c}*, g*, B) + n_2 = sBf[\tilde{c} * \tilde{x} * /B, \tilde{x} * /B] \tag{3.48}$$

$$\alpha(\tilde{x}*, \tilde{c}*, g*, B) + n_1 = \beta(\tilde{x}*, \tilde{c}*, g*, B) + n_2 \tag{3.49}$$

$$g* = sBf(\tilde{c} * \tilde{x} * /B, \tilde{x} * /B). \tag{3.50}$$

Figure 3.5 illustrates this steady state. Obviously, as far as $n_2 > n_1$, we have $\alpha(\tilde{x}*, \tilde{c}*, g*, B) > \beta(\tilde{x}*, \tilde{c}*, g*, B)$. Namely, in this case, more skill-biased innovation is introduced. Because the scarcity of skilled labor induces firms to promote more skill-augmenting technology. Moreover, purely skill-biased innovation can appear if there is large population growth of unskilled labor such that the upward line satisfying $\alpha + n_1 = \beta + n_2$ intersects with point α_0 on the innovation frontier in Fig. 3.5. In addition, we can see the consequences of the population growth in unskilled labor

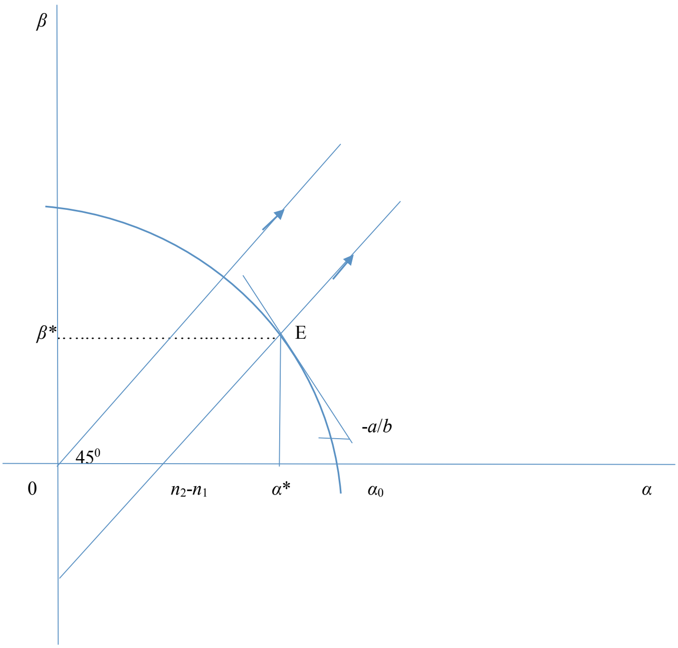


Fig. 3.5 Skill-biased innovation $\alpha^* > \beta^*(n_2 > n_1)$

in the steady state. Calculation of the comparative statics provides the following results[16]

$$\hat{\hat{x}}/\hat{n}_2 = \frac{1}{\Delta}\frac{n_2}{\beta+n_2}\left[\frac{-1}{\alpha+n_1}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{1}{1-b}(-a\sigma_2^{-1}-\kappa\sigma_1^{-1}+a+\kappa)+(1-\alpha_g)a\right] \tag{3.51}$$

$$\hat{\hat{c}}/\hat{n}_2 = \frac{1}{\Delta}\frac{n_2}{\beta+n_2}\left[\frac{1}{\alpha+n_1}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\frac{\kappa}{1-b}(\sigma_2^{-1}-\sigma_1^{-1})-(1-\alpha_g)a\right] \tag{3.52}$$

$$\hat{a}/\hat{n}_2 = \frac{1}{\Delta}\frac{n_2}{\beta+n_2}\left[\frac{-1}{\alpha+n_1}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\kappa(1-\sigma_2^{-1})(1-\sigma_1^{-1})+(1-\alpha_g)\frac{b}{1-b}(a\sigma_2^{-1}+\kappa\sigma_1^{-1}-a-\kappa)\right] \tag{3.53}$$

$$\hat{b}/\hat{n}_2 = \frac{1}{\Delta}\frac{n_2}{\beta+n_2}(1-\sigma_2^{-1})\left[\frac{-1}{\alpha+n_1}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}\kappa(1-\sigma_1^{-1})+(1-\alpha_g)a\right] \tag{3.54}$$

$$\hat{s}_L/\hat{n}_2 = \frac{1}{\Delta}\frac{n_2}{\beta+n_2}\kappa\left[\frac{-1}{\alpha+n_1}\frac{\beta_\alpha}{\beta_{\alpha\alpha}}(1-\sigma_2^{-1})(1-\sigma_1^{-1})+(1-\alpha_g)\frac{b}{1-b}\frac{a}{1-\kappa}(\sigma_1^{-1}-\sigma_2^{-1})\right] \tag{3.55}$$

$$\hat{a}/\hat{n}_2 - \hat{b}/\hat{n}_2 = \frac{1}{1-\beta_\alpha-\beta_g}\frac{\beta_{\alpha\alpha}}{\beta_\alpha}(1-\alpha_g)n_2 \tag{3.56}$$

[16]The details are available on request.

$$\hat{\alpha}/\hat{n}_2 = \frac{1}{1-\beta_\alpha-\beta_g}\frac{n_2}{\alpha} > 0. \tag{3.57}$$

$$\hat{\beta}/\hat{n}_2 = \frac{\beta_\alpha+\beta_g}{1-\beta_\alpha-\beta_g}\frac{n_2}{\beta} \tag{3.58}$$

$$\hat{\alpha}/\hat{n}_2 - \hat{\beta}/\hat{n}_2 = \frac{n_2}{\alpha\beta(1-\beta_\alpha-\beta_g)}(\beta-\beta_\alpha\alpha-\beta_g\alpha) \tag{3.59}$$

From these, we have the following proposition 3.3.

Proposition 3.3 *With endogenous factor-biased innovation that has the possible shift because of capital accumulation in the three-factor case, with some relevant capital-skill complementarity technology, as well as relatively larger population growth in unskilled labor, the following holds:*

If the population growth of unskilled labor is larger than that of skilled labor, skill-biased innovation can be introduced in the steady state. In the stable steady state, the increase in the population growth of unskilled labor can lead to expanding the frontier, resulting in more growth and skill-biased innovation if the expansion effect of unskilled-biased innovation is small. However, larger population growth of unskilled labor can lead to more increases in the skilled labor share than in the unskilled labor share, and thus can result in more labor income inequality as far as the expansion effect of skill-biased innovation is not so large.

Note that we obtain our results when the population of skilled labor is relatively scarce and the innovation possibility frontier is rather stationary. Therefore, if skilled labor supply is increasing, and the expansion effect of the innovation frontier is not so small, some outcomes may be reversed. However, if capital-augmenting technology such as Automation and Artificial Intelligence technologies can play the same roles for skilled workers and influence the expansion of the innovation possibility frontier and also change the capital-skill complementarity structure, the consequences of our analysis may lead to different outcomes.[17] Then, an alternative framework such as formulating endogenous capital-augmenting technological progress may be needed.

3.4 Concluding Remarks

Based on a three-factor, two-level CES production function that has capital-skill complementarity, we developed an endogenous skill-biased innovation model in the neoclassical growth model to analyze the dynamics of income inequalities. Extending the innovation possibility frontier that has a skill-biased innovation type, we showed that some relevant capital-skill complementarity technology and the possible shift of the innovation frontier because of capital accumulation play significant roles in

[17] See Acemoglu and Restrepo (2018), Berg et al. (2018).

the stability of the steady state and the behaviors of labor shares and inequalities. We also showed that the relatively large population growth of unskilled labor can introduce more skill-biased innovation and thus increasing income inequality. However, to investigate the implications of Automation and AI on growth, inequality and unemployment, an alternative formulation of endogenous capital-augmenting technical progress and its analysis are needed. These issues will be dealt with future research.

Acknowledgements The author is grateful to Hideyuki Adachi, Professor Emeritus of Kobe University. He is also grateful to Tamotsu Nakamura, Takeshi Nakatani, Atsushi Miyake, and Hideaki Uchida for their valuable comments on an earlier version of this paper.

Appendix

Aggregate Elasticity of Substitution

We can derive the endogenous elasticity of substitution as follows. $w/r$is given by the following equation:

$$\begin{aligned} w/r &= \{uw_1 + (1-u)w_2\}/\{f(l_1, l_2) - f_1(l_1, l_2)l_1 - f_2(l_1, l_2)l_2\} \\ &= v(l_1, l_2, u) \end{aligned} \tag{3.60}$$

In this case, $l_1 = u/k$ and $l_2 = (1-u)$ where $k \equiv K/L$. Denoting $\theta_{l1k} \equiv (k/l_1)(dl_1/dk)$, $\theta_{l2k} \equiv (k/l_2)(dl_2/dk)$, we have the following equations:

$$\theta_{l1k} = -1, \theta_{l2k} = -1 \tag{3.61}$$

Moreover, defining $\eta_1 \equiv (l_1/v)\partial v/\partial l_1$ and $\eta_2 \equiv (l_2/v)\partial v/\partial l_2$, the aggregate elasticity of substitution σ^A is then written as

$$\begin{aligned} \sigma^A &\equiv [(w/r)/k]dk/d(w/r) \\ &= \{(k/v)dv/dk\}^{-1} \\ &= (\eta_1\theta_{l1k} + \eta_2\theta_{l2k})^{-1} \\ &= (-\eta_1 - \eta_2)^{-1} \end{aligned} \tag{3.62}$$

Calculating each $\eta_1 \equiv (l_1/v)\partial v/\partial l_1$, $\eta_2 \equiv (l_2/v)\partial v/\partial l_2$, and specifying these in our two-level CES production function yield

$$\eta_1 = (b\sigma_2^{-1} - \sigma_1^{-1})a/(a + b\kappa), \eta_2 = -\sigma_2^{-1}b/(a + b) < 0 \tag{3.63}$$

Thus, the aggregate elasticity of substitution is given by

$$\begin{aligned}\sigma^A &= (-\eta_1 - \eta_2)^{-1} \\ &= (a + b\kappa)/(a\sigma_1^{-1} + b\kappa\sigma_2^{-1})\end{aligned} \qquad (3.64)$$

References

Acemoglu, D. 2002. Technical change, inequality, and the labor market. *Journal of Economic Literature* 40: 7–72.

Acemoglu, D. 2010. When does labor scarcity encourage innovation? *Journal of Political Economy* 118: 1037–1078.

Acemoglu, D. 2015. Localised and biased technologies: Atkinson and Stiglitz's new view innovation, and directed technological change. *Economic Journal* 125: 443–463.

Acemoglu, D., and P. Restrepo. 2017a. The Race between man and machine: Implications of technology for growth, factor shares and employment. *American Economic Review* (forthcoming).

Acemoglu, D., and P. Restrepo. 2017b. Robots and jobs: Evidence from US labor markets. NBER Working Papers, 23285.

Acemoglu, D., and P. Restrepo. 2018. Modeling automation. *American Economic Review*, *Paper and Proceedings*.

Adachi, H., T. Nakamura, K. Inagaki, and Y. Osumi. 2019. *Technological progress, income distribution, and unemployment -Theory and empirics-*, Springer Briefs in Economics.

Berg, A., E. Buffie, and F. Zanna. 2018. Should we fear the robot revolution? (The correct answer is yes), IMF Working Paper, 18/116.

Caselli, F., and W.J. Coleman II. 2006. The world technology frontier. *American Economic Review* 96: 499–523.

Ciccone, A., and G. Peri. 2005. Long-run substitutability between more and less educated workers: Evidence from a panel of countries. *Review of Economics and Statistics* 87: 652–663.

Drandakis, E.M., and E.S. Phelps. 1966. A model of induced invention, growth, and distribution. *Economic Journal* 76: 823–840.

Duffy, J., C. Papageorgiou, and F. Perez-Sebastian. 2004. Capital-skill complementarity? Evidence from a panel of countries. *Review of Economics and Statistics* 86: 327–344.

Graetz, G., and G. Michaels. 2015. Robots at work. CEPR Discussion Paper 1335.

Griliches, Z. 1969. Capital-skill complementarity. *Review of Economics and Statistics* 51: 465–468.

Hornstein, A., P. Krusell, and G. L. Violante. 2005. The effects of technical change on labor market inequalities. In *Handbook of economic growth*, vol. 1B, eds. Aghion, P. and S.N. Durlauf, 1275-1370, North-Holland.

Kennedy, C. 1964. Induced bias in innovation and the theory of distribution. *Economic Journal* 74: 541–547.

Kotlikoff, L., and J.D. Sachs. 2012. Smart machines and long-term misery. NBER Working Paper, 18629.

Korinek, A., and Stiglitz, J.E. 2017. Artificial intelligence and implications for income distribution and unemployment. NBER Working Paper 24174.

Krusell, P., L.E. Ohanian, J.-V. Rios-Rull, and G.L. Violante. 2000. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica* 68: 1029–1053.

Nordhaus, W.D. 1973. Some skeptical thoughts on the theory of induced innovation. *Quarterly Journal of Economics* 87: 208–219.

Samuelson, P. 1966. A theory of induced innovation along Kennedy-Weizacker lines. *Review of Economics and Statistics* 33: 133–146.

Sato, K. 1967. A two-level constant-elasticity-of-substitution production function. *Review of Economic Studies* 34: 201–218.

Solow, R.M. 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics* 43: 65–94.

Stiglitz, J.E. 2014. Unemployment and innovation. NBER Working Paper 20670.

Chapter 4
Ramsey's Conjecture in a Stochastically Growing Economy

Tamotsu Nakamura

4.1 Introduction

Becker (1980) confirms Ramsey's conjecture (Ramsey 1928) in a discrete-time model by proving that the most patient individuals own the entire capital of the economy, while others consume their wage income in the long-run. Using a continuous-time version of Becker's model, Mitra and Sorger (2013) not only confirm the conjecture, but also show that the exclusive possession by the most patient takes place after some finite time. Including these two contributions, Ramsey's conjecture has been examined in various environments. As a result, it turns out that the conjecture holds in the standard deterministic neoclassical growth model under moderate conditions as long as the economy reaches the stationary state without growth.

Using an endogenous growth model of learning-by-doing, Boyd III (2000) points out the importance of intertemporal substitution in the long-run income distribution and shows that Ramsey's conjecture does not always hold. Using a simple AK-type endogenous growth model, Nakamura (2014) also analyzes the role of intertemporal substitution to show the possibility that the most impatient individuals finally come to own almost all (but not the entire) capital in the long-run, if they have the highest intertemporal elasticity of substitution. Also, as Becker and Zilcha (1997) explain it by analyzing a stochastic version of Becker (1980) with borrowing constraints, uncertainty can be another cause of the failure of Ramsey's conjecture.

Taking uncertainty into account, risk aversion becomes to play an influential role in the preferences and hence in the long-run wealth distribution. To avoid future negative shocks to income, an individual will accumulate assets beyond her needs. Together with uncertainties, risk aversion determines this part of saving, which is referred to as *precautionary saving*. While an inverse of intertemporal elasticity of substitution measures the degree of consumption smoothing, risk aversion reflects

T. Nakamura (✉)
Kobe University, Kobe, Japan
e-mail: nakamura@econ.kobe-u.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_4

the human tendency to prefer certain gains to uncertain ones. Although the two concepts are different, traditional expected utility maximization models do not allow risk aversion to change independently from intertemporal substitution. In the case of the constant relative risk aversion utility function such as in Nakamura (2014), an inverse of intertemporal elasticity of substitution is equal to a degree of relative risk aversion.

Under uncertainty the subjective rate of return on capital that a risk-averse individual perceives is lower than the mathematical expectation. The risk-adjusted rate decreases as uncertainty or risk aversion increases. Hence, risk aversion and uncertainty increase the capital accumulation directly through an increase in *precautionary saving*, while they reduce the accumulation indirectly through a decline in the risk-adjusted rate of return. Thus, not only intertemporal substitution but risk aversion also plays an important role in the determination of saving and hence the capital accumulation under uncertainty via different channels.

To clarify the roles of the three different concepts, we adopt Kreps-Porteus non-expected utility preferences (Kreps and Porteus 1978, 1979), which allow independent variation in risk aversion and intertemporal substitution. Simplifying the production technology to a simple AK production function, this paper examines Ramsey's conjecture in a stochastically growing economy. As a result, it is shown that whether the intertemporal elasticity of substitution is higher or lower than one is crucial in the stochastic environment. If impatient individuals are sufficiently *less* risk-averse than the most patient, then the conjecture fails when the intertemporal elasticity is *higher* than one. Conversely, if impatient individuals are sufficiently *more* risk-averse than the most patient, then Ramsey's conjecture fails when the intertemporal elasticity of substitution is *lower* than one.

The rest of the chapter is organized as follows. Section 4.2 sets up a model with two types of heterogeneous individuals. Section 4.3 analyzes the growth, the dynamics and long-run wealth distribution in the economy. Section 4.4 concludes the chapter.

4.2 The Model

Consider the competitive equilibrium of an infinite horizon production economy that consists of a continuum of infinitely-lived heterogeneous individuals: each owns a firm and produces a homogenous good according to a stochastic technology.[1] The population of individuals is constant over time and normalized to one. They are uniformly distributed on a unit interval [0, 1]. For the sake of simplicity, it is assumed that only two types of individuals, type H and type L, exist in the economy.[2]

[1]This is to simplify the analysis of aggregate behavior. Since all individuals face the uncertainties having the same properties, the individual behavior does not change if we assume aggregate shocks instead of idiosyncratic shocks.

[2]The main results obtained in what follows remain unchanged for the case with an arbitrary number of types of individuals.

Individual i is type H if $i \in [0, \lambda)$, while individual j is type Lif $j \in [\lambda, 1]$. Hence, λ of the total population are type H, while $1-\lambda$ are type L.

4.2.1 Technology

To make the point clear, let us assume a simple AK technology. Output $y_i(t)$ is produced only by capital $k_i(t)$ according to the following stochastic production function:

$$dy_i(t) = Ak_i(t)[dt + \sigma dz_i(t)] \text{ with } A > 0 \text{ and } \sigma > 0 \tag{4.1}$$

In each period, the deterministic flow of each firm's production is $Ak_i(t)dt$. In addition to this deterministic part, the stochastic component of production $Ak_i(t)\sigma dz_i(t)$ also exists due to idiosyncratic technology shocks, where $dz_i(t)$ is a Wiener process with mean zero and unit variance, and the parameter σ is the instantaneous standard deviation of the shock. Assuming that capital does not depreciate, the rate of return on capital $r_i(t)$ becomes as follows:

$$r_i(t) = A[dt + \sigma dz_i(t)]. \tag{4.2}$$

Suppose that the shocks are uninsurable and idiosyncratic, and that the population is large enough to apply the law of large numbers. Then there is no aggregate uncertainty because idiosyncratic shocks are cancelled out at the aggregate level. Hence, the followings hold:

$$\int_0^\lambda k_i(t)dz_i(t)di = \int_\lambda^1 k_i(t)dz_i(t)di = 0, \text{ and hence, } \int_0^1 k_i(t)dz_i(t)di = 0 \; for all \; t \tag{4.3}$$

4.2.2 Preferences and Budget Constraints

Each individual's utility depends on consumption. To distinguish the effect of intertemporal substitution from that of risk-aversion, we employ a non-expected utility maximization setup.[3] At point in time t, individual i maximizes the intertemporal objective $V_i(t)$ by recursion,

$$f([1-\gamma_i]V_i(t)) = \left(\frac{1-\gamma_i}{1-1/\varepsilon_i}\right) c_i(t)^{1-1/\varepsilon_i} h + e^{-\rho_i h} f([1-\gamma_i]E_t V_i(t+h)), \tag{4.4}$$

[3]For detailed treatment of "ecursive utility, see Duffie and Esptein (1992).

where the function $f(x)$ is given by

$$f(x_i) = \left(\frac{1-\gamma_i}{1-1/\varepsilon_i}\right) x_i^{(1-1/\varepsilon_i)/(1-\gamma_i)}. \tag{4.5}$$

In (4.4), h is the economic decision interval, E_t is a mathematical expectation conditional on time-t information, and $\rho_i > 0$ is the time preference rate. The parameter $\gamma_i > 0$ measures the relative risk-aversion, while the parameter $\varepsilon_i > 0$ is the intertemporal elasticity of substitution.[4] When $\gamma_i = 1/\varepsilon_i$, so that $f(x_i) = x_i$, the setup becomes the standard state- and time-separable.[5]

The flow budget constraint of individual i is given by

$$c_i(t)dt + dk_i(t) = r_i(t)k_i(t)dt \text{ with } k_i(0) > 0, \tag{4.6}$$

where $k_i(0)$ is the initial capital of individual i. Taking (4.2) into account, (4.6) is rewritten as

$$dk_i(t) = [Ak_i(t) - c_i(t)]dt + \sigma A dz_i(t)k_i(t) \text{ with } k_i(0) > 0. \tag{4.7}$$

Without a borrowing constraint or a debt constraint, the maximization problem is trivial because individuals can increase their utility as much as they want by accumulating infinite debt. We impose the following borrowing constraint, known as the no-Ponzi game condition:

$$k_i(t) \geq 0 \text{ for any } t. \tag{4.8}$$

Since an individual earns no wage income, once she has debt, i.e., $k_i(t)$ becomes negative, she is unable to repay it. In other words, she cannot have debt at any point in time.

Since the individual's instantaneous felicity function satisfies the Inada condition, the consumption at any point in time must be positive on any equilibrium path, i.e.,

$$c_i(t) > 0 \text{ for any } t. \tag{4.9}$$

When the individual's capital becomes zero, i.e., $k_i(t) = 0$, then she cannot have positive consumption because she cannot borrow in equilibrium. Taking the inequality (4.9) into account, (4.8) must hold with strict inequality as follows[6]:

$$k_i(t) > 0 \text{ for any } t. \tag{4.10}$$

---

[4] For a detailed discussion on the roles of these parameters, see, for example, Kreps and Porteus (1978, 1979), Epstein and Zin (1989, 1991), Weil (1989), and Obstfeld (1994a, b).

[5] This paper analyzes the individual's behavior in the limit as h becomes infinitesimally small. When $\gamma = 1/\varepsilon$, then (4.4) implies that, as $h \to 0$, $V(t)$ becomes the standard setup of discounted sum of utilities, i.e., $V_i(t) = E_t\left[(1-\gamma_i)^{-1}\int_t^\infty c_i(t)^{1-\gamma_i} e^{-\rho_i(s-t)} ds\right]$.

[6] The inequalities (4.9) and (4.10) constitute the feasibility condition, and, in turn, impose the parameter restrictions. See also footnote 7.

4.2.3 Individual Optimization and Saving Behavior

Let $J_i(k_i(t))$ denote the maximum feasible level of the expected sum of discounted utilities. Equation (4.4) suggests that the value function $J_i(k_i(t))$ takes the following form:

$$J_i(k_i(t)) = (b_i k_i)^{1-\gamma_i}/(1-\gamma_i), \tag{4.11}$$

with the consumption given by

$$c_i(t) = u_i k_i(t) \quad \text{with } u_i = b_i^{1-\varepsilon_i}, \tag{4.12}$$

where b_i is a positive constant to be determined. Solving the maximization problem (see Appendix for the detailed derivation), we obtain

$$u_i = A - \varepsilon_i(A-\rho_i) - \frac{(1-\varepsilon_i)\gamma_i(\sigma A)^2}{2}, \tag{4.13}$$

and,

$$b_i = \left[A - \varepsilon_i(A-\rho_i) - \frac{(1-\varepsilon_i)\gamma_i(\sigma A)^2}{2}\right]^{1/(1-\varepsilon_i)}. \tag{4.14}$$

Substitution of (4.13) into (4.12) gives the following consumption function[7]:

$$c_i = u_i k_i = \left[A - \varepsilon_i(A_i-\rho_i) - \frac{(1-\varepsilon_i)\gamma_i(\sigma A)^2}{2}\right] k_i. \tag{4.15}$$

(Time arguments are suppressed when no ambiguity results.)

Substituting (4.15) into (4.7), we have the individual capital accumulation equation as follows:

$$dk_i = s_i k_i dt + \sigma A k_i dz_i, \tag{4.16}$$

where

$$s_i = \varepsilon_i(A-\rho_i) + \frac{(1-\varepsilon_i)\gamma_i(\sigma A)^2}{2} = \varepsilon_i(\tilde{r}_i - \rho_i) + \frac{\gamma_i(\sigma A)^2}{2} \text{with } \tilde{r}_i = A - \frac{\gamma_i(\sigma A)^2}{2}, \tag{4.17}$$

where $\tilde{r}_i$ is the risk-adjusted rate of return on capital for individual i. As (4.16) shows, the capital accumulation dk_i consists of two parts; one is the deterministic part $s_i k_i dt$, and the other is the stochastic part, i.e., unanticipated gains or losses $\sigma_i A k_i dz_i$.

---

[7]In order to ensure the optimality of the proposed consumption rule in (4.15), the feasibility and transversality conditions must be imposed. The feasibility requires that both consumption and capital be positive, i.e., $0 < u_i < 1$. The transversality condition that guarantees the convergence of value function is $\varepsilon_i(\gamma_i - 1)[A - \gamma_i(\sigma A)^2/2] + \varepsilon_i(\rho_i - \gamma_i) + 1 > 0$. (See, for example, Smith 1996a, b, and Stokey and Lucas 1989.) Although some restrictions are required on such parameters as A, ε_i, γ_i, ρ_i and σ, the analysis remains valid even under the restrictions.

The saving rate of the deterministic part s_i consists of two effects: one is the *profitability effect* $\varepsilon_i(\tilde{r}_i - \rho_i)$, and the other is the *precautionary effect* $\gamma_i(\sigma A)^2 k_i/2$. When the risk-adjusted rate of return $\tilde{r}_i$ exceeds the time preference rate ρ_i, then individual i is willing to postpone a part of current consumption for future profits, and hence for future consumption. Conversely, if $\tilde{r}_i$ is lower than or equal to ρ_i, then the individual has no incentive to accumulate capital for future profits. Hence, the term $\varepsilon_i(\tilde{r}_i - \rho_i)$ can be considered as the *profitability effect*. However, when $\tilde{r}_i = \rho_i$, she saves $\gamma_i(\sigma A)^2 k_i/2$ against unexpected future shocks. The term $\gamma_i(\sigma A)^2/2$ can therefore be interpreted as the *precautionary effect*.

We should note that the size of the *profitability effect* $\varepsilon_i\ (\tilde{r}_i - \rho_i)$ is a product of ε_i and $\tilde{r}_i - \rho_i$. If the risk-adjusted rate exceeds the time preference rate, i.e., $\tilde{r}_i > \rho_i$, then individual i saves and hence accumulates capital, and vice versa. However, the size of the saving depends not only on the gap between $\tilde{r}_i$ and ρ_i but also the adjustment cost. Since the intertemporal elasticity of substitution is the degree at which individuals adapt to changes in consumption over time, the inverse $1/\varepsilon_i$ is considered as the adjustment cost of capital accumulation measured in terms of *utility*. Putting it differently, the intertemporal elasticity ε_i presents the flexibility of the individual over time: the higher the elasticity is, the larger the changes in consumption are.

An increase in the degree of risk aversion γ_i and/or the uncertainty parameter σ raises the saving rate through the *precautionary effect*, while it reduces the saving rate through the *profitability effect* because it lowers the risk-adjusted rate of return. As is discussed in the above, the intertemporal elasticity of substitution ε_i plays an important role in determining the size of the *profitability effect*. The total impact therefore depends crucially on ε_i. In fact, from (4.17), we have the following relationships:

$$\frac{\partial s_i}{\partial \gamma_i} \overset{>}{\underset{<}{=}} 0 \quad \Leftrightarrow \quad \varepsilon_i \overset{<}{\underset{>}{=}} 1 \quad , \quad \frac{\partial s_i}{\partial \sigma} \overset{>}{\underset{<}{=}} 0 \quad \Leftrightarrow \quad \varepsilon_i \overset{<}{\underset{>}{=}} 1. \tag{4.18}$$

A small ε_i implies that individual i prefers smooth consumption. Since the individual tries to avoid large changes in consumption over time, the *profitability effect* is smaller than the *precautionary effect* when ε_i is smaller than one. This implies that an increase in γ_i and/or σ increases saving when $\varepsilon_i < 1$, and vice versa.

Proposition 4.1 *When the intertemporal elasticity of substitution of an individual is smaller than one, then an increase in risk-aversion and/or in uncertainty increases the saving rate, and vice versa.*

4.2.4 Macroeconomic Equilibrium

Denoting the total capital stock of the economy by k, the capital stock that individuals H have by k_H, and the capital stock that individuals L have by k_L, they are as follows:

$$k = \int_0^1 k_i di = \int_0^\lambda k_i di + \int_\lambda^1 k_i di, \;\; k_H = \int_0^\lambda k_i di, \;\; k_L = \int_\lambda^1 k_i di, \;\; and \;\; k = k_H + k_L. \tag{4.19}$$

By use of (4.16) and (4.17), we can analyze the dynamics of k_H, k_L and k. For k_H,

$$dk_H = \int_0^\lambda dk_i di = \int_0^\lambda (s_H k_i dt + \sigma A k_i dz_i) di = s_H \int_0^\lambda k_i dt di + \sigma A \int_0^\lambda k_i dz_i di.$$

Taking (4.3) into account, the above becomes

$$dk_H = s_H k_H dt, \;\; dk_H = s_H k_H dt, \;\; \dot{k}_H = dk_H/dt = s_H k_h \;\; or \;\; g_H = \dot{k}_H/k_H = s_H, \tag{4.20}$$

where

$$s_H = \varepsilon_H(\tilde{r}_H - \rho_H) + \frac{\gamma_H(\sigma A)^2}{2} = \varepsilon_H\left(A - \frac{\gamma_H(\sigma A)^2}{2} - \rho_H\right) + \frac{\gamma_H(\sigma A)^2}{2}. \tag{4.21}$$

Similarly,

$$dk_L = s_L k_L dt, \;\; \dot{k}_L = dk_L/dt = s_L k_L, \;\; or \;\; g_L = \dot{k}_L/k_L = s_L, \tag{4.22}$$

where

$$s_L = \varepsilon_L(\tilde{r}_L - \rho_L) + \frac{\gamma_L(\sigma A)^2}{2} = \varepsilon_L\left(A - \frac{\gamma_L(\sigma A)^2}{2} - \rho_L\right) + \frac{\gamma_L(\sigma A)^2}{2}, \tag{4.23}$$

and therefore,

$$dk = dk_H + dk_L = s_H k_H dt + s_L k_L dt \;\; or \;\; \dot{k} = dk/dt = s_H k_H + s_L k_L. \tag{4.24}$$

The above can be written in growth rates as follows:

$$g = \frac{\dot{k}}{k} = s_H \frac{k_H}{k} + s_L \frac{k_L}{k} = s_H\theta_H + s_L\theta_L = s_H\theta_H + s_L(1-\theta_H) = g_H\theta_H + g_L(1-\theta_H), \tag{4.25}$$

where $\theta_i = k_i/k$ is the capital share of individuals $i\,(i = H, L)$, and hence $\theta_H + \theta_L = 1$.

The feasibility of the optimal solution (see footnotes 6 and 7) ensures that both types of individuals continue to accumulate capital, i.e.,

$$s_H(= g_H) > 0, \quad s_L(= g_L) > 0. \tag{4.26}$$

The above conditions also guarantee that $g(= g_H\theta_H + g_L(1-\theta_H)) > 0$, i.e., the economy continues to grow. The dynamics of the economy is determined by (4.20), (4.22) and (4.24).

4.2.5 The Dynamics of Wealth Distribution

By the definition of θ_H, we have

$$\dot{\theta_H}/\theta_H = \dot{k}_H/k_H - \dot{k}/k_H == (s_H - s_L)(1-\theta_H) \ or \ \dot{\theta}_H = (s_H - s_L)(1-\theta_H)\theta_H. \tag{4.27}$$

As is shown in Panel (a) of Fig. 4.1, if $s_H > s_L$, then θ_H increases over time and approaches to one asymptotically. Since individuals L also keep accumulating capital, the entire capital of the economy will not be possessed by individuals H even in the long-run. However, it surely implies that individuals H's behavior becomes a dominant determinant of the long-run growth rate in the sense that the economy-wide growth rate g approaches the capital accumulation rate of individuals H, g_H. This can be confirmed by substituting $\theta_H = 1$ (or $\theta_L = 0$) into (4.25):

$$g = \dot{k}/k = s_H = g_H \quad .$$

If, in contrast, $s_H < s_L$, then θ_H decreases over time and approaches to zero asymptotically, as Panel (b) of Fig. 4.1 shows. In other words, θ_L increases toward one over time and hence individuals L come to have the almost all (but not the entire) capital of the economy. As a result, the growth rate of the economy eventually approaches to the saving rate of individuals L. It also can be confirmed by substituting $\theta_L = 1$(or $\theta_H = 0$) into (4.25):

$$g = \dot{k}/k = s_L = g_L.$$

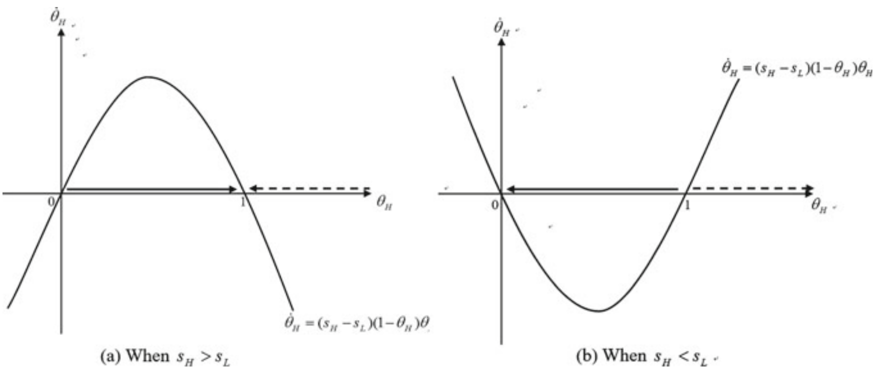


Fig. 4.1 Transitional dynamics of capital share of individuals

4.3 On Ramsey's Conjecture

In what follows, without loss of generality, we assume that individuals H are less patient than individuals L, i.e., $\rho_H > \rho_L$. In a neoclassical growth model without long-run growth, Ramsey's (1928) conjecture holds that the most patient individuals, i.e., individuals L, come to own the entire capital of the economy in the long-run. The conjecture holds in a growing economy only when the most patient individuals accumulate capital, while others do not in the long-run. This is the case when $s_L > 0 \geqslant s_H$. For, however, an economy in which all individuals accumulate capital in the long run, we cannot use Ramsey's conjecture in the strict sense. As the analysis in the previous section shows, the capital share of one group will approaches to one even if all individuals have positive savings in the long-run. Hence, let us introduce the following definition of Ramsey's conjecture in a growing economy.

Definition: Ramsey's conjecture in a growing economy
The *most patient* individuals come to own the almost all (but not the entire) capital of the economy in the long-run. In other words, the share of capital owned by the *most patient* asymptotically approaches to one.

4.3.1 Ramsey's Conjecture in a Stochastically Growing Economy

Ramsey's conjecture holds when $s_L > s_H$. From (4.17), we have

$$s_L - s_H = \varepsilon_L(\tilde{r}_L - \rho_L) - \varepsilon_H(\tilde{r}_H - \rho_H) + (1/2)(\gamma_L - \gamma_H)(\sigma A)^2,$$

or

$$s_L > s_H \quad \Leftrightarrow \quad \varepsilon_L(A - \rho_L) > \varepsilon_H(A - \rho_H) - \frac{[(1-\varepsilon_L)\gamma_L - (1-\varepsilon_H)\gamma_H](\sigma A)^2}{2}. \tag{4.28}$$

Proposition 4.2 *Ramsey's conjecture in a stochastically growing economy holds if one of the following two conditions is satisfied:*

(i) The preferences of all individuals are the same except for time preference, i.e., patience.
(ii) The intertemporal elasticities of substitution of all individuals are one.

Proof (i) Suppose that $\gamma_H = \gamma_L = \gamma$ and $\varepsilon_H = \varepsilon_L = \varepsilon$. Then, the latter inequality in (4.28) becomes $\rho_H > \rho_L$, which is true by definition.

(ii) Suppose that $\varepsilon_H = \varepsilon_L = 1$. Then, the latter inequality in (3.14) becomes $\rho_H > \rho_L$.

On one hand, condition (i) is self-evident in the sense that time preference is an only factor to explain differences in capital accumulation. On the other hand, condition (ii) needs some explanation. When $\varepsilon_i = 1$, then the *precautionary effect* $\gamma_i(\sigma A)^2/2$ is exactly cancelled out by the *profitability effect* $\varepsilon_i\ (\tilde{r}_i - \rho_i)$ because it becomes $\tilde{r}_i - \rho = A - \gamma_i(\sigma A)^2/2 - \rho$. In other words, the term $\gamma_i(\sigma A)^2/2$ disappears in s_i. Since the two effects are cancelled out, no room exists for risk aversion and uncertainty to influence the saving behavior. Therefore, differences in saving behavior are explained only by differences in time preference.

4.3.2 Failure of Ramsey's Conjecture (1): Difference in Intertemporal Substitution

If $s_H > s_L$, then Ramsey's conjecture fails. In order to see the effect of intertemporal substitution clearly, let us eliminate the effect of risk aversion. Suppose that individuals of both types are equally risk averse, i.e., $\gamma_H = \gamma_L = \gamma$. Then, since $\tilde{r}_H = \tilde{r}_L = \tilde{r}$ from (4.17),

$s_H - s_L = \varepsilon_H(\tilde{r} - \rho_H) - \varepsilon_L(\tilde{r} - \rho_L)$ with $\tilde{r} = A - [\gamma(\sigma A)^2/2]$.

Hence, the following relationships hold:

$$s_H > s_L \quad \Leftrightarrow \quad \frac{\varepsilon_H}{\varepsilon_L} > \frac{\tilde{r} - \rho_L}{\tilde{r} - \rho_H} \quad \Leftrightarrow \quad \varepsilon_H > \left(\frac{\tilde{r} - \rho_L}{\tilde{r} - \rho_H} \right) \varepsilon_L. \tag{4.29}$$

Since, by assumption, $(\tilde{r} - \rho_L)/(\tilde{r} - \rho_H) > 1$, the above shows that Ramsey's conjecture fails when the intertemporal elasticity of substitution of *impatient* individuals ε_H is sufficiently larger than that of the most patient ε_L, even if both types of individuals are equally risk-averse.[8] Hence, the possibility exists that the *most impatient* individuals, individuals H in the model, come to own the almost all (but not the entire) capital of the economy in the long-run, as is shown in Panel (a) of Fig. 4.1. These observations are summarized in the following proposition.[9]

Proposition 4.3 *Other things being equal, if the most impatient individuals have the largest intertemporal elasticity of substitution, then they can come to own the almost all (but not the entire) capital of the economy in the long-run. In other words, Ramsey's conjecture fails.*

The intuition behind Proposition 4.3 is straightforward. The size of the *profitability effect* is a product of two terms: ε_i and $\tilde{r} - \rho_i$. By definition, $\tilde{r} - \rho_L > \tilde{r} - \rho_H$. If, however, ε_H is sufficiently larger than ε_L, then $\varepsilon_H(\tilde{r} - \rho_H)$ can be larger than $\varepsilon_L(\tilde{r} - \rho_L)$.

[8] Putting it the other way around, the conjecture remains true as long as the elasticity of the most patient individuals is *not* smaller than that of others.

[9] Using a deterministic model, Nakamura (2014) shows the same proposition.

4.3.3 Failure of Ramsey's Conjecture (2): Difference in Risk Aversion

To examine the effect of risk aversion, let us suppose that the intertemporal elasticities of substitution are the same across individuals, i.e., $\varepsilon_H = \varepsilon_L = \varepsilon$. Then, from (4.17),

$$s_H - s_L = \varepsilon(\rho_L - \rho_H) + (1/2)(1-\varepsilon)(\gamma_H - \gamma_L)(\sigma A)^2.$$

Therefore, the following relationship is derived:

$$s_H > s_L \quad \Leftrightarrow \quad (1/2)(1-\varepsilon)(\gamma_H - \gamma_L)(\sigma A)^2 > \varepsilon(\rho_H - \rho_L).$$

Hence, when $\varepsilon < 1$, i.e., $1 - \varepsilon > 0$, then

$$s_H > s_L \quad \Leftrightarrow \quad \gamma_H > \gamma_L + 2\varepsilon(\rho_H - \rho_L)/(1-\varepsilon)(\sigma A)^2. \tag{4.30}$$

Also, when $\varepsilon > 1$ i.e., $\varepsilon - 1 > 0$, then

$$s_H > s_L \quad \Leftrightarrow \quad \gamma_H < \gamma_L - 2\varepsilon(\rho_H - \rho_L)/(\varepsilon - 1)(\sigma A)^2. \tag{4.31}$$

The above relationships lead to the following proposition.

Proposition 4.4 *When the intertemporal elasticity of substitution is lower than one, then Ramsey's conjecture fails if impatient individuals are sufficiently more risk-averse than the most patient individuals. When, in contrast, the intertemporal elasticity is higher than one, then the conjecture fails if impatient individuals are sufficiently less risk-averse than the most patient individuals.*

An increase in risk-aversion reduces the risk-adjusted rate of return and hence reduces saving through the *profitability effect*, while it increases saving through the *precautionary effect*. The magnitude of the *profitability effect* depends on the intertemporal elasticity of substitution. If the elasticity is larger than one, then the *profitability effect* dominates the *precautionary effect*, and vice versa. Thus, high risk aversion implies a low saving rate when the elasticity is larger than one, and vice versa. Suppose, therefore, that the elasticity is larger than one. Then the saving rate of impatient individuals can be higher that of the most patient individuals if the impatient are sufficiently *less risk averse* than the most patient. As a result, Ramsey's conjecture fails. Suppose, conversely, that the elasticity is smaller than one. Then the conjecture fails if the impatient are sufficiently *more risk averse* than the most patient.

Propositions 4.3 and 4.4 lead us to the following proposition.

Proposition 4.5 *Suppose that the intertemporal elasticity of substitution is lower than one and the most impatient individuals are sufficiently more risk-averse than* *others. Then they can own the almost all (but not the entire) capital of the economy in the long-run. Suppose, conversely, that the elasticity is higher than one and the most impatient individuals are sufficiently less risk-averse than others. Then they can own the almost all capital of the economy in the long-run.*

It is counterintuitive that the possibility exists that almost all (but not the entire) capital will be possessed by the *most impatient* individuals with the *least risk aversion* in the long run. If, however, it is taken into account that the *profitability effect* is larger than the *precautionary effect* with a high intertemporal elasticity of substitution, then we can see it. The *profitability effect* is a product of the elasticity ε_i and the net-benefit $\tilde{r}_i - \rho_i$. Since the risk-adjusted rate of return for the impatient $\tilde{r}_H$ is higher than that for the patient $\tilde{r}_L$ if the impatient are less risk-averse than the patient, it is possible even under $\rho_H > \rho_L$ that $\tilde{r}_H - \rho_H > \tilde{r}_L - \rho_L$. Although the *precautionary effect* is larger for the patient (individuals L) than for the impatient (individuals H), with a large elasticity (ε_H), the *profitability effect* ($\varepsilon_H(\tilde{r}_H - \rho_H)$) can dominate the *precautionary effect*. This explains the latter half of Proposition 4.5.

4.4 Concluding Remarks

Undoubtedly, uncertainties, risk-aversion and intertemporal substitution have inherent importance in the wealth accumulation. Introducing Kreps-Porteus non-expected utility preferences, which allow independent variation in risk aversion and intertemporal substitutability, this paper has examined Ramsey's conjecture in a stochastically growing economy. As a result, we have put forward several propositions to show that the conjecture does not necessarily hold.

Among the propositions Proposition 4.5 is the most particular to the model presented in this paper. Suppose the intertemporal elasticity is substantially high. Then the possibility exists that the *most impatient and least risk averse* individuals come to own the almost all capital of the economy in the long-run. The conventional wisdom suggests that risk-averse people save more than the less risk-averse via *precautionary saving* in uncertain environments. However, as this paper shows, the inference is not always true. The subjective rate of return adjusted by uncertainty and risk aversion for less risk-averse individuals is higher than that for more risk-averse individuals. Hence, the less risk-averse individuals have more incentive to save than the more risk-averse. With high intertemporal substitution, this effect dominates the precautionary saving. Therefore, the less risk-averse individuals save more and accumulate more capital.

The analysis in this paper is based on a simple setup. It is hoped that this paper stimulates future research on the wealth distribution under uncertain environments in more general setups.

Acknowledgements The author would like to thank Professors Kentaro Iwatsubo, Tetsugen Haruyama, Daishin Yasui and seminar participants at the Kobe University. The financial support of Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (15K03431 and 17K18564) is gratefully acknowledged. The usual disclaimer applies.

Appendix

Applying Ito's lemma to the maximization of $V_i(t)$ in (4.4), we get the following stochastic Bellman equation:

$$0 = \max_{c_i}\{[(1-\gamma_i)/(1-1/\varepsilon_i)]c_i^{1-1/\varepsilon_i} - \rho_i f([1-\gamma_i]J_i(k_i)) \\ + (1-\gamma_i)f'([1-\gamma_i]J_i(k_i))[J_i'(k_i)(Ak_i - c_i) + (1/2)J_i"(k_i)(\sigma A k_i)^2]\}. \quad (4.32)$$

From (4.32), the first-order condition for c_i is

$$c_i^{-1/\varepsilon_i} = f'([1-\gamma_i]J_i(k_i))J_i'(k_i). \quad (4.33)$$

Equation (4.4) suggests that $J_i(k_i)$ takes the following form:

$$J_i(k_i) = (b_i k_i)^{1-\gamma_i}/(1-\gamma_i), \quad (4.34)$$

where b_i is a positive constant to be determined. Now, (4.33) becomes

$$c_i = u_i k_i \;\; with \;\; u_i = b_i^{1-\varepsilon_i}. \quad (4.35)$$

Substitution of (4.34) and (4.35) into (4.32) gives

$$u_i = A - \varepsilon_i(A - \rho_i) - \frac{(1-\varepsilon_i)\gamma_i(\sigma_i A)^2}{2}, \quad (4.36)$$

and therefore,

$$b_i = \left[A - \varepsilon_i(A - \rho_i) - \frac{(1-\varepsilon_i)\gamma_i(\sigma_i A)^2}{2}\right]^{1/(1-\varepsilon_i)}. \quad (4.37)$$

Substitution of (4.37) into (4.35) gives the following consumption function:

$$c_i = u_i k_i = \left[A - \varepsilon_i(A_i - \rho_i) - \frac{(1-\varepsilon_i)\gamma_i(\sigma_i A)^2}{2}\right]k_i. \quad (4.38)$$

References

Becker, R.A. 1980. On the long-run steady state in a simple dynamic model of equilibrium with heterogeneous households. *Quarterly Journal of Econonomics* 95 (2): 375–82.

Becker, R.A., and I. Zilcha. 1997. Stationary Ramsey equilibria under uncertainty. *Journal of Economic Theory* 75: 122–140.

Boyd III, J.H., 2000. *Sustained growth with heterogeneous households*, International Florida University. http://faculty.fiu.edu/~boydj/papers/growth.pdf.

Duffie, D., and L.G. Epstein. 1992. Stochastic differential utility. *Econometrica* 60 (2): 353–94.

Epstein, L.G., and S.E. Zin. 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica* 57: 937–969.

Epstein, L.G., and S.E. Zin. 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: an empirical analysis. *Journal of Political Economy* 99: 263–286.

Kreps, D.M., and E.L. Porteus. 1978. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46: 185–200.

Kreps, D.M., and E.L. Porteus. 1979. Dynamic choice theory and dynamic programming. Econometrica 47: 91–100.

Mitra, T., and G. Sorger. 2013. On Ramsey's conjecture. *Journal of Economic Theory* 148 (5): 1953–1976.

Nakamura, T. 2014. On Ramsey's conjecture with AK technology. *Economics Bulletin* 34 (2): 875–884.

Obstfeld, M. 1994a. Evaluating risky consumption paths: the role of inter temporal substitutability. *European Economic Review* 38: 1471–1486.

Obstfeld, M. 1994b. Risk-taking, global diversification, and growth. *American Economic Review* 84: 1310–1329.

Ramsey, F.P. 1928. A mathematical theory of saving. *Economic Journal* 38: 543–559.

Weil, Philippe. 1989. The equity puzzle and the risk-free puzzle. *Journal of Monetary Economics* 24: 401–421.

Smith, W.T. 1996a. Feasibility and transversality conditions for models of portfolio choice with non-expected utility in continuous time. *Economics Letters* 53 (2): 123–31.

Smith, W.T. 1996b. Taxes, uncertainty, and long-term growth. *European Economic Review* 40 (8): 1647–64.

Stokey, N., and R.E. Lucas Jr. 1989. *Recursive methods in economic dynamics*. Harvard University Press.

Chapter 5
Impact of Copyright Protection on Re-creation of Digital Contents When Expression and Idea Are Divisible

Takuya Nakaizumi

Abstract We consider the case in which the expression and the idea are completely divisible and copyright protection cannot be applied to very closely substitutable derivative work. Even if the ideas of these works are very similar, the substitute is regarded as an independent work, and not as a copyright infringement. Therefore, it is possible to supply closely substitutable goods based on similar ideas. Thus, the monopoly rent of the copyright holders might decrease, so they face the risk of failing to recover the initial cost. Assuming Bertrand competition with zero marginal cost, considering the nature of digital contents, we show that imitators have no incentive to enter the market in vertical product differentiation except in the case of a sufficiently high quality of imitation. Even if they can enter the product-differentiated market, they do not compete and divide the market to produce differentiated works. Thus, the original producer's monopolistic profit is maintained under the general copyright system.

5.1 Introduction

According to the nature of informational goods, copyrighted works have attributes of both non-excludability and non-competitiveness. Therefore, because of its low marginal cost, the prices of copyrighted works are preferably set low, or even at zero, if they are digital contents and the marginal cost is zero. However, if the price of a copyrighted work is equal to the marginal cost, the creation cost cannot be recovered and is ex ante considered a fixed cost as an incentive for content creation. Thus, the copyright law gives producers the exclusive right to use the works and to gather the creation cost within a certain period of copyright protection. Thus, it is illegal to sell copyrighted materials without any permission from the producers.

T. Nakaizumi (✉)
Kanto Gakuin University, Yokohama, Japan
e-mail: nakaizum@kanto-gakuin.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_5

In principle, on the other hand, copyright protects only expressions and not ideas. Even if the ideas of copyrighted works are very similar, the imitation is regarded as an independent work, and not as a copyright infringement, as long as the expressions are quite different. In other words, it is possible to supply closely substitutable goods based on similar ideas if the expressions are different. Thus, this effect might decrease the monopoly rent of copyright holders, who may have concerns about losing their monopolistic benefit easily.

Apparently, the depreciation of the monopolistic profit of the initial copyright holder is determined by several factors, such as production cost of similar goods and the amount of substitute goods. Because existing works are conceived from ideas, while alternative goods are obtained from existing works and only the expressions change, the cost of creation is naturally lower for substitute works than initial works. Thus, the profit on the original work might be lower than the creation cost.

Thus, the creation cost of an initial work might not be recovered, and initial works not produced. In this paper, we set up a benchmark model to check whether initial works are created or not.

The outcome actually depends on various factors, not only the cost difference but also the quality difference between the original and derivative works and the time required for re-creation. In this paper, we develop a simple Stackerberg market with Bertrand competition and zero marginal cost within which both initial works are first produced and then the imitator considers entering the market to produce a closely substitutable work. We check whether the monopoly profit can be obtained and whether or not the initial work is supplied, with both vertical and horizontal product differentiation.

We derive an astonishing result: with the strong pressure of Bertrand competition, the entrants obtain no profit even if they enter the market; thus, they have no incentive to produce an imitation work in vertical product differentiation. In horizontal product differentiation, they have an incentive to create a derivative work, but they differentiate the product sufficiently to avoid competing with each other unless strong pressure from price competition reduces the entrant profit to zero.

We also refer to an important aspect: expressions and ideas are closely related. In reality, it seems difficult to isolate ideas from expressions. Therefore, we also refer to the extent to which the inseparability of ideas and expressions influences monopoly profits.

As Arai Arai (2018) pointed out, no previous theoretical work has analyzed the divisibility of expressions and ideas.

In Sect. 5.2, we consider vertical product differentiation. In Sect. 5.3 we derive the market structure and the degree of horizontal product differentiation using the Hotelling model. In Sect. 5.4, we extend the model in which ideas and expressions are partially dispensable. Section 5.5 presents the concluding remarks.

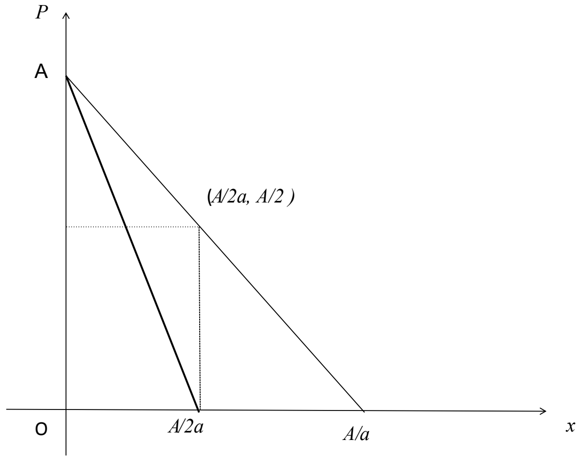


Fig. 5.1 Optimal contents provision without price discrimination

5.2 The Vertical Product Differentiation Model and the Benchmark Outcome

Suppose a market of digital content in which demand is simplified as a linear function $D(P)$; P is the price of the content and is described as a linear inverse demand function, $P^D(x) = A - ax$. The marginal cost is $= 0$. Further, the initial producer's cost to create the content is an ex-ante fixed cost, F.

In Proposition 5.1, we first set the condition that the initial producer creates the content when s/he could sell it as a monopoly product (Fig. 5.1).

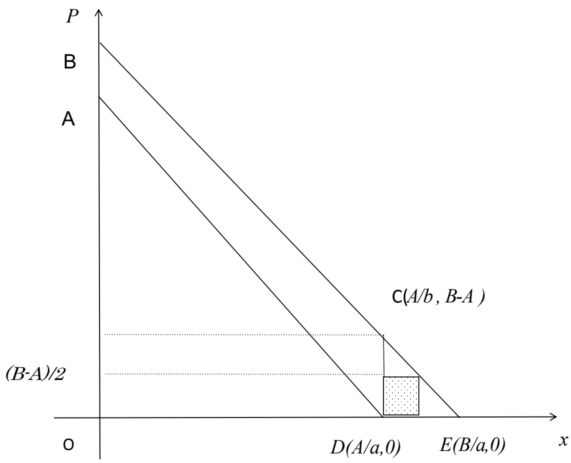
Proposition 5.1 *If the initial producer sells it as a monopoly product without price discrimination, s/he produces it if and only if* $\frac{A^2}{4b} \geq F$. *With perfect price discrimination, s/he produces it if and only if* $\frac{A^2}{b} \geq F$.

In Proposition 5.1, there are no competitors. We now turn to the case in which it is possible to produce closely substitutable content by imitating the idea of the initial content, while the expression is quite different. This paper assumes that only one competitor enters the market and provides the same kind of content, which is a perfect substitute. We then consider vertical product differentiation. The difference is the fixed cost F'. We obtain an astonishing result through a quite well-known analytical method.

Proposition 5.2 *Suppose that both the initial producer and an imitator produce the same kind of content in the market. If they compete on price in Bertrand competition, the equilibrium price is zero because the marginal cost is zero. Thus, the imitator has no incentive to enter the market.*

Bertrand competition is too strong for the recovery of either the initial producer's or the entrant's creation cost. Imitators would enter the market only if their product

Fig. 5.2 Optimal contents provision without price discrimination



quality is sufficiently better than the original to compensate for the fixed cost of producing the re-creation. Thus, we consider vertical product differentiation.

Proposition 5.3 *Suppose the quality of the imitation is better than that of original content and the inverse demand function of the imitation is simplified as* $P^D = B - ax,\ B > A$. *The imitators have an incentive to enter the market if and only if their creation cost F' is*

$$F' \leq \frac{(B-A)^2}{4a} \tag{5.1}$$

By Fig. 5.2, if the fixed cost is less than the area of the small square described by Fig. 5.2, the fixed cost is recovered, and the imitator has an incentive to enter the market. The original producer, however, has no reward and is not compensated for the fixed cost. Thus, s/he has no incentive to create the original.

Proposition 5.4 *If Eq. (5.1) holds, the original producer has no reward and is not compensated for the fixed cost. Thus, s/he has no incentive to create the original.*

Interestingly, the necessary condition for the imitator to enter the market is that the imitation should have higher quality than original. Naturally, the creation cost of the imitation is lower than that of the original, even if the expression is completely changed. However, even if the cost of the imitation is lower than that of the original, the imitators cannot enter the market if the quality or attractiveness of the imitation is higher than that of the original and it seems difficult to produce. Thus, a prohibition against the use of expressions is all that is needed to allow enough monopoly power to compensate for the original producer's creation cost, although preventing imitations that are more attractive than the original might be a "related right."

The time-lag to produce an imitation is worth considering. The initial producer obtains some profit if the imitator delays to enter the market even if the competitor's product quality is sufficiently higher. Thus, the more the delay, the greater the incentive to create original works.

5.3 Horizontal Product Differentiation of the Hotelling Model

Derivative works are characterized by not only vertical product differentiation but also horizontal differentiation. In this section we consider horizontal product differentiation following the Hotelling model based on the sequential location choice of Fleckinger and Lafay (2003).

We study a market where consumers belong to a segment [0, 1]. Consumers are spread over the market according to a uniform distribution. There are two stages in the leader-and-follower game. First, the initial producer, as the leader, chooses the location. We simplify the location choice of the initial producer by assuming 0. Although it seems a strong assumption in a location choice game, it is more natural for quality differentiation of the content. We discuss this point later. Then, the imitator chooses location $ain[0, 1]$. After the location choice stage, both choose the price simultaneously based on Bertrand competition. If they compete with each other, they are supposed to engage in Bertrand competition, and the price goes to zero. However, if they do not have to compete, because other suppliers cannot satisfy the users' preferences (since they too difficult to reflect in the content), one producer manufactures the product monopolistically. In addition, we assume no price discrimination for the initial content producer and imitator, although they can set a single different price in a segmented market.

The consumer's content evaluation is constant for any $xin[0, 1]$ and v. Further, the difference in preferences is expressed as the distance. The marginal disutility to consume the content from the difference location is assumed to be v.

Thus, if there is no entrant, the original content producer faces the downslope demand function as in the previous section, $P^d(x) = v - vx, x \in [0, 1]$. The imitator enters in sequence and chooses the location based on the leader's position, 0. The leader and the follower have full and complete knowledge about the system.

We should note several points. First, quality choice is modeled in this paper as location choice. The initial content producer's quality is first fixed at 0. The leader usually chooses the location freely, most often at the center of the distribution, in the general Hotelling model. In this quality choice model, however, it is difficult to consider the notion of the side. If the incumbent chooses the center of the interval, only the entrant chooses either side of the interval. Even if the quality is very similar, the incumbent gets at most half of the share and is definitely ensured at least half of the market. That a half share is ensured to the incumbent even without the support of copyrights is an interesting and not trivial result. In order to avoid trivial results, we assume that the initial producer's content quality is chosen at 0. Second, considering

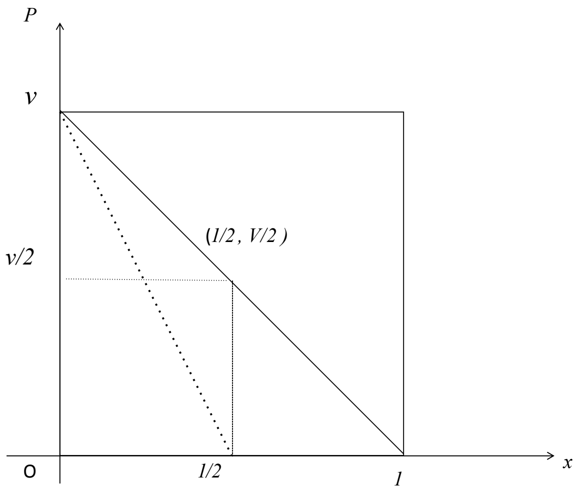


Fig. 5.3 Optimal contents provision without price discrimination

the nature of the content, location decisions are naturally costly and are made once and for all; relocating is considered prohibitively costly and is not permitted.

We solve the problem through backward induction. At the second stage, the imitator decides the price and whether to compete or not based on the location choice, a. We present this situation in Definition 5.1 (Fig. 5.3).

Definition 5.1 If the imitator and initial producer, but not the border agent, compete, both set the price at zero. If no agent, other than the border agent, can access both producers, the market is defined as segmented. Then, the imitator and initial producer never compete, and each sets a single price different from the other's.

Then, they choose the location as stated in Lemma 5.1.

Lemma 5.1 *The incumbent sets location a far enough to prevent competition between producers for any user.*

If they compete, the price falls to zero, which is definitely lower than the competition-based price. Thus, the entrant's optimal strategy is to set a price higher than 0 by choosing a, where no customer can access both producers. Thus, we have the following lemma for prices based on a

Lemma 5.2 *The incumbent sets the price* $p_i = v/2$ *and obtains the monopoly rent* $\pi_i = v/4$*, while the entrant provides only* $[1/2, 1]$*. Thus,* $1/2 \geq a \geq 1$*, and the price is set at* $p_e = \min va, v/2 - va$*.*

Thus, we derive the following proposition.

Proposition 5.5 *Thus, the entrant chooses* $a = 3/4$ *and sets* $p_2 = 3v/4$*. The market is segmented and the entrant receives more revenue than the incumbent.*

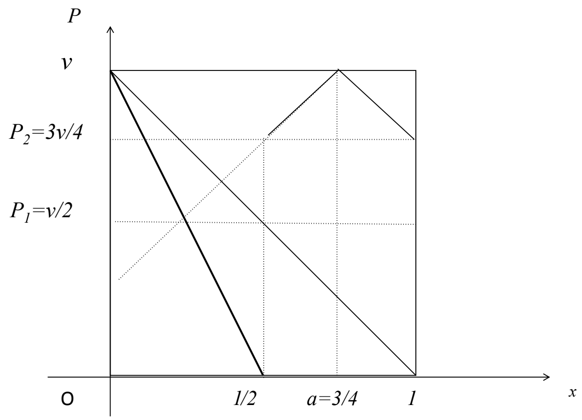


Fig. 5.4 Location choice and profit of entrant

Thus, we derive the following proposition:

Proposition 5.6 *The incumbent is ensured a monopolistic profit with the price set at* $p_1 = v/2$. *Thus, even under horizontal product differentiation, the initial producer is ensured a monopoly rent and has no disincentive for ex-ante content production.*

In fact, there is a very complicated boundary problem: if $a \geq 1/2$, customer $x = 1/2$ can access both content types. In this model, we avoid the boundary problem by assuming that the boundary customer chooses either content with the same probability (Fig. 5.4).

A strong disincentive of Bertrand competition is that the entrant refuses to compete with the incumbent in both vertical and horizontal product differentiation. Thus, the initial producer is ensured a monopoly rent unless the entrant's product is of a sufficiently higher quality to allow the imitation cost to be recovered in the face of Bertrand competition.

5.4 When Expressions and Ideas are Partially Divisible

In the previous section, it is assumed that ideas and expressions are dispensable. Assuming Bertrand competition with zero marginal cost, considering the nature of digital contents, we show that unless the quality of the imitation is sufficiently high, imitators have no incentive to enter the market under vertical product differentiation. Even if they can enter a product-differentiated market, they do not compete and divide the market to produce differentiated works. Thus, no entrant appears, and the monopoly power of the initial producer can be maintained.

From a more realistic viewpoint, on the other hand, expressions and ideas are partially inseparable. In this section, we consider a possibility of investment for the re-creator to increase the indivisibility while the original author has an incentive to increase the indivisibility to deter the entrant by producing re-creative work.

Naturally, there is a certain threshold to decide whether any work is a copyright infringement or not. If the original author can easily set the threshold high enough to deter re-creation by an entrant, entry is impossible; similar results apply as for other intellectual property rights.

However, where such a threshold is too costly to create, it is easy to enter the market by imitating the original idea; the discussion then remains the same as in the previous section.

Then, what about the middle case? First, even if you do so, the re-creator cannot generate revenue unless s/he creates a more valuable work. It follows that the same results as in previous section apply in case ideas and expressions are partially dispensable.

5.5 Concluding Remarks

We consider the case in which the expression and the idea are completely divisible and copyright protection cannot be applied to very closely substitutable derivative work. Even if the ideas of the derivative work are very similar, it is regarded as an independent work, and not as a copyright infringement. Thus, it is possible to supply closely substitutable goods based on similar ideas. This might decrease the monopoly rent of the copyright holders, so they face the risk of failing to recover the initial cost.

Because of the nature of digital content, we assume Bertrand competition with zero marginal cost in the market. We show that except in the case of a sufficiently high quality of imitation, imitators have no incentive to enter the market in vertical product differentiation. Even if they can enter the product-differentiated market, they do not compete and divide the market to produce differentiated works. This is because Bertrand competition is too strong for the entrant to recover the imitation cost. It is also applicable where the idea and the expression are partially dispensable. Thus, the original producer's monopolistic profit is maintained under the general copyright system.

References

Anderson, S. 1987. Spatial competition and price leadership. *International Journal of Industrial Organization* 5: 369–398.

Arai, Y. 2018. Intellectual property right protection in the software market. *Economics of Innovation and New Technology* 27(1): 1–13.

D'Aspremont, C., J. Jaskold Gabszewicz, and J.F. Thisse. 1979. On Hotelling's stability in competition. *Econometrica* 47: 1145–1150.

Fleckinger, Pierre and Thierry Lafay. 2003. Horizontal differentiation and price competition with sequential entry. *mimeo*

Hotelling H. 1929. Stability in competition. *The Economic Journal* 39: 41–57.

Loertscher, Simon and Gerd Muehlheusser. 2011. Sequential location games. *The RAND Journal of Economics* 42(4): 639–663 (Winter 2011).

Younies, Hassan, and H.A. Eiselt. 2011. *Sequential Location Models H, 163–178*. Foundations of Location Analysis Springer: A. Eiselt and Vladimir Marianov Edited.

Part II
Inequality, Redistribution, and Intergenerational Transfers

Chapter 6
An International Comparison of Inequality of Educational Opportunity Using TIMSS

Changhui Kang and Sungjae Oh

6.1 Introduction

In recent years, there has been a growing interest in the effects of economic inequalities on the economy and society as a whole. Prior studies on economic inequality have addressed inequalities of economic achievement such as wages and incomes. Considering the discussion of modern political philosophers on inequalities, recent economic studies have distinguished between inequalities in economic achievement, on the one hand, and inequality of opportunity and inequality of effort, on the other (Roemer 1993, 1998; Fleurbaey 2008; Fleurbaey and Maniquet 2011; Roemer and Trannoy 2016). An economic achievement of an individual is a final outcome that combines effects of the individual's environment (or opportunity) with effects of her effort. Inequalities of achievement can be divided by the inequality of opportunity and the inequality of effort.

Recent views of political philosophers and economists dealing with inequalities suggest that, among the two components of inequality in achievement, inequality of opportunity is morally inacceptable (Dworkin 1981a, b; Arneson 1989; Cohen 1989), since inequality of opportunity is irrelevant to the will of an individual. Such inequality resulting from impacts of an endowed environment is inacceptable because the environment is beyond the control of an individual's volition or voluntary choices. On the other hand, economic inequalities arising from an individual's voluntary choices or differences in effort (i.e., inequality of effort) can be morally acceptable. The achievement of a person who has put more effort is likely to be high, while the achievement of a person who has not is likely to be low. This is the result of voluntary choices of an individual. Of the two types of inequalities, inequality of

C. Kang (✉)
Chung-ang University, Seoul, South Korea
e-mail: ckang@cau.ac.kr

S. Oh
Seoul National University, Seoul, South Korea

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_6

opportunity needs to be corrected through appropriate social policies, but inequality of effort needs not to be socially corrected. Correcting the latter inequality is rather morally inadequate and can worsen economic efficiency.

Policies aiming to mitigate economic inequalities have traditionally attempted to ease inequalities of achievement through redistribution of economic outcomes (i.e., taxation on the rich). Such a redistribution policy is limited since it weakens legitimate efforts and the incentive system of individuals. On the other hand, policies that mitigate inequality of opportunity recognize individual efforts and the importance of incentive systems. Accordingly, recent economic studies of inequality focus on inequality of opportunity, not inequality of achievement alone.

Since Roemer (1998), various methods for empirically measuring the degree of inequality of opportunity for a given society have been developed by many researchers. Ferreira and Peragine (2015), Roemer and Trannoy (2016) and Ramos and Van de Gaer (2016) outline a variety of such methods of measurement. Using such methods, several studies conducted empirical analysis to measure the degree of inequality of opportunity of a society at a given time. For example, Ferreira and Gignoux (2011) analyze Latin American countries and Lefranc et al. (2008) and Bricard et al. (2013) European countries. Cogneau and Mesple-Somps (2008) measured the degree of inequality of opportunity in economic achievement for African countries.

For Korea, Ko and Lee (2011) and Kim et al. (2016) have tried to measure the degree of inequality of opportunity. While such studies play a leading role in showing the degree of inequality of opportunity in Korea, no study compares Korea's inequality of opportunity with other Asian countries and OECD countries with similar levels of economic development.

There are some international studies that measure the degree of inequality of opportunity in a country and compare it to other countries (Ferreira and Gignoux 2011, 2014; Cogneau and Mesple-Somps 2008; Gamboa and Waltenberg 2012; Lefranc et al. 2008; Salehi-Isfahani et al. 2014). Unfortunately, however, they do not report Korea's level of inequality of opportunity. In this paper, we measure the inequality of opportunity for Korea, in particular the inequality of opportunity of education, from 1995 to 2015, comparing Korea's figures with those of other countries. Using such figures, we first trace the changes in the inequality of educational opportunity in Korea over the past several decades. Next, we compare Korea's level of the inequality of educational opportunity with other countries. In so doing we evaluate Korea's relative position in terms of inequality of educational opportunity and suggest directions for future improvement of inequality.

Specifically, employing data from the Trends in International Mathematics and Science Study (TIMSS), an international comparative mathematics and science achievement assessment, we measure the inequality of educational achievement that is proxied by TIMSS test scores of middle school students (8th grade in international standards and 2nd grade in junior high school in Korea). We trace the changes in inequality of educational opportunity for Korea over the past decades and compare Korea's inequality of educational opportunity with other countries.

This paper proceeds as follows: Sect. 6.2 summarizes previous studies that theoretically and empirically analyze inequality of opportunity. Section 6.3 explains our empirical method in detail, and Sect. 6.4 summarizes the main results. Finally, Sect. 6.5 concludes this chapter.

6.2 Previous Literature

John Roemer (1993, 1998) introduced the concept of inequality of opportunity suggested by political philosophers into the framework of economic analysis. Since then, economists have attempted to empirically measure the degree of inequality of opportunity in a given society. Bourguignon et al. (2007) use Brazilian "National Household Survey (1996)" data to calculate the proportion of the inequality of opportunity in hourly wages among Brazilian adults. They consider the birth region of an individual, race, education of the parents, and the occupation group of the father as environmental variables. Their study suggests that in 1996 about 10 to 37 percent of Brazil's total hourly wage inequality is accounted for by inequality of opportunity. Ferreira and Gignoux (2011) use household survey data from six countries in South America (Brazil, Colombia, Ecuador, Guatemala, Panama and Peru) to compare the inequality in adult household spending between 1996 and 2006. They suggest that inequality of opportunity accounts for about 1/4 to 1/2 of the inequality of consumption expenditures. Based on the "Survey on Income and Wealth of Italian Households, 1993, 1995, 1998, 2000," Checchi and Peragine (2010) find that inequality of opportunity (the environment is the education level of the parents) accounts for about 20 percent of the inequality of annual income. Aaberge et al. (2011) and Almås (2011) estimate the inequality of opportunity in Norwegian society. Björklund et al. (2012) measure inequality of opportunity in Sweden.

A study on Korea was first conducted by Kim and Lee (2009). They transform the model of Roemer et al. (2003) into a form suitable for empirical analysis for Korea and calculate the optimal tax rate for leveling opportunity in real world. Based on this, they develop an index of opportunity leveling for a given tax policy. They find that the inequality of opportunities in Korea during the first half of the 2000s was gradually increasing. Ko and Lee (2011) apply a regression method of Ferreira and Gignoux (2011) based on "Korean Labor and Income Panel Study (KLIPS)" data for 2003–2006. They disaggregate income inequalities of 30–55 year old male heads into inequality of opportunity and inequality of effort. When the education of one's father is used as the only environmental variable, about 16–59% of educational inequality and about 2–12% of wage inequality and income inequality of the head are explained by inequality of opportunity. Kim et al. (2016) apply the method of Bourguignon et al. (2007) to the data of KLIPS individuals (30–60 years old) for 2000-2012. About 18–23% of the income inequality in Korea is explained by the inequality of opportunity.

As economic theories of inequality of opportunity and relevant empirical methods are developed, a question naturally arise as to how the degree of inequality of opportunity differs across different countries. To answer this question, a few international comparative studies have been conducted. For example, Ferreira and Gignoux (2011) calculate the share of inequality of opportunity in inequality of household spending in each of six South American countries (Brazil, Colombia, Ecuador, Guatemala, Panama and Peru). Cogneau and Mesple-Somps (2008) measure the proportion of inequality of opportunity in five African countries (Ghana, Uganda, Cote d'Ivoire, Guinea and Madagascar).

While international comparative studies of inequality of opportunity provide useful information for outlining similarities and differences in inequality of opportunity among different countries, these studies use a different type of surveys for different countries. In addition, they calculate inequality of opportunity with different variables set as environment. It is therefore difficult to distinguish whether the disparities in inequality of opportunity across societies represent genuine inequality of opportunity or due to differences in the data or the list of environmental variables. In order to address such limitations, one needs to standardize both survey questionnaires and empirical models for all the countries under study.

Based on this idea, Lefranc et al. (2008) conducted a survey of nine Western countries (Belgium, France, West Germany, UK, Italy, Netherlands, Norway, Sweden, (2003) with the same questionnaires. They use such data to calculate the inequality of opportunity of each country. Bricard et al. (2013) use data from SHARELIFE, a retrospective survey of Europeans over the age of 50 to measure inequality of opportunity of health in European countries (Austria, France, Spain, Germany, Sweden, Poland, Belgium, Denmark, the Netherlands, Switzerland, Greece, Czech Republic and Italy). While such previous studies provide very useful international comparative data, the analysis is limited to European countries. In order to compare the level of inequality of opportunity across countries around the world, we need a survey data set that is standard among many and different countries.

The international student assessment such as the Programme for International Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS), etc. and its sub-questionnaire surveys are likely to be useful for this purpose. They could provide a useful means of measuring inequality in a country and comparing it across different countries. For example, Ferreira and Gignoux (2014) and Gamboa and Waltenberg (2012) use PISA data and Salehi-Isfahani et al. (2014) use TIMSS data. They measure the inequality of opportunity of educational achievement in each country. Ferreira and Gignoux (2014) and Gamboa and Waltenberg (2012), however, study only South American countries. Salehi-Isfahani et al. (2014) analyze only Middle Eastern and North African countries.

Using TIMSS data, we extend such previous studies by including a wider range of countries including Korea and Asian countries and a longer period from 1995 to 2015. Then we measure the inequality of opportunity of education in each country and examine their evolution over time.

6.3 Empirical Framework

6.3.1 Data

For empirical analysis we employ data from TIMSS, which measures educational achievement globally for a group of countries. TIMSS is an international educational performance assessment conducted under the auspice of the International Association for the Evaluation of Educational Achievement (IEA). TIMSS is conducted for students in the 4th grade (elementary school) and 8th grade (junior high school) who live in about 50 countries every four years from 1995 to 2015. For each study year the identical tests of mathematics and science subjects are given to relevant students to measure the level of student academic achievement. We use math and science test scores from the 8th grade and measure the inequality of educational achievement at each point for each country. In addition to student achievement scores, TIMSS collects educational environment data for students, schools, teachers and parents separately. It contains information about student environment which is important for our study of the inequality of opportunity. The number of countries available in TIMSS data is 34 countries in 1995, 34 countries in 1999, 45 countries in 2003, 47 countries in 2007, 42 countries in 2011 and 40 countries in 2015. We have access to data on both educational achievement and family background for a total of 74 countries across all survey years.

The TIMSS data set has merits for an international comparison for inequality of opportunity. Studies measuring economic inequality of income or wealth are based on heterogeneous data sets produced by different institutions in different countries. Therefore, results of such studies are limited in that the measurement criterion and the point of the survey are different from each other and the composition of the data is very heterogeneous. To the extent that TIMSS builds data based on the same test questions and questionnaire items from the same institution for each cycle, the data are fairly homogeneous for different countries. Given that the current study explores differences and changes in inequality of opportunity among different countries. TIMSS offers appropriate data for our purpose.

6.3.2 Calculation of the Inequality Index

In this study we use mathematics and science test scores of the students provided by TIMSS as an achievement variable. The variance (and standard deviation) of each subject score is used as an inequality index for each country and each year. In previous studies, various statistics such as variance, coefficient of variation, Gini coefficient, and mean log deviation were used as inequality indicators of achievement. In the current study, "variance" in particular is used as an inequality measure. The reasons for such a choice are given as follows.

TIMSS measures science and mathematics achievement of students around the world with the same test questions globally. Given that each country has different textbooks and curricula, however, the questionnaires that students in a country encounter differ from country to country. TIMSS converts test scores obtained from different countries into the same scale, using the Item Response Theory (IRT) model of education statistics, and calculates five scores for each subject. In the course of applying IRT, one can use several theoretical combinations of assumptions about percent correctness and statistical models of each test item. Thus, TIMSS provides five types of conversion scores as a common scale score for each subject. Each conversion score is once again standardized so that the global average is 500 points and the global standard deviation is 100 points.

A researcher may use the original conversion score with a common measure when she calculates the inequality of educational achievement in each country at a particular point in time. Alternatively, she could use a standardized score whose global average is 500 points. No matter which score one uses, the rank of inequality calculated between countries should not change significantly. When using Gini coefficients or mean log deviations as an inequality index, however, the ranking of inequality among countries depends on which score one uses. To avoid such a problem, Ferreira and Gignoux (2014) and Salehi-Isfahani et al. (2014) use "variance" as an inequality index. Following their decision, the current study also uses "variance" as an inequality index. In addition, "variance" makes it easier to break down total variance into inequality of opportunity and inequality of effort.

In this study, we use five kinds of environmental variables: (1) gender, (2) the number of books held by the household, (3) the education level of the father, (4) availability of computers in the home, and (5) the size of city in which the student resides. Gender is divided into two groups: (1) boys and (2) girls. The number of books in household is divided into three groups: (1) Less than 10 volumes, (2) 11 to 100 volumes, (3) More than 101 volumes. The education level of the father is divided into three groups: (1) completion of elementary and junior secondary education, (2) completion of secondary education and/or vocational education, and (3) admission to tertiary education. Having a computer is a dummy variable.

We apply a parametric estimation method proposed by Ferreira and Gignoux (2011). We calculate two indicators of inequality of opportunity, direct and indirect. The first way to calculate the direct indicator is as follows.

Define C_i as student i's environment vector consisting of the five environment variables selected above. C_i consists of one sex dummy variable, a variable that transforms the number of books into consecutive volumes, a variable that transforms the education of the father into education years, the presence of computers in the household, and the city scale dummy variables of a student's residential area. Suppose that student i's test score is generated by the following model:

$$y_i = C_i\beta + u_i. \quad (6.1)$$

Calculate an estimate $\hat{\beta}$ by applying OLS to Eq. (6.1), and the fitted value of the score predicted only by C_i (i.e., explained only by environmental differences) using the following Eq. (6.2).

$$\hat{y}_i = C_i \hat{\beta}. \tag{6.2}$$

The probability distribution of test scores constructed by $\hat{y}_i$ is called a smoothed distribution. The ratio of the following Eq. (6.3) is used as a direct indicator to express the degree of relative inequality of opportunity

$$\frac{Var\left(\hat{y}_i\right)}{Var\left(y_i\right)} \quad . \tag{6.3}$$

Second, the indirect inequality index proposed by Ferreira and Gignoux (2011) is calculated as follows.

Based on the estimated value $\hat{\beta}$ and the residual $\hat{u}_i$ obtained from Eq. (6.1), one can derive the following hypothetical test score:

$$\tilde{y}_i = \bar{C}_i \hat{\beta} + \hat{u}_i. \tag{6.4}$$

where $\bar{C}_i$ consists of the average value of each environment variable included in C_i.

$\tilde{y}_i$ can be interpreted as an estimate of the test score, which is explained only by the difference in effort in a hypothetical state where the influence of the environment is kept equal (at the average level) to all students. The probability distribution of the scores constructed by $\tilde{y}_i$ is called a standardized distribution.

The ratio of the following Eq. (6.5), calculated using the $\tilde{y}_i$ derived from the above process, is presented as an indirect indicator of relative inequality of opportunity

$$1 - \frac{Var\left(\tilde{y}_i\right)}{Var\left(y_i\right)} \tag{6.5}$$

Due to the space limitation of the paper, we explicitly present only the indirect indices of Eq. (6.5). The values of direct indicators are in general similar to those of indirect indicators. The estimates of inequality of opportunity (both direct and indirect) presented in this study measure the lower limit of the true inequality of opportunity in a country. The environment variables encompassed by vector C_i are only a subset of all the environmental variables that may influence inequality of opportunity. Since unobservable environmental variables are included in u_i, the estimates calculated in this analysis show smaller values than the true inequality of opportunity.

6.4 Estimation Results

Table 6.1 and Fig. 6.1 show the inequalities (standard deviations) of the math score for each country. Table 6.1 presents standard deviations of the mathematical score of a country at each survey year (1995–2015). Column (7) shows the average of the standard deviations at each survey year. For convenience of comparison, stan-

Table 6.1 Inequality (standard deviation) of math test scores

Country	Standard deviation of math test scores						
	1995	1999	2003	2007	2011	2015	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Algeria				56.7			56.7
Armenia			80.2	82.6	83.7		82.1
Australia	91.8	74.3	76.5	71.5	75.1	74.1	77.2
Austria	84						84.0
Bahrain			68.9	74.9	86.7	72.6	75.8
Bosnia				72.8			72.8
Botswana			65.7	69.3		77.8	70.9
Bulgaria		77.4	78	91.7			82.4
Canada	83.2	68.5				63.2	71.6
Chile		73.9	67.9		70.3	72.2	71.1
Colombia	63.7			69			66.3
Cyprus	80.8	73.5	74.5	79.6			77.1
Czech	88.4	74.7		64.1			75.7
Denmark	79.3						79.3
Egypt			83.9	95		91	90.0
El Salvador	86.1			66.8			76.5
England			70		74.8	69.4	71.4
Estonia			63				63.0
Finland		60.5			59.1		59.8
France	74.8						74.8
Georgia				91.4	96.9	82.8	90.4
Germany	81.7						81.7
Ghana			85.8	85.6	82		84.5
Greece	83.1						83.1
Hong Kong	94.8	69.7	66.6	86.6	76.6	71.2	77.6
Hungary	87.8	74.7	66.6	71.8	71.9	71.9	74.1
Iceland	71.1						71.1
Indonesia		97.4	87.8	82.3	78.4		86.5
Iran	57.1	74.9	66.1	75.9	79.4	84	72.9
Ireland	85.1					62.4	73.8
Israel	84.8	83.1	78.3	91.6	83.9	89.9	85.3
Italy		79.5	70.9	70.2	66.8	66.4	70.8
Japan			73.9	78.2	76.9	79.3	77.0
Jordan		94.2	79.8	93.2	89.1	85.2	88.3
Kazakhstan					75.5	91.2	83.3
Kuwait	54.3			73.4		81.5	69.7
Latvia	78.1	74.4	68.3				73.6

Table 6.1 (continued)

Country	Standard deviation of math test scores						
	1995	1999	2003	2007	2011	2015	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lebanon			61	68.3	67.2	71.2	66.9
Lithuania	77.9	66.2	73.3	69.7	68.9	68.5	70.8
Macedonia		80.6	76.6		93.9		83.7
Malaysia		75.5	69.6	71.3	84.9	78.6	76.0
Malta				84.5		77.1	80.8
Moldova		80	77.4				78.7
Mongolia				75.4			75.4
Morocco		90	66.3	75.9	80.4	73.7	77.3
Netherlands	80.3	65.3	61.1				68.9
New Zealand	84.9	79.7	68.7		76	77	77.3
Norway	80.8		64	58.5	57.2	62.8	64.7
Oman				82.9	91.2	91.5	88.5
Palestine			85.6	95.2	94.6		91.8
Philippines		87.1	79.9				83.5
Portugal	59.4						59.4
Qatar				86.3	102.1	89.2	92.5
Romania	84.3	87.3	80.4	87.1	86		85.0
Russia	87.4	79.7	71.6	76.8	75.2	77.9	78.1
Saudi Arabia			74.3	69.2	85.9	77.8	76.8
Servia			81.5	80.8			81.2
Singapore	90.7	74.7	72.6	81.8	75	73	78.0
Slovakia		68	71.3				69.7
Slovenia	83.2	76.9	65.2	65.3	63.1	61.8	69.2
South Korea	99.7	71.1	74.5	82.9	77.7	77	80.5
South Africa			88.9			77.9	83.4
Spain	70.2						70.2
Sweden	84.8		63	62	60.6	61.7	66.4
Switzerland	86.9						86.9
Syria			73	79.5	92.3		81.6
Taiwan		91.9	88.7	92	92.8	83.3	89.7
Thailand	81	77.9		83.7	78.8	84.1	81.1
Tunisia		60	57	62.5	70		62.4
Turkey		81.2			98.2	88.2	89.2
UAE					77.7	85.2	81.4
Ukraine				80.7	80.7		80.7
USA	84.8	73.6	69.6	67.2	67.6	74.6	72.9
Average	80.8	77.0	73.1	77.3	79.2	76.9	77.0

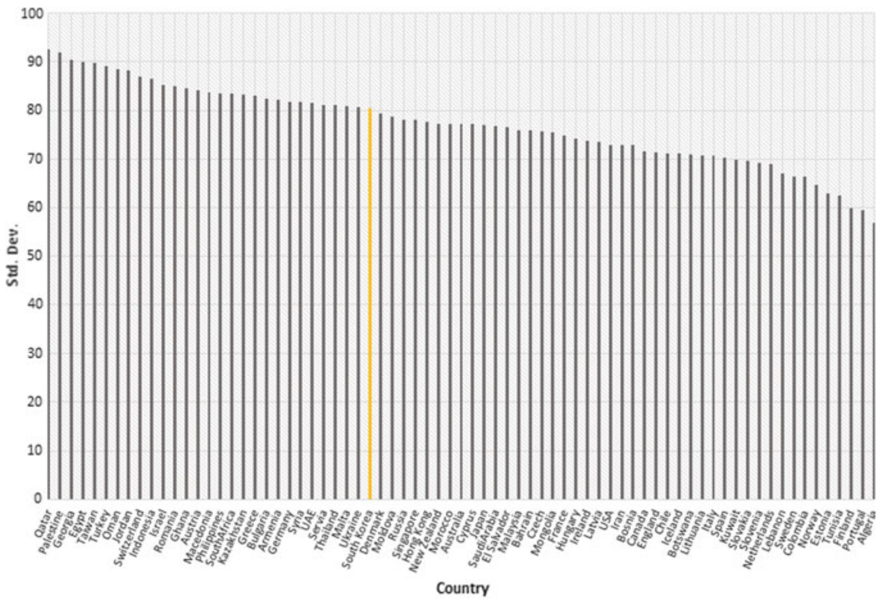


Fig. 6.1 Standard devation of math score: average for period 1995 to 2005

dard deviations are presented instead of variances as an inequality measure. Korea's standard deviation is usually between 71.1 and 82.9. It is located midway among the 73 countries under study. Korea's standard deviation was exceptionally high at 99.7 in 1995. According to Fig. 6.1, the average of standard deviations of math score of Korea from 1995 to 2015 is 80.5. This is the 29th highest of the 73 countries under study. In terms of inequality of achievement, Korea usually stands at a slightly higher level than the median.

Table 6.2 and Fig. 6.2 show the proportion of inequality of educational opportunity for each country. Inequality of educational opportunity of Korea is 14.6% in 1995, 18.2% in 1999, 19.4% in 2003, 19.3% in 2007, 16.1% in 2011 and 14.6% in 2015, respectively. Inequality of educational opportunity rose from the mid-1990s to the mid-2000s and then tended to decline since the mid-2000s. From 1995 to 2015, the average value of Korea's inequality of opportunity is 17.0%. In terms of the average of inequality of opportunity from 1995 to 2015, Korea's inequality of opportunity is the 13th highest among the 73 countries under study. However, for developing countries in the Middle East, Africa, and South America, the attendance rate and TIMSS participation rate for 8th grade students are significantly lower than those of other developed countries. Thus, inequality of educational opportunity in those countries will be significantly underestimated. In order to understand the true relative position of Korea, it is necessary to compare the indices among the group of largely homogeneous countries.

Table 6.2 Proportion of inequality of educational opportunity

Country	Proportion of inequality of educational opportunity (percent)						
	1995	1999	2003	2007	2011	2015	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Algeria				1.1			1.1
Armenia			5.8	3	9.1		6.0
Australia	10.8	10.6	15.1	20.5	21.1	18.7	16.1
Austria	9.9						9.9
Bahrain			11.9	10.9	19.3	7.1	12.3
Bosnia				9.2			9.2
Botswana			7.6	6		6.9	6.8
Bulgaria		16.7	12.2	16.8			15.2
Canada	2.4	6.1				12.9	7.1
Chile		22.4	30.5		20.9	16.9	22.6
Colombia	4.6			19			11.8
Cyprus	8.9	10.1	12.3	14			11.3
Czech	8	9.9		17.2			11.7
Denmark	4.5						4.5
Egypt			13	6.8		9.6	9.8
El Salvador	7.6			16.2			11.9
England			16.1		20.7	23.1	20.0
Estonia			14.4				14.4
Finland		11.8			9.2		10.5
France	4.5						4.5
Georgia				8	15.3	15.1	12.8
Germany	13.1						13.1
Ghana			13.1	6.4	6.1		8.6
Greece	9.3						9.3
Hong Kong	8.1	8.2	10.3	13.3	13.9	14.6	11.4
Hungary	15.1	26.4	25.8	24.4	28.7	40.6	26.8
Iceland	5.3						5.3
Indonesia		12.1	6.9	9.6	7.5		9.0
Iran	7.4	13.5	16.4	20	23.1	21.1	16.9
Ireland	9.6					23.8	16.7
Israel	15	19.7	12.7	12.6	20.3	21.6	17.0
Italy		8.2	13.1	11	10.7	18.2	12.2
Japan			13.4	16.7	14.3	17.5	15.5
Jordan		8.1	13.4	10.2	12.7	9.5	10.8
Kazakhstan					7.8	3.9	5.8
Kuwait	3.2			5.7		9.6	6.2
Latvia	5.8	13	9.3				9.4

Table 6.2 (continued)

Country	Proportion of inequality of educational opportunity (percent)						
	1995	1999	2003	2007	2011	2015	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lebanon			12.4	13.5	15.5	7.9	12.3
Lithuania	11.2	25	16.5	18.1	20.4	21.6	18.8
Macedonia		21.6	14.9		19.9		18.8
Malaysia		22.2	13.6	17.3	23.4	12.7	17.8
Malta						19.6	19.6
Moldova		4.5	5.5				5.0
Mongolia				13.7			13.7
Morocco		4.5	3.3	7.5	11.5	7.2	6.8
Netherlands	10.2	8.8	14.6				11.2
New Zealand	9.3	14.9	13.8	13.8	22.4	24.6	16.5
Norway	5.8		9.8	11.1	14.8	15	11.3
Oman				13.6	21.9	7.5	14.4
Palestine			8.3	8.2	8.9		8.4
Philippines		8.2	10.4				9.3
Portugal	8.8						8.8
Qatar				13.7	15.8	15	14.8
Romania	10.2	12.5	18.6	20.8	25.5		17.5
Russia	9.8	12.1	12.1	14.6	10.7	6.7	11.0
Saudi Arabia			8.9	10.5	6.8	8	8.6
Servia			16.5	15.4			16.0
Singapore	8.5	9.8	14.8	18.2	16.6	17.3	14.2
Slovakia		16.1	20.4				18.3
Slovenia	10.1	14	10.3	13.7	13.8	12.2	12.4
South Korea	14.6	18.2	19.4	19.3	16.1	14.6	17.0
South Africa			20.3			16.8	18.6
Spain	8.8						8.8
Sweden	9.2		13.8	12.2	10.8	17.6	12.7
Switzerland	6						6.0
Syria			6.1	4.7	3.2		4.7
Taiwan		18.5	19.9	20.6	20.1	24	20.6
Thailand	8.5	15		20.4	13	16.1	14.6
Tunisia		10.7	12.5	14.9	14.7		13.2
Turkey		5.6		23.1	23.3	23.3	18.8
UAE					10.6	7.3	9.0
Ukraine				16.1	17.3		16.7
USA	13.4	17.5	17.5	17.1	17.1	17.2	16.6
Average	8.8	13.4	13.5	13.6	15.6	15.5	12.4

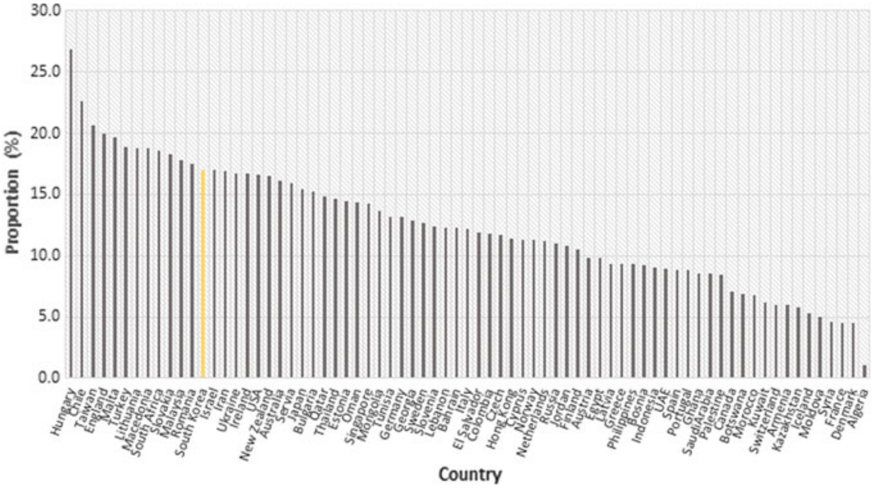


Fig. 6.2 Proportion of inequality of opportunity: average for period 1995 to 2015

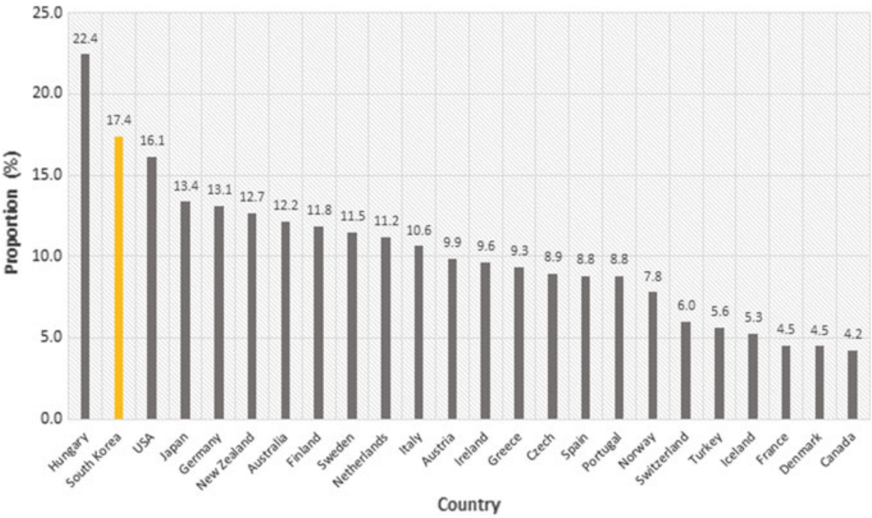


Fig. 6.3 Proportion of inequality of opportunity: average for period 1995–2003, OECD countries

Figures 6.3, 6.4 show the proportion of inequality of educational opportunity for OECD countries, a group of relatively homogeneous countries. Figures 6.3, 6.4 select 24 countries that have been OECD member countries since 1996 and have participated in TIMSS from 1995 to 2015. The list of the 24 countries is as follows: Australia, Austria, Canada, Czech Republic, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, South Korea, Spain, Sweden, Switzerland, Turkey, USA.

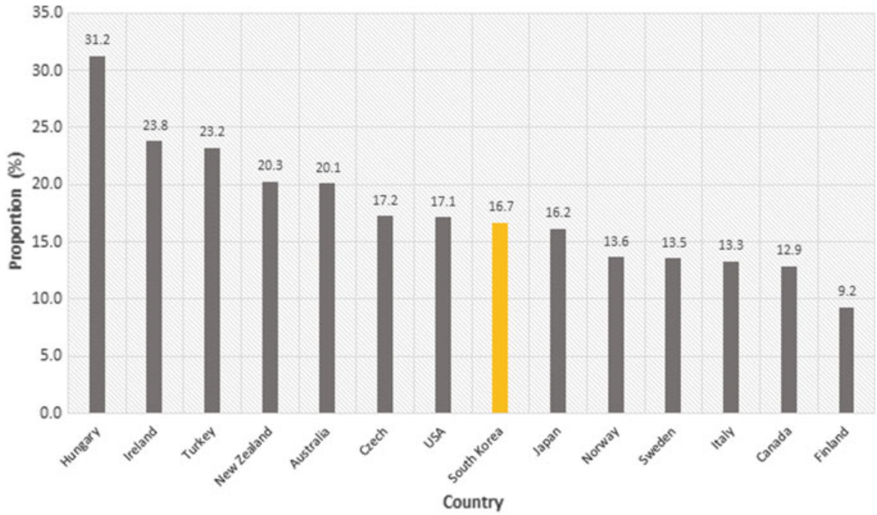


Fig. 6.4 Proportion of inequality of opportunity: average for period 2007–2015, OECD countries

Korea's average proportion of inequality of opportunity during period 1995 to 2003 was 17.4%. Among OECD countries participating in TIMSS, Korea stands at the second highest among the 24 OECD countries. As shown in Table 6.2, Korea's inequality of opportunity has eased since the late 2000s. Korea's average proportion of inequality of opportunity during period 2007 to 2015 is 16.7%, which stands at the eighth position among 14 OECD countries. Even if we confine a comparison to OECD countries, Korea's inequality of opportunity is quite high from 1995 to 2003. Since 2007, however, it has been low and has remained moderate.

Figures 6.3, 6.4 show the results of the comparison only among OECD countries which are in similar economic development stages. On the other hand, we could alternatively set up comparison countries based on the geographic location. Figures 6.5, 6.6, 6.7, 6.8 show the proportion of inequality of educational opportunity among countries in each broad region.

Figure 6.5 compares inequality of opportunity for East Asian countries (Hong Kong, Japan, Korea, Malaysia, Singapore, Thailand, Taiwan). Figure 6.6 compares it for Western European countries (UK, Finland, Italy, Netherlands, Norway, Turkey and Sweden). Figure 6.7 compares among North and Oceania countries (Canada, USA, Australia, New Zealand). Figure 6.8 compares among Middle Eastern and North African countries (Bahrain, Egypt, Iran, Israel, Jordan, Saudi Arabia and Tunisia). Figures 6.5, 6.6, 6.7, 6.8 show inequality of opportunity for countries participating in TIMSS more than three times between 1995 and 2015.

According to Fig. 6.5 which compares East Asian countries, Korea's level of inequality of educational opportunity (average of 17.0% of six survey years) is intermediate. Overall, Hong Kong has the lowest inequality of opportunity (11.4% on

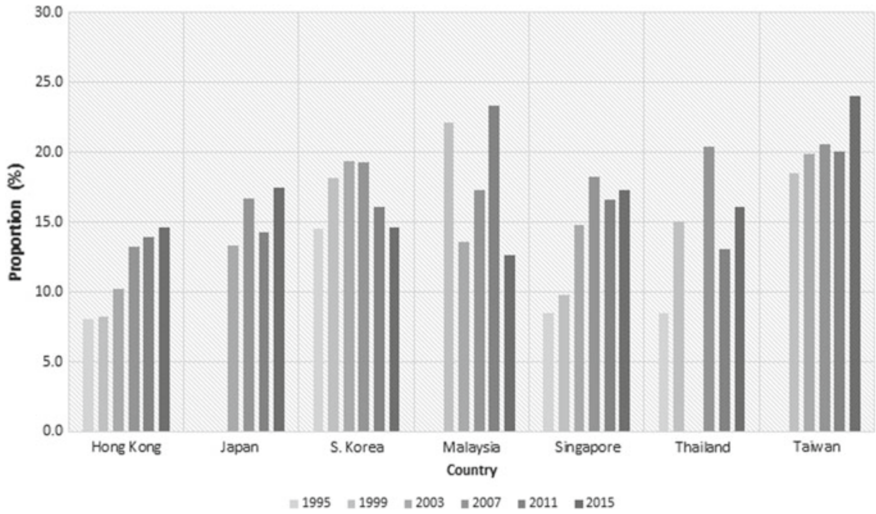


Fig. 6.5 Proportion of inequality of opportunity: East Asia

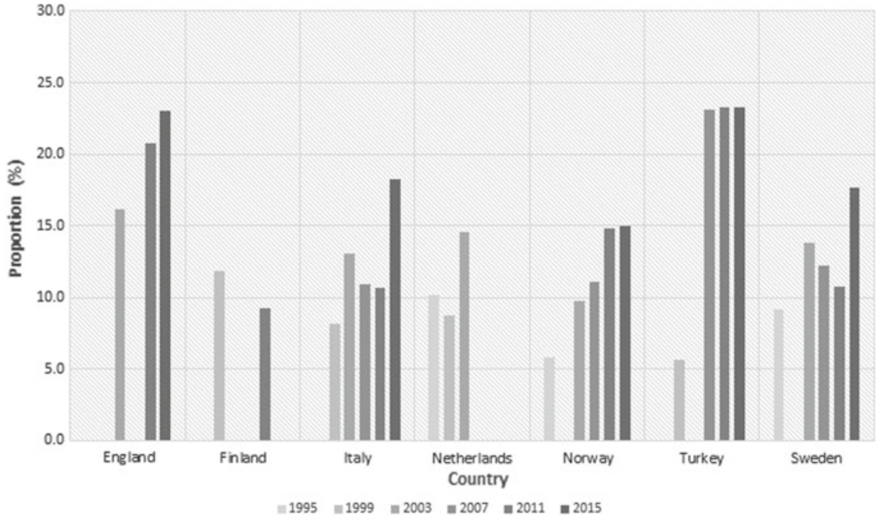


Fig. 6.6 Proportion of inequality of opportunity: W. Europe

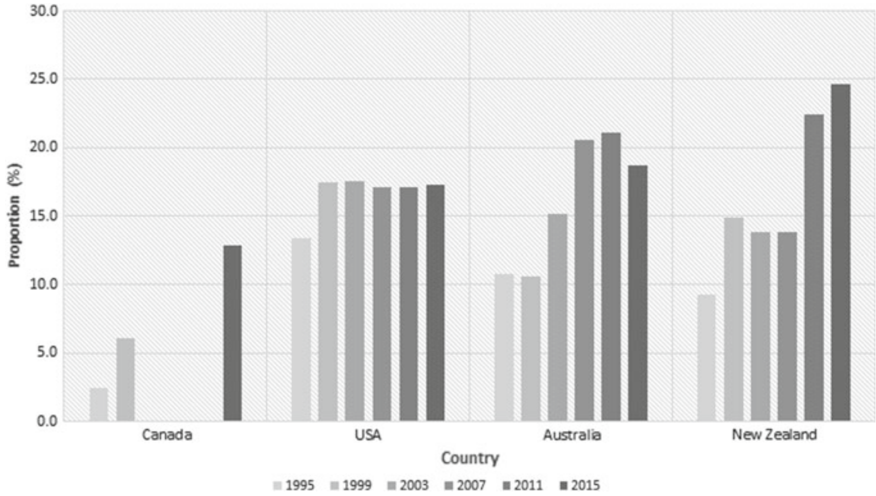


Fig. 6.7 Proportion of inequality of opportunity: North America and Oceania

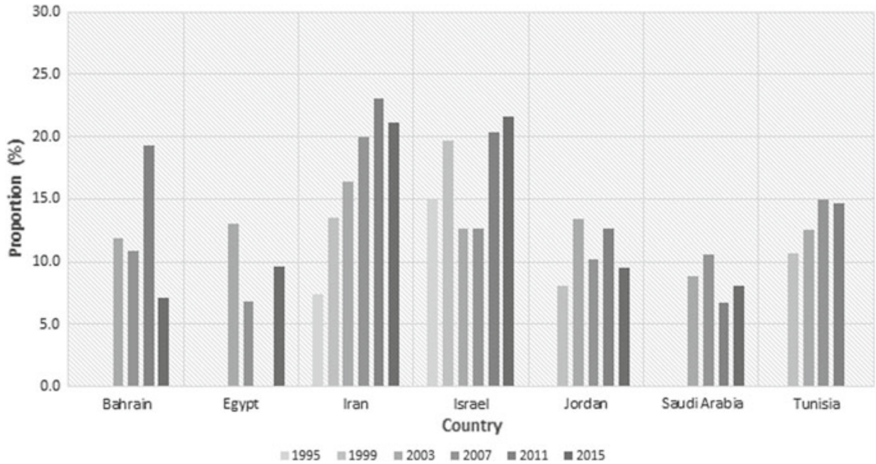


Fig. 6.8 Proportion of inequality of opportunity: Middle East and North Africa

average in six survey years). On the other hand, Taiwan's inequality of opportunity is the highest (20.6% on average for six survey years).

Figure 6.6 shows that western European countries also have variation in inequality of opportunity as in East Asia. Compared with Korea's inequality of opportunity, Norway, Finland (average of 10.5% in two survey years), Netherlands (average of 11.2% in three survey years), Norway (average of 11.3% in five survey years), Sweden (average of 12.7% in five survey years) is relatively low. Inequality in the UK (England, average of 20.2% in three survey years) and Turkey (average of 18.8% in four survey years) is higher than Korea's.

Figure 6.7 shows that among North American and Oceanian countries, inequality of educational opportunity in Canada (average of 7.1% in three survey years) is lower than in Korea. However, inequality of educational opportunity in the United States (average of 16.6% for six survey years), Australia (average of 16.1% for six survey years) and New Zealand (average of 16.5% for six survey years) is similar to Korea's.

Figure 6.8 suggests that among Middle East and North African countries, inequality of opportunity of Iran (average of 16.9% in six survey years) and Israel (average of 17.0% in six survey years) is similar to Korea's. However, inequality of opportunity in Bahrain (average of 12.3% for four survey years), Egypt (average of 9.8% for three survey years), Jordan (average of 10.8% for five survey years), Saudi Arabia (average of 8.6% for four survey years) and Tunisia (average of 13.2% for four survey years) is lower than for Korea's. Since the rate of secondary school enrollment in Egypt, Iran, Jordan, and Tunisia (especially the rate of female enrollment) is very low at some point in time, however, it is possible that inequality of educational opportunity in these countries is estimated to be much lower than the true value. According to education statistics in the World Development Indicators of the World Bank, between 1999 and 2015, the rate of secondary school enrollment in Egypt (77–82% of men, 68–81% of women), Iran (77–89% of men, 64–89% of women), and Tunisia (male 60–90%, female 54–94%) is considerably lower than that in Korea (90–101% for males and 90–101% for females). Such low enrollment rates cast doubt on validity of the estimated degree of inequality of educational opportunity in these countries.

6.5 Conclusion

In recent years, inequality issues have emerged as an important concern in our society. Inequality of opportunity of education is also being a subject of interest. Korea has a large private tutoring market outside public education. Such a society always faces a risk that inequality of opportunity of education can deteriorate. The public education sector has contributed to the equalization of educational opportunities through the traditional secondary school equalization policy in Korea. As private tutoring has been rampant recently, however, inequality of educational opportunity is likely to deteriorate gradually. Since inequality of educational opportunity is closely related to social mobility through education, an analysis of inequality of educational opportunity can contribute to a policy for enhancing social integration. Using TIMSS data, we estimate changes in inequality of educational opportunity of Korea between 1995 and 2015 and compare such inequality of Korea with other countries.

Overall, Korea's inequality of educational opportunity is found to stand at an intermediate position from the international perspective. From 1995 to 2015, Korea's average portion of inequality of educational opportunity is 17.0%, which is the 13th highest among the 73 countries under study. Compare with OECD countries participating in TIMSS, while Korea's average inequality of opportunity in 1995–2003 stands the fourth highest of the 28 OECD countries, that average tends to decline since the mid-2000s and stands at the tenth position among the 16 OECD countries.

Although Korea's inequality of educational opportunity currently stands at an intermediate position, it is likely to deteriorate over time, since the private tutoring sector has recently been expanding in Korea. Various policies need to be conceived so as to weaken the size of the private education sector and its influence on educational inequality. We hope our study can become a step forward to this end.

Acknowledgements This chapter heavily draws on Kang and Oh (2017), which is a part of a research project organized by the Korea Development Institute and the Korea Labor Institute in collaboration with the Korea National Research Council for Economics, Humanities and Social Sciences (Project 17-22-01).

References

Aaberge, Rolf, Magne Mogstad, and Vito Peragine. 2011. Measuring long-term inequality of opportunity. *Journal of Public Economics* 95 (3–4): 193–204.

Almås, Ingvild, Alexander W. Cappelen, Jo Thori Lind, Erik Ø. Sørensen, and Bertil Tungodden. 2011. Measuring unfair (in)equality. *Journal of Public Eonomics* 95: 488–499.

Arneson, R. 1989. Equality and equal opportunity of welfare. *Philosophical Studies* 56: 77–93.

Björklund, Anders, Markus Jäntti, and John E. Roemer. 2012. Equality of opportunity and the distribution of long-run income in Sweden. *Social Choice and Welfare* 39 (2–3): 675–696.

Bourguignon, François, Francisco H.G. Ferreira, and Marta Menéndez. 2007. Inequality of opportunity in Brazil. *Review of Income and Wealth* 53 (4): 585–618.

Bricard, Damien, Florence Jusot, Alain Trannoy, and Sandy Tubeuf. 2013. Inequality of opportunities in health and the principle of natural reward: Evidence from European countries. In *Health and inequality*, ed. Pedro Rosa Dias, and Owen O'Donnell, 335–370. Bingley, UK: Emerald Group Publishing.

Checchi, Daniele, and Vito Peragine. 2010. Inequality of opportunity in Italy. *The Journal of Economic Inequality* 8 (4): 429–450.

Cogneau, Denis, and Sandrine Mesple-Somps. 2008. Inequality of opportunity for income in five countries of Africa. In *Inequality and opportunity: Papers from the second ECINEQ society meeting*, ed. John Bishop, and Buhong Zheng, 99–128. Bingley, UK: Emerald Group Publishing.

Cohen, G.A. 1989. On the currency of egalitarian justice. *Ethics* 99: 906–944.

Dworkin, Ronald. 1981a. What is equality? part 1: equality of welfare. *Philosophy and Public Affairs* 10 (3): 185–246.

Dworkin, Ronald. 1981b. What Is equality? part 2: equality of resources. *Philosophy and Public Affairs* 10 (4): 283–345.

Ferreira, Francisco H.G., and Jérémie Gignoux. 2011. The measurement of inequality of opportunity: Theory and an application to Latin America. *Review of Income and Wealth* 57 (4): 622–657.

Ferreira, Francisco H. G., and Jérémie Gignoux. 2014. The measurement of educational inequality: achievement and opportunity. *The World Bank Economic Review* 28 (2): 210–246.

Ferreira, F.H., and V. Peragine. 2015. Equality of opportunity: Theory and evidence. Discussion Paper No. 8994, Institute for the Study of labor (IZA), Bonn.

Fleurbaey, Marc. 2008. *Fairness, responsibility, and welfare*. Oxford and New York: Oxford University Press.

Fleurbaey, Marc, and Francois Maniquet. 2011. *A theory of fairness and social welfare*. Cambridge and New York: Cambridge University Press.

Gamboa, Luis, and Fábio Waltenberg. 2012. Inequality of opportunity for educational achievement in Latin America: Evidence from PISA 2006–2009. *Economics of Education Review* 31 (5): 694–708.

Kang, Changhui, and Sungjae Oh. 2017. An international comparison of inequality of opportunity in education. In *Social integration and economic growth: Lowering inequality and new growth*, ed. Joohoon Kim, and Sungjae Cho, 100–121, National Research Council for Economics, Humanities and Social Sciences Collaboration Research Paper 17-22-01.

Kim, Joonil, Seo Hwan-Joo, and Woojin Shin. 2016. An analysis on the contribution of the inequality of opportunity to income inequality: Evidence from Korea in the 2000s. *Order Economics Journal* 19 (1): 41–67.

Kim, Woocheol, and Woojin Lee. 2009. Roemer's conception of equal opportunity and Korea's income tax-benefit policy. *Economy Analysis* 15 (3): 129–168.

Ko, Ja-Yee, and Woo Jin Lee. 2011. Father's education and son's achievement. *Korea Journal of Public Finance* 4 (2): 47–87.

Lefranc, Arnaud, Nicolas Pistolesi, and Alain Trannoy. 2008. Inequality of opportunities vs. inequality of outcomes: Are Western societies all alike? *Review of Income and Wealth* 54 (4): 513–46.

Ramos, Xavier, and Dirk Van de Gaer. 2016. Approaches to inequality of opportunity: Principles, measures, and evidence. *Journal of Economic Surveys* 30 (5): 855–83.

Roemer, J. 1993. A pragmatic theory of responsibility for the egalitarian planner. *Philosophy and Public Affairs* 10: 146–166.

Roemer, J. 1998. *Equality of opportunity*. Cambridge: Harvard University Press.

Roemer, J., and A. Trannoy. 2016. Equality of opportunity: Theory and measurement. *Journal of Economic Literature* 54 (4): 1288–1332.

Salehi-Isfahani, Djavad, Nadia Hassine, and Ragui Assaad. 2014. Equality of opportunity in educational achievement in the Middle East and North Africa. *Journal of Economic Inequality* 12 (4): 489–515.

Chapter 7
Political Economics of Income Inequality, Redistributive Policy, and Interregional Migration

Moriki Hosoe, Tadahisa Ohno and Akio Kawasaki

7.1 Introduction

This investigation reveals how regional redistributive policies in two regions and intra-regional migration are developed with income inequality in the framework of fiscal federation. Generally speaking, the main objectives of government are provision of public goods and services and easing of economic inequality among residents. We confine ourselves to discussion of economic inequalities and redistributive policy under a fiscal federation in order to focus on income disparity issue. Generally, fiscal federalism is said to promote competition for mobile resources with the division of policy responsibilities among different levels of government (see Waldasin 2008).

We consider political economic analysis of tax competition between two regions. Here, redistribution decisions are made by the majority group in regions from the point of their own interests. We designate such a society as a selfish-democracy society or populism society. In a mature society, governments are concerned not only about direct supporters but also to some extent the supporters of opposite parties. Nevertheless, we consider this naive and selfish society in these analyses. This selfishness might be one aspect of modern society despite advocacy of humanity and human rights.

For the analysis of regional economics, important and early discussion of redistribution in a decentralized system is found in a report of a study by Oates (1972), who emphasizes that mobility of households is likely to undermine attempts by

The original version of this chapter was revised: The author name "Akira Kawasaki" has been changed to "Akio Kawasaki". The correction to this chapter can be found at https://doi.org/10.1007/978-981-13-7036-6_22

M. Hosoe (✉)
Kyushu University (Emeritus Professor), Fukuoka, Japan

T. Ohno
Kumamoto University, Kumamoto, Japan

A. Kawasaki
Oita University, Oita, Japan

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_7

local governments to redistribute income. Positive analysis of redistribution policies undertaken by local governments under direct democracy is offered by Brown and Oates (1972). Orr (1976) emphasizes a concern by wealthy people for poor people as a factor of income redistribution policies. Epple and Romer (1991) examine how much redistribution occurs when only local governments can have tax/transfer instruments, and individuals can move freely among jurisdictions. Recently, Arawatari and Ono (2013) investigate intergenerational mobility through educational investment using a political economy redistribution model. Luca and Gulcin (2013) also analyze the implications of income inequality and mobility for demand for redistribution and social insurance under a three-period stochastic overlapping-generations model. But they do not mention about interregional movement.

From the social dynamics of the democracy, Acemoglu and Robinson (2006) investigate a regional analysis in the context of the dynamics of democratic society, where low-income group has a political power. In Acemoglu et al. (2015), it is shown that democracy does not lead to a uniform decline in post-tax inequality, but can result in changes in fiscal redistribution and economic structure that have ambiguous effects on inequality. But their analysis does not include the factor of migration. Related to regional analysis with income inequalities, Flamand (2015) studies the relative merits of centralized and decentralized redistribution in a political economy context assuming cross-regional heterogeneity in average income and identity. Her work includes the assumption that a wealthy group has policy-making power. These papers are very interesting. We will apply these discussions to the context of interregional movement and regional policy.

In this paper, there are two income groups, high and low, in two regions and the low income group is a majority in each region. They live three periods. The first period is childhood. By ignoring the care cost of child, we can restrain our discussion to two-period analysis. In the second and last period, people in each group earn a constant and same income. People in the beginning of the second period vote for the power of local governments looking for representatives of their interests. Since low income group is a majority, it takes the power of the local government in each region. The local government has the remaining two periods horizon for the generation. In the selfish-democracy society, low-income group tries to make policy decision in pursuit of self-profit. We assume that while high-income people can move between regions, low-income people cannot move because of the high cost of migration. Therefore, if the local policy in one region is harsh for high-income group, high-income people may migrate to the other region. Then the low-income group must make decisions, taking the migration into consideration. The unique way for the high-income group to resist is migration. This is a concept of voice and exit in Hirschman (1970). We make this assumption in order to emphasize the selfish democracy society or the populism society.

We show characteristics of two-period policy, depending on the population rates of the two income groups in two regions and the dynamics of migration of high-income group people. In a stable long-run equilibrium, the total amounts of post-tax income of low-income group in respective periods in the two regions are the same. We also find an unstable two-period cycle equilibrium where there is an extreme population disparity of high-income group between two regions. Lastly, when there is more extreme population disparity of high-income group between two regions, explosive cycle equilibrium of migration equilibrium is shown to occur as a kind of corner solution,

Our paper is organized as follows. Section 7.2 presents a model of tax competition between two regions. Section 7.3 presents analyses of the two-period policy of two regions in the shadow of migration. Section 7.4 presents an investigation of the dynamics of migration in short-run and long-run. Section 6 concludes this paper.

7.2 Model

Presuming two regions A, B, with respective high-income and low-income groups, superscripts H and L respectively denote the high-income and low-income groups. People live three periods. People in t-generation are born in period t-1, earn a constant income in period t and period $t + 1$, and have their children in period t. The high-income group and low-income group have respective constant incomes y^H and y^L for two periods, common to the two regions. The high-income population in region i is denoted as θ_i^H. For simplicity of discussion, we neglect their care activity of childhoods. So we can analyze their behaviors in two-period model. Our model is a special case of overlapping generation model. Only high-income people in t-generation are assumed to be able to migrate between two regions at the beginning of period $t + 1$. They move with their children to a region. Low-income people are assumed to be not able to migrate due to high migration cost. The low-income population in each region is denoted as θ_i^L $(i = A, B)$. In the framework of the fiscal federation, each region has the authority to independently implement the redistributive policy with taxes and income transfers.

We assume that the lower-income group population is larger than the higher-income group population in each region ($\theta_i^L > \theta_i^H$ $(i = A, B)$). At period t, only t-generation people are assumed to participate in the political decision. Then the low-income group has political power related to tax policy and redistributive policy in each region as a majority group. In order to emphasize the situation of a selfish democracy, we assume that the low-income group takes a discriminatively redistributive policy. In particular, the low-income group levies an income tax only on the high-income group as a discriminative tax policy and receives the total amount of tax as a discriminative transfer policy. We denote a tax rate on each income in region i as τ_i. Then tax collection costs are assumed to be incurred. We represent the tax collection cost for one dollar tax as a function of the tax rate on income. This function is specialized by $C(\tau_i) = \frac{c}{2}\tau_i^2$, where c is a positive constant. The total tax collection cost is assumed to burden the low-income group in each region. Then, for region i $(i = A, B)$ the utility in the period t of the representative of each income group of t-generation is the level of consumption itself. Then two utilities of the period t are described as follows.

$$U_i^H(t) = U(c_i^H) = c_i^H = y^H - \tau_i y^H \tag{7.1}$$

$$U_i^L(t) = U(c_i^L) = c_i^L = y^L + (\tau_i - C(\tau_i))\frac{y^H\theta_i^H}{\theta_i^L} \tag{7.2}$$

The second term in the right-hand side of the utility Eq. (7.2) stands for the net transfer of total tax of high income measured by the low-income group per capita.

From the definition, the difference of utility for the high-income people between the two regions depends on the difference of the tax rate levied in the two regions. The region with a higher tax rate on the high-income group has a higher first-period utility of the low-income group.

The low-income group of t-generation, as the power-holder, considers the maximization of the two-period utility $W^L(t)$ of the low-income group in each region, starting from period t. Then, they must consider the possibility of the migration of people between two regions.

$$\begin{aligned} W_i^L(t) = U_i^L(t)+\delta U_i^L(t+1) &= y^L + (\tau_i(t) - C(\tau_i(t)))\frac{y^H\theta_i^H(t)}{\theta_i^L} \\ &+ \delta\left(y^L + (\tau_i(t+1) - C(\tau_i(t+1)))\frac{y^H\theta_i^H(t+1)}{\theta_i^L}\right), \end{aligned} \tag{7.3}$$

where $\tau_i(t)$ is a tax rate in region i at period t. $\theta_i^L(t)$ and $\theta_i^H(t)$ is the population of low-income group and high-income group in region i at period t, respectively and δ is a discount factor.

Migration Equation

The migration of t-generation occurs in the beginning of period $t+1$. The ability of individuals to move from one region to another one generally will depend on the amount of income that they receive after implementing the redistributive policy. For our migration equation for t-generation, we assume that people make a myopic forecast of the future. Namely, the migration of high-income people occurs from the region with lower utility in period t for high-income group to the region with higher utility in period t for high-income group. Therefore, when $U_i^H(t) > U_j^H(t)$, some people of the high-income group in region j will move to region i in the beginning of period $t+1$. When $U_i^H(t) > U_j^H(t)$, we specify a migration equation of t-generation as follows.

$$\theta_i^H(t+1) = \theta_i^H(t) + b\frac{(U_i^H(t) - U_j^H(t))}{U(y^H)}\theta_j^H(t) \tag{7.4}$$

where b is a positive coefficient of migration and $U(y^H)$ is denoted as y^H. It is noteworthy that in the two migration dynamics, the migration scale depends on the population base of the out-flow region. Here,

$$\frac{(U_i^H(t) - U_j^H(t))}{U(y^H)}$$

is introduced to cancel the income effect of migration. On the other hand, when $U_i^H(t) < U_j^H(t)$, some people of the high-income group in region i will move to region j under a myopic forecasting. Therefore, when $U_i^H(t) < U_j^H(t)$, the following migration equation is described.

$$\theta_i^H(t+1) = \theta_i^H(t) - b\frac{(U_j^H(t) - U_i^H(t))}{U(y^{H)}}\theta_i^H(t)$$
$$= \theta_i^H(t) - b(\tau_i(t) - \tau_j(t)\theta_i^H(t) \quad (7.5)$$

Therefore, the difference in tax rates in two regions determines the migration rate in out-flow region. Migration coefficient b represents the sensitivity of the influence of the difference of tax rates on migration ratio. We assume that $b < 1$, which means that the sensitivity is low.

When $U_i^H(t) > U_j^H(t)$, substituting (7.1) for (7.4), we obtain the following equation.

$$\theta_i^H(t+1) = \theta_i^H(t) + b(-\tau_i(t) + \tau_j(t))\theta_j^H(t) \quad (7.6)$$

One-Period Policy

As a reference point, we consider the case of $\delta = 0$ in (7.3), which means that the low-income group has the time horizon of only one period. Then the following problem must be resolved by the government of the low-income group.

$$Max_{\tau_i} \ \ U(c_i^L)$$

We denote the tax policy under the one-period horizon as one-period policy. Then the optimal tax rate for the low-income group satisfies the following first-order condition.

$$1 = C'(\tau_i)$$

Therefore the optimal tax rate is

$$\tau_i^o = \frac{1}{c}$$

This implies that the optimal tax should be levied so as to make the marginal revenue of tax equal to marginal cost of tax. We assume that c>1 in order to have an inner solution. This tax rate is called the optimal one-period tax rate. We note that it is an efficient tax rate in terms of the tax collection cost. The optimal tax rate decreases with the tax collection cost per capita. From our definition of utility of high-income group, in this case, the utility is common to two regions. Considering the migration equation, we have the following obvious lemma:

Lemma 1 *Under one-period policy, there is no migration of high-income group between two regions.*

Then, from the definition of utility of low-income group, we have

$$U_i^L = (\tau^o - C(\tau^o))y^H \frac{\theta_i^H}{\theta_i^L} = \frac{1}{2c} y^H \frac{\theta_i^H}{\theta_i^L}$$

Therefore Lemma 2 is held.

Lemma 2 *Under one-period policy, the utility of low-income group in each region increases as the population of high-income group increases.*

This is why the population of high-income group implies the increase of the tax base. Therefore, from the point of view of the low-income group, the in-flow of high-income group is desirable.

7.3 Two-Period Policy Under the Shadow of Migration

Back to the two-period horizon of t-generation, the low-income group wants to increase the tax rate for the high-income group, but it might cause some people of the high-income group to migrate to the other region in the period $t + 1$, which might then decrease the tax base in the period $t + 1$. Let us first begin to analyze the case of migration from region B to region A. In the situation of the migration from region B to region A, $U_A^H(t) > U_B^H(t)$ holds.

Then, the low-income group of t-generation in region A resolves the following problem:

$$Max_{\tau_A(t),\tau_A(t+1)} \;\; W_A^L(t)$$

$$s.t. \;\; \theta_A^H(t+1) = \theta_A^H(t) + b(-\tau_A(t) + \tau_B(t))\theta_B^H(t).$$

Then we have two first-order conditions. From the first-order condition for $\tau_A(t + 1)$, we obtain

$$\frac{\partial W_A^L(t)}{\partial \tau_A(t+1)} = 1 - C'(\tau_A(t+1)) = 0$$

Therefore, the optimal tax rate $\underline{\tau}_A(t + 1)$ on high-income group of t-generation in region A is obtained.

$$\underline{\tau}_A(t+1) = \frac{1}{c} \tag{7.7}$$

We assume $c > 1$ to satisfy the inner solution condition of the tax rate. Then, the per capita net transfer for low-income group in region A is

$$\underline{\tau}_A(t+1) - C(\underline{\tau}_A(t+1)) = \frac{1}{2c}$$

The first-order condition for $\tau_A(t)$ is obtained as follows.

$$\frac{\partial W_A^L(t)}{\partial \tau_A(t)} = (1 - C'(\tau_A(t)))\frac{y^H \theta_A^H(t)}{\theta_A^L} + \frac{\delta \theta_B^H(t) b(-1) y^H}{2c\theta_A^L} = 0$$

From this, we have the following optimal tax rate $\underline{\tau}_A(t)$ on high-income group in region A.

$$\underline{\tau}_A(t) = \left(1 - \frac{\delta b \theta_B^H(t)}{2c\theta_A^H(t)}\right)\frac{1}{c} \tag{7.8}$$

Therefore, we understand that the optimal tax rate in period t in the in-flow region under the shadow of migration is smaller than the efficient tax rate in terms of tax-collection cost. Then if

$$\frac{\theta_A^H(t)}{\theta_B^H(t)} < \frac{\delta b}{2c}, \tag{7.9}$$

this policy is an subsidy policy. This implies that only providing a subsidy induces the in-flow of migration, when the population of high-income group is considerably small.

For this optimal tax rate in period t, the following properties are obviously held.

$$\frac{\partial \underline{\tau}_A(t)}{\partial \theta_A^H(t)} > 0, \frac{\partial \underline{\tau}_A(t)}{\partial \theta_B^H(t)} < 0,$$

The increase of population of high-income group in region A has the incentive for the low-income group in region A to raise the tax rate. On the other hand, the increase of population of high-income in region B has the incentive for the low-income group in region A to increase the migration by decreasing the tax.

$$\frac{\partial \underline{\tau}_A(t)}{\partial \delta b} < 0$$

The increase of the discount factor and migration coefficient give the incentive for the low-income group in region A to decrease the tax in order to increase the migration. The increase of marginal cost of tax collection obviously makes it difficult to raise the tax.

The partial derivative of $\underline{\tau}_A$ regarding c is

$$\frac{\partial \underline{\tau}_A(t)}{\partial c} = \frac{1}{c^2}\left(-1 + \frac{\delta b \theta_B^H}{2c\theta_A^H}\right)$$

From (7.8), when the optimal tax is positive (negative), this partial derivative is positive (negative).[1]

From the above discussion, the following proposition is held.

[1]We will thank referee 1 for pointing the possibility of subsidy policy out. By incorporating the possibility of subsidy policy explicitly, we were able to make the analysis of dynamics of migration clearer.

Proposition 1 *In the case of in-flow migration, the optimal tax policy in period t has the following properties.*

(1)The optimal tax rate in in-flow region is smaller than the efficient tax rate in terms of tax-collection.

(2)The optimal tax rate in in-flow region increases with the population of high-income group in in-flow region and decreases with that of the level of high-income in out-flow region.

(3)The optimal tax rate in in-flow region decreases with the migration coefficient and the discount factor.

(4) When the optimal tax is positive (negative), the optimal tax rate in in-flow region decreases (increases) with the tax-collection coefficient.

On the other hand, we consider the problem of the low-income group in region B in the situation of the migration from region B to region A. Then the migration equation for region B is

$$\theta_B^H(t+1) = \theta_B^H(t) - b\frac{(U_A^H(t) - U_B^H(t))}{U(y^H)}\theta_B^H(t)$$

$$= \theta_B^H(t) - b(\tau_B(t) - \tau_A(t))\theta_B^H(t)$$

Then two-period utility of the low-income group of t-generation in region B is

$$W_B^L(t) = U_B^L(t) + \delta(U_B^L(t+1)) = y^L + (\tau_B(t) - C(\tau_B(t)))\frac{y^H\theta_B^H(t)}{\theta_B^L)}$$
$$+ \delta\left(y^L + (\tau_B(t+1) - C(\tau_B(t+1)))\frac{y^H\theta_B^H(t+1)}{\theta_B^L)}\right) \quad (7.10)$$

The first-order condition for $\tau_B(t+1)$ in order to maximize $W_B^L(t)$ is the same as that for $\tau_A(t+1)$. Therefore, the optimal second-period tax rate in the out-flow region is also smaller than the efficient rate in terms of tax-collection cost.

On the other hand, the first-order condition for $\tau_B(t)$ is shown as follows.

$$\frac{\partial W_B^L(t)}{\partial \tau_B(t)} = (1 - C'(\tau_B(t)))\frac{y^H\theta_B^H(t)}{\theta_B^L} + \frac{\delta\theta_B^H(t)b(-1)y^H}{2c\theta_B^L} = 0$$

From this first-order condition,

$$(1 - c\tau_B(t)) - \frac{\delta b}{2c} = 0$$

Therefore we have the optimal tax rate in period t for region B.

$$\bar{\tau}_B(t) = \left(1 - \frac{\delta b}{2c}\right)\frac{1}{c} \quad (7.11)$$

This is positive because it is assume that b<1 and c>1. Let us make a comparison with optimal tax rates for in-flow region and out-flow region. Obviously the optimal tax rate for the in-flow region A is influenced by the population rate between the high-income group in two regions and the optimal tax rate for the out-flow region B is not influenced by the population of the high-income group in the out-flow region. Let us summarize the above discussion.

Proposition 2 *(1) The optimal tax rate in period* $t+1$ *on the high-income group of t-generation in the out-flow region is also smaller than the efficient rate in terms of tax-collection cost.*

(2)While the optimal tax rate in period t on the high-income group of t-generation in the in-flow region is influenced by the population rate between the high-income group in two regions, the optimal tax rate in the out-flow region is not influenced by the population rate of the high-income group.

Condition of Migration

Here, since we assume that region A is an in-flow region and region B is an out-flow region, we have to confirm whether the in-flow condition is satisfied ex-post. This means that our policies are Nash equilibrium ones. This condition is $U_A^H(t) > U_B^H(t)$.

$$U_A^H(t) - U_B^H(t) = (-\underline{\tau}_A + \bar{\tau}_B)y^H$$

$$= \left(-\left(1 - \frac{\delta b \theta_B^H(t)}{2c\theta_A^H(t)}\right)\frac{1}{c} + \left(1 - \frac{\delta b}{2c}\right)\frac{1}{c}\right)y^H$$

$$= \left(\frac{\theta_B^H(t)}{\theta_A^H} - 1\right)\frac{\delta b}{2c^2}y^H > 0$$

Therefore, the condition that region A is an in-flow region is shown to be $\theta_A^H(t) < \theta_B^H(t)$. That is,

Proposition 3 *The condition that region A is an in-flow region of t-generation in Nash equilibrium is equivalent to satisfy the following inequality:*

$$\theta_A^H(t) < \theta_B^H(t) \tag{7.12}$$

7.4 Dynamics of Migration

From the migration Eq. (7.6), t-generation people can start migration with their children in the beginning of period $t+1$. To consider a dynamics of migration over generations, we assume that people of t-generation in each region area each have only one child in period t for the simplicity of description. We also assume that the child of a parent having low income (high income) earns low income (high income) by becoming adult. This is to emphasize the durability of the income disparity. Then, the high-income group of $(t+1)$-generation have a population composition $(\theta_A^H(t+1), \theta_B^H(t+1))$ in period $t+1$, which is equal to the population composition of t-generation of high-income groups in period $t+1$ after migration. Since the population of low-income group is fixed, the discussion about population dynamics for the low-income group is neglected. At period $t+1$, the local governments for new generation $t+1$ decide two-period optimal policies. Then the wave of new migration occurs at the period $t+2$ and the new high-income population composition is determined. Then the local governments for new $(t+2)$-generation in two regions decide their optimal policies. And so on. This shows a dynamics of migration over generations between two regions. In the following, we investigate the properties on this dynamics of migration between two regions. Note that this dynamics of population can be said to be a special case of overlapping generation model.

For the simplicity of discussion, we normalize the total population in high-income group as $1 = \theta_A^H(t) + \theta_B^H(t)$ without the loss of generality. Therefore, when two regions have the equal population in high-income group, $\theta_A^H(t) = \theta_B^H(t) = 1/2$. Here, let us assume that $\theta_A^H(t) < \theta_B^H(t)$. Then the region A has the inflow of migration as shown in the previous section.

Then the dynamics of migration in region A with in-flow migration is described as follows.

$$\theta_A^H(t+1) = \theta_A^H(t) + b(-\tau_A(t) + \tau_B(t))\theta_B^H(t)$$

$$= \theta_A^H(t) + \left(\frac{\theta_B^H(t)}{\theta_A^H(t)} - 1\right)\frac{\delta b^2}{2c^2}\theta_B^H(t)$$

This is described as an autonomous dynamics of population θ_A^H in region A because the total population of high-income group is fixed as unity.

$$\theta_A^H(t+1) = \theta_A^H(t) + \left(\frac{1 - 2\theta_A^H(t)}{\theta_A^H(t)}\right)\frac{\delta b^2}{2c^2}(1 - \theta_A^H(t)) \equiv F(\theta_A^H) \qquad (7.13)$$

The function F is easily shown to be a convex function of θ_A^H with $F(1/2) = 1/2$ and $F(1) = 1$. Since the partial derivative of F with respect to θ_A^H is

$$\frac{dF}{d\theta_A^H(t)} = 2\frac{\delta b^2}{2c^2} + 1 - \frac{\frac{\delta b^2}{2c^2}}{\theta_A^H(t)^2}$$

Therefore, from $\frac{dF}{d\theta_A^H} = 0$, the value $\bar{\theta}_A^H$ of the minimal point is

$$\bar{\theta}_A^H = \sqrt{\frac{\frac{\delta b^2}{2c^2}}{2\frac{\delta b^2}{2c^2} + 1}} \qquad (7.14)$$

From this, we can show that if $\frac{\delta b^2}{c^2} \le 1$, $\bar{\theta}_A^H \le \frac{1}{2}$. Considering that $\delta < 1$, $c > 1$, and $b < 1$, the inequality is obviously satisfied. Therefore, $\bar{\theta}_A^H \le \frac{1}{2}$ is held.

We can also define $\theta_A^H = \theta_A^{1H} (\neq 1/2)$ satisfying $F(\theta_A^H) = 1/2$. Then

$$\frac{1}{2} = \theta_A^H(t) + \left(\frac{1 - 2\theta_A^H(t)}{\theta_A^H}\right) \frac{\delta b^2}{2c^2}(1 - \theta_A^H(t))$$

Rearranging this equation, we have

$$\left(\left(2\frac{\delta b^2}{2c^2} + 1\right)\theta_A^H - 2\frac{\delta b^2}{2c^2}\right)\left(\theta_A^H - \frac{1}{2}\right) = 0$$

From this, $\theta_A^H (\neq 1/2)$ satisfying $F(\theta_A^H) = 1/2$ is

$$\theta_A^{1H} = \frac{2\frac{\delta b^2}{2c^2}}{2\frac{\delta b^2}{2c^2} + 1} \tag{7.15}$$

Lastly, we seek the initial population $\theta_A^{H0} (= 1)$ of high-income group satisfying $F(\theta_A^{H0}) = 1$. Then we can easily obtain

$$\theta_A^{H0} = \frac{\frac{\delta b^2}{2c^2}}{(2\frac{\delta b^2}{2c^2} + 1)} \tag{7.16}$$

Therefore, if $\theta_A^H(t) < \theta_A^{H0}$, $\theta_A^H(t+1) = 1$ is held.

Then as seen in Fig. 7.1, the following proposition is held.

Proposition 4 *(1)The $(t+1)$-period population of high-income group in region A decreases as the t-period population of high-income group in region A increases, when it is smaller than $\bar{\theta}_A^{1H}$,*

(2)The $(t+1)$-period population of high-income group in region A increases as the t-period population of high-income group in region A increases when it is larger than $\bar{\theta}_A^{1H}$.

In Fig. 7.1, if the initial population of region A satisfies $\theta_A^H \in (\theta_A^{1H}, 1/2)$, then the dynamics of population in region A stably converges to 1/2. If $\theta_A^H \le \theta_A^{1H}$, then the dynamics of population in region A might change drastically from phase of in-flow to that of out-flow. To clarify this possibility, we must describe the dynamics of population of region A with out-flow of migration.

For out-flow of migration in region A, the dynamics of population in region A is shown in (7.5). From (7.5), we obtain

$$\theta_A^H(t+1) = \theta_A^H(t) - \left(\frac{\theta_A^H(t)}{\theta_B^H} - 1\right)\frac{\delta b^2}{2c^2}\theta_A^H(t).$$

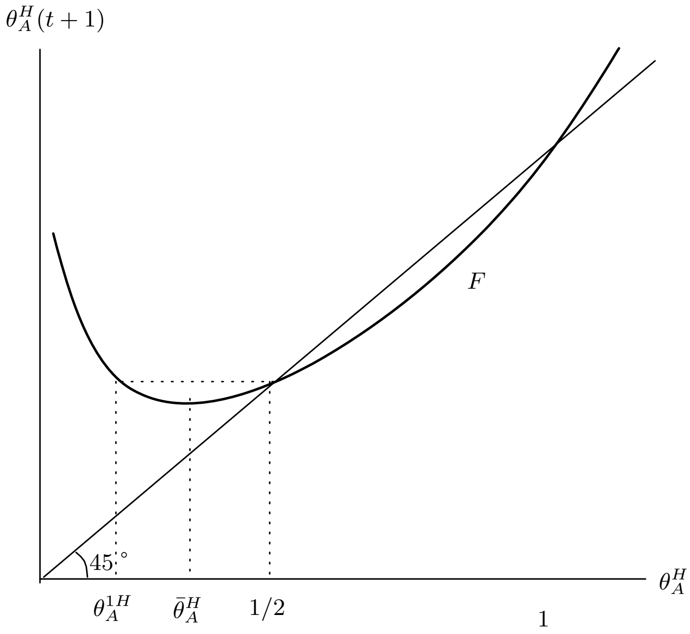


Fig. 7.1 A Basic Dynamics of Population in Region *A*

Therefore, the dynamics of population in region A with out-flow of migration is described as a following autonomous differential equation.

$$\theta_A^H(t+1) = \theta_A^H(t) - \left(\frac{2\theta_A^H(t) - 1}{1 - \theta_A^H}\right)\frac{\delta b^2}{2c^2}\theta_A^H(t) \equiv G(\theta_A^H) \tag{7.17}$$

Then the curve of G is a concave function of $\theta_A^H(t)$ with $G(1/2) = 1/2$. The value of the maximal point is greater than one-half.

If initial population θ_A^H is within an interval $(1/2, \theta_A^{2H})$ in Fig. 7.2, then the dynamics of population in region A might converge stably to equal population 1/2. If the initial population of region A satisfies $\theta_A^H \in (\theta_A^{2H}, 1)$, then the dynamics of population in region A changes dramatically from phase of out-flow to that of in-flow.

From Figs. 7.1 and 7.2, when the high-income group in a region is larger than a level of the high-income population in region A in period t and when it is sufficiently small, a big migration from region B occurs. In turn, the big out-flow from region A occurs and then region A become a minority in high group population.

To investigate the dynamics of migration in detail, we introduce the two-period migration function $\theta_A^H(t+2) = F(G(\theta_A^H(t)))$, by composing two migration Eqs. (7.13) and (7.17).

This function represents the high-income population in period t+2 in region A as a function of that in period t, which is depicted in Fig. 7.3. From Fig. 7.3, we have two

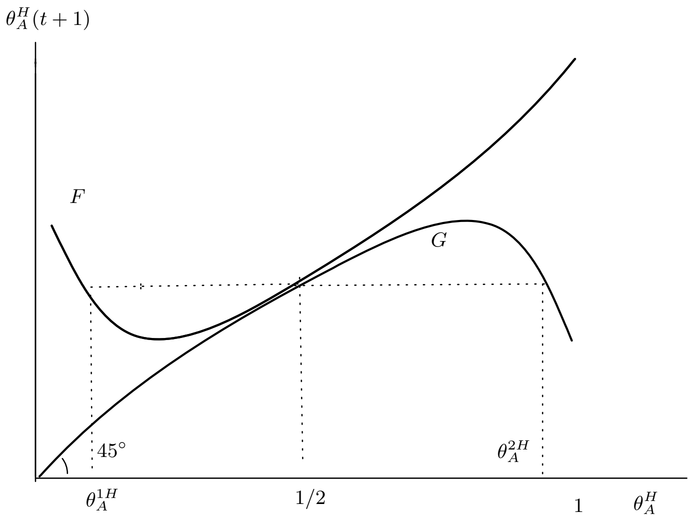


Fig. 7.2 An In-flow and Out-flow Dynamics of Population in Region A

equilibria where $\theta_A^H = G(F(\theta_A^H)$ is satisfied. This equilibrium is shown as follows.

$$\theta_A^{H*} = \frac{1}{2}, \theta_A^{H**} = \frac{\frac{\delta b^2}{2c^2}}{\frac{\delta b^2}{2c^2} + 1}$$

The first equilibrium is trivial. Two regions have equal-sized population of high-income group. In each period, this situation continues. This is called equal-population equilibrium of high-income group. Obviously this equilibrium is stable. Second equilibrium is shown at the point T where the 45 line and the curve $(F(\theta_A^H))$ intersects in Fig. 7.3. The population of high-income group in region A corresponding to the point T is denoted as θ_A^{H**}. This equilibrium implies that when the population of high-income group in region A in period t is θ_A^{H**}, θ_A^{H**} occurs again after big out-flow occurs in the next period. Therefore this equilibrium is a kind of cyclical equilibrium. Let us name it two-period cycle equilibrium. But this equilibrium is unstable. Then if the initial population of high-income population in region A is larger than θ_A^{H**}, high-income population in region A converges to 1/2, cyclically or not. If the initial population of high-income population in region A is smaller than θ_A^{H**}, the population of high-income population in region A repeats big out-flow and big in-flow. Once the high-income population in region A becomes smaller than θ_A^{H0}, all high-income population concentrates in region B in the next period. Furthermore, all high-income population in turn concentrates in region A in the next period. Therefore this extreme-concentration cycle occurs. This third equilibrium can be said to be an extreme-cycle equilibrium.

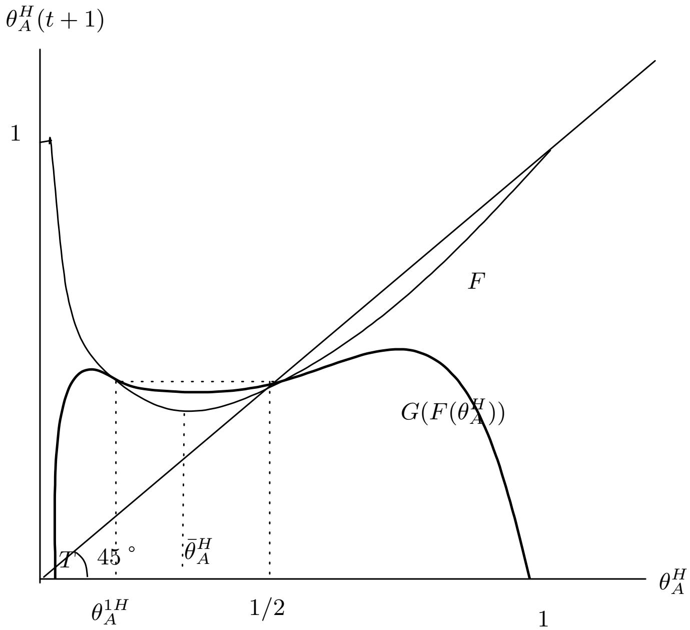


Fig. 7.3 Two-period Cycle Equilibrium and Equal-population Equilibrium

Note that when the high-income population of t-generation in one region is smaller than $\frac{\frac{\delta b}{2c}}{1+\frac{\delta b}{2c}}$, the redistributive policy $\tau(t)$ becomes a positive subsidy. When all high-income population concentrates in region B, then the optimal subsidy policy is

$$\tau_A^L = \frac{b(1-\frac{\delta b}{2c})\frac{1}{c} - 1}{b}$$

[2]Therefore, this extreme-cycle equilibrium is attained through a subsidy competition.[3]

From these considerations, we obtain the following proposition.

Proposition 5 *We have three migration equilibria in long-run.*

(1) First equilibrium is an equal-population equilibrium of high-income group. This equilibrium is a stable in the dynamics of high-income group population in each region.

---

[2]Considering $\theta_A^H(t) = 0$ and $\theta_B^H(t) = 1$, this optimal subsidy policy in region A is obtained from the following maximization problem.

Max $W_A^L(t) = y^L + \delta(y^L + (\tau_A(t+1) - c(\tau_A(t+1))\frac{y^H \theta_A^H(t+1)}{\theta_A^L}$

s.t. $\theta_A^H(t+1) = b(-\tau_A(t) + \tau_B(t))$.

[3]When $\frac{b}{c}$ is smaller than one, it is satisfied that $\frac{\frac{\delta b}{2c}}{1+\frac{\delta b}{2c}} > \theta^{H**}$.

(2) Second equilibrium occurs at a large level of the population disparity of high-income group between two region is large. At a small level of the population of high-income group in region A in period t, region A has a big in-flow of high-income population and the small level of high-income population again occurs after big out-flow in the next period. This equilibrium is a kind of cyclical equilibrium and is unstable.

(3) In a region with sufficiently small population of high-income group, the big in-flow of migration occurs and this may bring a big population of high-income group in the region, and finally full concentration of high-income group into one region may occur. Then one-pole concentration is repeated alternately in two regions. This extreme-cycle equilibrium is attained through a subsidy competition.

7.5 Concluding Remarks

This chapter presented an investigation of how regional redistributive policies in two regions. Intra-regional migration in two regions with income inequalities were developed in a framework of fiscal federation. We presented a model of tax competition between two regions and analyzed a dynamics of migration between two regions.

We obtained a unique Nash equilibrium of tax policies in two regions and clarified the characteristics of the tax policy in each region. We also investigated the migration dynamics of high-income group. We found three long-run equilibria. First equilibrium is an equal-population equilibrium of high-income group which is stable. Second equilibrium is a kind of cyclical equilibrium where big in-flow and big out-flow of high-income group are repeated, which occurs at a large level of the population disparity of high-income group between two region is large. This is unstable. Thirdly, an extreme-cycle equilibrium may occur, where one-pole concentration is repeated alternately in two regions. This is attained through a subsidy competition.

A remaining task is how the system of fiscal federation functions when the high-income group has political power. It is important to design incentives of the high-income group adequately to produce redistributive policies.

References

Acemoglu, D., and J.A. Robinson. 2006. *Economic origins of dictatorship and democracy*. Cambridge: Cambridge University Press.

Acemoglu, D., S. Naidu, P. Restrepo and J. Robinson. (2015). Democracy, redistribution and inequality. In *Handbook of income distribution,* ed. by A.B. Atkinson, vol. 2B. Chapter 21. Amsterdam: Elsevier.

Arawatari, R. and T. Ono. 2013. Inequality, mobility, an redistributive politics. *Journal of Economic Theory 148*.

Epple, D., T. Romer. 1991. Mobility and redistribution. *Journal of Political Economy 96*.

Flamand, S. 2015. Interregional transfer, group loyalty and the decentralization of redistribution. *Economics of Governance 16*.

Hirschman, A.O. 1970. *Exit, voices, and loyality: responses to decline in firms, organization, and state*. Cambridge: Harvard University Press.

Luca, B., and G. Gulcin. 2013. Income inequality, mobility, and the welfare state: a political economy model. *Macroeconomic Dynamics* 17 (6):

Oates, W. 1972. *Fiscal federalism*. New York: Harcourt Brace.

Orr, L.L., 1976. Income transfers as a public good: an application to AFDC. *American Economic Review 66*.

Waldasin, D.E. 2008. Fiscal federation. In *The new Palgrave dictionary of economics*, ed. S.N. Doulauf, and L.E. Blume. Basingstoke: Palgrave Macmillan.

Chapter 8
Population Aging and Tax Policy in Korea

Kiseok Hong

8.1 Introduction

Korea has long been considered a country with relatively stable public finances compared to other major countries. However, due to the shrinking tax base after years of low economic growth and the increasing need for welfare spending, concerns are rising about the nation's long-term fiscal sustainability. Also, the accelerating pace of population aging is expected to exacerbate fiscal pressures. Considering these structural changes and the relatively low tax burden in Korea, it seems inevitable that the tax rate will be raised by a substantial margin in the future. An important question, then, is which tax to raise. Against this backdrop, this study attempts to examine what impact population aging may have on the nation's taxation structure and to draw implications for the direction of future tax reforms.

There exist diverse views on the impact of population aging on the composition of taxation. According to common perceptions, political pressures may develop towards replacing capital income tax with labor income tax as the proportion of older people increases, because the older generation tends to rely more on capital income than labor income. However, Razin et al. (2004) and Mateos-Planas (2010) each illustrates a mechanism by which an increase in the proportion of the elderly makes people vote for a higher, not lower, capital income tax rate. In Mateos-Planas (2010), an aged population makes the economy's marginal product of capital lower, which in turn leads voters to decrease their savings and prefer higher taxes on capital. Razin et al. (2004) develops an overlapping-generations model where capital income tax transfers resources from the elderly to the young, and shows that the young median voter prefers a higher capital income tax when the elderly proportion is higher. Razin et al. (2004) also provides supporting empirical evidence using panel data gathered from 10 European countries.

K. Hong (✉)
Ewha Womans University, Seoul, South Korea
e-mail: khong@ewha.ac.kr

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_8

While Razin et al. (2004) and Mateos-Planas (2010) examine the determination of taxation structure from the political economy perspective, many other studies approach the issue from the perspective of optimal taxation theory (Mankiw and Weinzierl 2006; Aiyagari 1995; Imrohoroglu 1998; Nishiyama and Smetters 2005; and Park 2012). Since Judd (1985) and Chamley (1986), the theory of optimal taxation has traditionally contended that it is most efficient to abolish capital income tax altogether. Recently, however, studies based on the life-cycle model have proposed that the optimal rate of capital income tax may be greater than zero when there are uninsurable idiosyncratic income shocks. For example, Aiyagari (1995) argued that maintaining the capital income tax rate above zero can be Pareto improving because capital may be over-accumulated due to the precautionary savings motive under uninsurable idiosyncratic income risk. Nishiyama and Smetters (2005) also assumed cases in which individuals were exposed to uninsurable idiosyncratic shocks, even though they focused more on the insurance role of progressive income taxes rather than the possibility of over-accumulation due to precautionary savings. Their calibration shows that replacing income tax with consumer excise tax may lower individual utility because, even though the economy-wide income may rise as in the conventional theory of optimal taxation, individuals will face increased uncertainty. While differing from Aiyagari (1995) in details, Nishiyama and Smetters (2005) similarly contend that the optimal capital tax rate is greater than zero. More recently, Conesa et al. (2009) showed that, in an overlapping-generations general equilibrium model that features age dependence of labor supply elasticity and no age-dependent tax, the optimal capital income tax rate can be as high as 36%.

If the uninsurable idiosyncratic shocks faced by individuals are mainly caused by labor income uncertainty, the welfare effect of a tax reform may differ between working population (young generation) and retired population (aged generation). Also, according to Nishiyama and Smetters (2005), substantial inter-generational resource reallocation may occur if income tax is replaced by consumption tax. Given that the present retired people have already paid income taxes during their working years, a higher consumption tax (combined with an income tax cut) will make the old generation pay more lifetime taxes and effectively transfer resources to the younger generation. Nishiyama and Smetters (2005) explain that the long-term welfare increase generated by the replacement of income tax with consumption tax in Auerbach and Kotlikoff (1987) is largely due to the transfer effect of capital levy imposed on the current generations for the future generations. Also, unlike Conesa et al. (2009) who only consider steady states, Fehr and Kindermann (2015) explicitly take into account the welfare of all individuals including transitional cohorts and conclude that the optimal capital income tax rate in terms of pure economic efficiency is still zero. These studies suggest that a discussion of optimal tax will have to specify how to compare welfare between different generations and that the optimal taxation structure may critically depend on the country's demographic structure.

There are several studies that examine welfare impacts of tax reforms with respect to the Korean economy, mostly from the perspective of optimal taxation theory (Cheon and Choi 2006; Chang 2009). However, these studies focus on the steady state, with little consideration given to changes in the age structure of population. An exception is the study by Choi and Cheon (2003), which compares between lifetime tax burdens of current and future generations under various tax scenarios to finance the current national pension and health insurance schemes of Korea. According to this study, a higher income tax (as opposed to a higher value-added tax) will lighten the burden of the present generation at the expense of future generations. Similarly, Kim (2006) compares between changes in the lifetime utility of different generations that may follow a tax reform, using a CGE (Computable General Equilibrium) model. The calibration results indicate that a lower capital income tax can be beneficial, increasing the welfare of all future generations and most of the present generations.

According to the abovementioned studies, the effect of aging on taxation may materialize through both the political economic channel and the optimal taxation channel. Also, in both channels, the age structure of population is expected to play a key role. In this paper, we reexamine the implications of population aging on the composition of taxation using a CGE model with idiosyncratic income shocks calibrated to the Korean economy. In particular, we calibrate the CGE model to see how lifetime utility changes for each generation as a result of a specific tax reform, under the current and future demographic structure of Korea. The model of the paper is a standard Auerbach and Kotlikoff (1987) closed-economy model extended to incorporate idiosyncratic income shocks. While most calibration studies examine the impacts of a tax reform from the perspective of optimal taxation theory, this study takes a different angle by examining the lifetime utility of different age groups to see how political choices will be made on the tax regime under the majority voting system. Also, there exist methodological differences between this study and previous calibration studies on tax reforms in Korea. Among studies similar to ours, Choi and Cheon (2003) performs a generational accounting analysis instead of a full general equilibrium analysis and thus does not take various general equilibrium effects into account. Kim (2006) uses a general equilibrium framework but with no reference to income shocks.

8.2 Model

We develop an overlapping-generations general equilibrium model that consists of 60 generations (equivalent to real life ages of 21 through 80) and covers the period from 2011 forward. The model assumes that individual consumers in each generation make optimal consumption/saving decisions for the rest of their lives and leave no bequests as in typical Auerbach-Kotlikoff models. For the sake of simplicity, it is also assumed that the labor supply of each consumer is constant during his/her entire working periods before retirement. However, in order to incorporate a trend increase in economy-wide labor productivity into the model, we introduce an exogenous growth of the amount of endowed time over generations (across cohorts). In other words, we assume that, while the labor supply throughout an individual's lifetime is constant, there is a difference in the effective labor supply (or endowed time) across generations.

Also to be noted in the model is the consideration of income risks. Existing studies show that the influence of taxation on individuals' utility may greatly differ depending on the nature of income risks. Accordingly, the model in this study explicitly considers the existence of an uninsurable idiosyncratic income risk, despite the resulting complications in the model's structure. The idiosyncratic income risk is represented by the probability of unemployment and employment: individuals face the probability of employment/unemployment in the next period conditional on this period's employment/unemployment status and maximize their expected utility taking the probability into consideration. There is no macroeconomic risk assumed in the model.

The rest of the model is fairly conventional. The firm sector is assumed perfectly competitive, which dictates that the interest and wage rates are equal to the marginal productivity of capital and labor, respectively. The government levies taxes on labor income, capital income, and consumption, and uses up the revenue each period, keeping the fiscal budget always balanced. The details of the model are as follows.

(1) Consumers

The utility of a consumer relies only on c, the consumption of goods and services, and is independent of leisure. Utility is given by a CRRA (Constant Relative Risk Aversion) function. Each consumer expects to survive until 60 years of age with no uncertainty in life expectancy. The existence of income uncertainty (the possibility of unemployment) implies that the consumer's optimization problem can be represented by a Bellman equation as follows:

$$\begin{aligned} v_{i,t}(A_{i,t}) &= u(c_{i,t}) + (1+\theta)^{-1} E_t[v_{i+1,t+1}], \\ u(c_{i,t}) &= c_{i,t}^{(1-1/\rho)}/(1-1/\rho), \quad i = 1, \ldots, 60 \end{aligned} \tag{8.1}$$

where $v_{i,t}$= value function of an individual aged i in period t

$A_{i,t}$= asset of an individual aged i in period t

$c_{i,t}$= consumption of an individual aged i in period t

θ= time preference rate

$E_t\,()$ = expectation conditional on all information available in period t

$1/\rho$ = degree of relative risk aversion.

The one-period budget constraint of an individual aged i in period t is given as follows:

$$A_{i+1,t+1} = A_{i,t}(1 + r_t(1-\tau_k)) + w_t(1-\tau_l)l_{i,t}h_{i,t} + b_{i,t} - c_{i,t}(1+\tau_c) \tag{8.2}$$

where $A_{i,t}$ = asset of an individual aged i in period t

r_t = interest rate in period t

w_t = wage rate in period t

τ_k = capital income tax rate

τ_l = labor income tax rate

$h_{i,t}$ = efficiency of labor of an individual aged i in period t

$l_{i,t}$ = labor supply of an individual aged i in period t

$c_{i,t}$ = consumption of an individual aged i in period t

b_{it} = transfer income (unemployment insurance) of an individual aged i in period t.

The size of assets is always equal to or greater than zero ($A_{i,t} \geqslant 0$) due to liquidity constraints, and $A_{61,t} = 0$ and $A_{1,t} = 0$ for all t as there is no bequest. Also, as mentioned above, each individual is assumed to believe that the probability of mortality is zero before the age of 60 and 1 at the end of age 60. In the actual population statistics used in calibrations, however, the mortality probability is always greater than zero and smaller than 1 for each cohort at each age, making the size of a cohort decrease with age. This means that the budget constraint given in Eq. (8.2) does not hold exactly between periods t and $t+1$ when aggregated over all individuals in the same cohort. To resolve this problem, we assume that the assets belonging to individuals who die between t and $t+1$ are distributed evenly to other individuals in period $t+1$.

As is well known, there is no analytical solution to the expected utility maximization problem given in Eqs. (8.1) and (8.2). We obtain the solution to the problem numerically through a grid search.

(2) Firms

The firm sector is perfectly competitive and production technology is given by the standard Cobb–Douglas function:

$$Y_t = K_t^{\beta} L_t^{1-\beta}, \tag{8.3}$$

where $K_t = \sum_i A_{i,t-1}$ and $L_t = \sum_i h_{i,t} l_{i,t}$. The total capital stock within the economy is wholly owned by the private sector and firms produce output Y by combining total capital stock K and labor L. The equilibrium interest rate and wage rate are as follows:

$$r_t = \beta(K_t/L_t)^{\beta-1}$$
$$w_t = (1-\beta)(K_t/L_t)^{\beta} \tag{8.4}$$

where $r$and w are the interest rate and the wage rate, respectively.

(3) Government

The government makes consumptions expenditures G and transfer payments E, using revenues raised from taxes on labor income, capital income, and consumption. The fiscal budget is assumed to be in balance for every period.

$$G_t + E_t = \tau_l w_t L_t + \tau_k r_t K_t + \tau_c C_t, \tag{8.5}$$

where $C_t = \sum_i c_{i,t}$ and $\tau_l,\ \tau_k,\ \tau_c$ denote tax rates on labor income, capital income, and consumption, respectively.

(4) Income Risk

The employment status of an individual in every period is determined by the following Markov transition matrix:

$$[p_{11}, 1-p_{11}; p_{21}, 1-p_{21}], \tag{8.6}$$

where p_{11} denotes the probability that an individual who is currently employed will remain employed in the next period, and p_{21} the probability that an individual who is currently unemployed will be employed in the next period. As the transition matrix is assumed to be constant over time, the aggregate unemployment rate in each period is fixed at $(1-p_{11})/(1-p_{11}+p_{21})$.

Unemployment benefits paid to an unemployed individual are assumed to be lower than market wage, so that there is no perfect insurance against labor income risks. As mentioned above, these risks are purely idiosyncratic and pose no uncertainty for the economy as a whole.

(5) Solutions to the Model

The model covers two centuries of period from 2011 to 2210, a timespan long enough to ensure the existence of the steady state. Once population projections are provided for the entire period, the model can be solved numerically to generate equilibrium paths of key macroeconomic variables. As the National Statistical Office of Korea provides population projections only for the years up to 2065, we fill the gap for the latter years by constructing our own projections. Specifically, we assume that the probability of mortality becomes zero for all individuals starting from year 2016, except for those aged 60 whose survival rate into the next period is zero by construction. We also assume that the number of newborns (individuals aged 1) remains the same each year at the 2065 value. Under these assumptions, individuals aged i in year $2065+t$ would turn into individuals aged $i+1$ in $2065+t+1$ with no dropout, and eventually the age distribution of population would become completely uniform in year 2144 and forward. As will be seen below, the steady state of the economy is achieved soon after the age structure becomes uniform.

Another set of exogenous values to be put into the model is the initial asset holdings of current generations. Individuals who are currently alive in 2011 (the initial year) must have made consumption/savings decisions from the first year of their lives, taking the then current and future series of macroeconomic variables as given. Therefore, in order to endogenously derive the size of assets held by individuals currently alive, one needs to extend the model's time span by 60 years from 2011 backwards. But then, in order to determine the optimal choice of each age group 60 years ago, the model's time span would have to be extended backwards even further. It is clear that this approach will reach no solution. The only feasible option would be to exogenously assume the size of initial asset holdings for each age group in 2011. We will discuss in the next section how this is done in our calibrations.

The model's solutions are obtained by Gauss–Seidel iterations. We first start with arbitrary series of the interest and wage rates for the entire 200 years and arbitrary values for initial asset holdings by each age group in 2011, and derive individuals' optimal consumption and saving schedules. From the individual choices, the size of the economy-wide capital stock and labor can be calculated, which in turn can be plugged back into Eq. (18.2) to generate new series of interest and wage rates. This process is repeated until the fixed point of interest and wage rates is reached.

Due to income uncertainty, an individual's optimal consumption/saving decisions can be solved backwards from the last period of life using the Bellman equation given in Eq. (8.2). Given the domain of the state variable to be used in the grid search, a, the maximization problem in Eq. (8.2) can be viewed as a mapping process that finds the optimal value of a in the next period for each value of a in the present period. In the last period of life, the value function is identical with the one-period utility function and thus can be easily defined. Once the value function for the last period is given, value functions for all preceding periods can be obtained sequentially in reverse order.

As income uncertainty causes intra-generational heterogeneity among individuals, one needs to take aggregation not only over different age groups but also over different individuals within each age group to obtain the total capital stock and total labor supply for the economy. The distribution of individuals within an age group can be expressed by $\lambda(a, s)$, where a and s represent asset size and employment status, respectively, and $\sum_s \sum_a \lambda(a,\ s) = 1$. The probability distribution $\lambda(a, s)$ will change over time as follows:

$$\lambda_t\left(a^{'},\ s^{'}\right) = \sum_s \sum_{a:a' \in M_t(a,s)} \prod(s^{'},\ s)\lambda_{t-1}(a,\ s) \tag{8.7}$$

In the above equation, a' and s' (a and s) represent assets and employment status in period t(period $t-1$), respectively, and M_t is a mapping that assigns the optimal asset level in period t to each asset level in period $t-1$. The transition matrix $\prod(s^{'},\ s)$

represents the probability that the employment status will shift from s in period $t-1$ to s' in period t. For retirees who have no employment risks, the transition matrix becomes irrelevant and there is no heterogeneity across individuals within the same age group.

(6) Parametrization of the Model

First, the transition matrix in Eq. (8.6) that determines the stochastic nature of the model is assumed as follows: $\prod\left(s', s\right) = [0.98, \quad 0.02; \quad 0.5, \quad 0.5] \qquad (8.6)'$

This assumption is based on estimates of Nam et al. (2005) which made use of Korea's Economically Active Population Survey data. The above transition matrix implies that the unemployment rate in each period will be 3.85%.

The domain of the state variable (assets), a, is set to be between 0 and 100 based on preliminary calibration results. Preliminary analyses indicate that an individual's lifetime asset accumulation reaches up to 50–60, sufficiently smaller than the range of 0–100. For the consumer's objective function, we assume logarithmic utility and the time preference rate of 0.015.

The supply of labor is exogenously fixed at 1 if employed, and 0 if unemployed or retired (aged 41 or older). The efficiency of labor, on the other hand, is assumed to vary with age even during working years in order to allow for age related differences in experience and skill. We utilize Miles (1999)'s estimates of labor efficiency given as follows:

$$\text{efficiency of labor}_i = 0.05 * (i + 20) - 0.0006 * (i + 20)^2, \qquad (8.8)$$

where i is the model age, ranging from 1 through 40. In addition, we implement economy-wide labor productivity growth by assuming that the amount of effective time endowment increases over each generation by 2%.

The cross section of the initial assets held by each generation as of 2011 is assumed to be the same with the (expected) lifetime asset holding pattern of individuals aged 1 in 2011, adjusted for the aforementioned intergenerational labor productivity differences. With the initial asset holdings assumed this way, the lifetime asset holding pattern of age 1 individuals in 2011 can be derived anew from the model, which in turn can be used to revise the cross section of initial asset holdings. This process can be repeated until there is no revision to be made.

The capital share of income in the Cobb–Douglas production function, β, is assumed to be 1/3. Our baseline values for fiscal variables, i.e., the capital income tax rate, consumer excise tax rate, and labor income tax rate, are 12, 10, and 18%, respectively, in accordance with the effective tax rates estimated in Kim and Kim (2007). The unemployment benefits are assumed to be 20% of wages, based on Chang (2009). As unemployment benefits are tax exempt, the proportional labor income tax can be progressive.

8.3 Calibration Results

The determination of taxation structure may critically depend on what alternatives are examined. For instance, the choice between a capital-income tax hike and a labor-income tax hike may change, depending on the current rate of each income tax, accompanying changes in other taxes, the availability of a progressive system, etc. In this study, after considering the current fiscal conditions of Korea and the structural limitations of our model, we examine the following two scenarios. The first scenario considers a revenue-neutral tax reform where capital income tax is replaced with labor-income tax or consumption tax. Many previous studies on tax reforms address a similar situation. In the second scenario, with the recognition that tax hikes will be inevitable in Korea due to the progress of population aging, we consider raising additional revenue through capital income tax, labor income tax, or consumption tax. Although the impact of a tax reform may change depending on how the revenue is used, we simply assume that all tax receipts, apart from unemployment benefits, would be consumed by the government, and ignore any possible alternatives.

1. Certainty Case

Even though our model features uncertainty over employment status and income, we first calibrate the model for a special case where unemployment risk is zero. By comparing the results from the certainty case with those from a more realistic setup in the next section, we will be able to understand the implications of employment/income risks on taxation choices. Also, while the solutions to the uncertainty model are obtained only approximately through grid search, exact solutions can be derived analytically for the certainty model. Therefore, for the certainty case, we can use both approaches to obtain exact solutions as well as approximate solutions and confirm their consistency.

Baseline Results

Calibration results obtained from the certainty model under baseline assumptions are as follows. First, as Fig. 8.1 shows, asset accumulation over an individual's lifetime (60 years) exhibits inverted U shape, a characteristic feature of the lifecycle model. Although Fig. 8.1 depicts the decision of an individual whose age is 1 in 2011, similar patterns are obtained for individuals in other cohorts. Second, Fig. 8.2 shows that the interest rate, a key macroeconomic series in the model, is predicted to keep decreasing for the next 20–30 years and then gradually increase until reaching the steady state. The predicted decrease in the interest rate is driven by population aging and the resulting increase in the capital/labor ratio. Under an aging population, the aggregate supply of labor falls while the accumulation of capital (assets) increases, generating aforementioned changes in the capital/labor ratio and marginal product of capital. Figures 8.1 and 8.2 confirm that our model is capable of predicting macroeconomic implications of population aging within the framework of the life-cycle model.

Fig. 8.1 Asset accumulation over lifetime

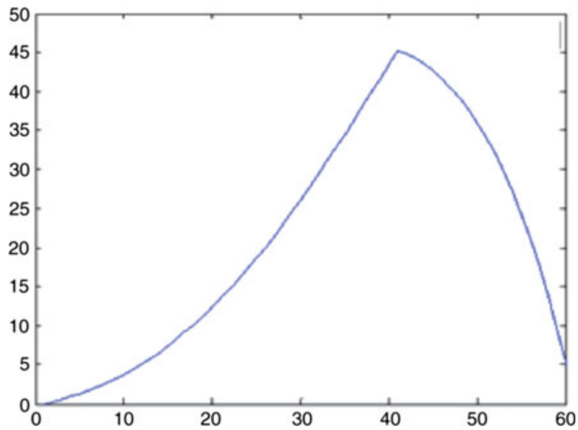
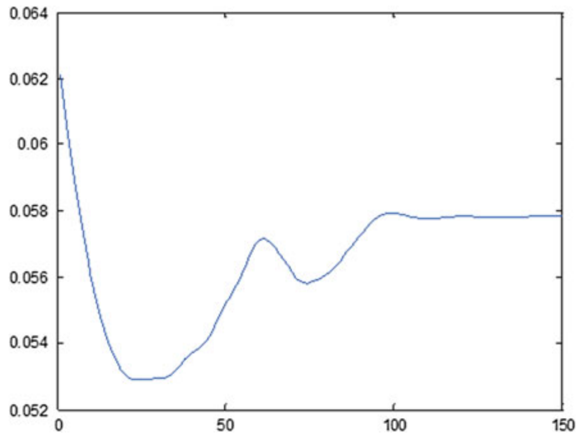


Fig. 8.2 Interest rate series



Scenario 1

As explained above, the baseline tax system consists of 12% capital income tax, 10% consumption tax, and 22% labor income tax. Scenario 1 considers a case in which the capital income tax rate is reduced to zero and either the consumption tax or labor income tax rate is raised to maintain the total tax revenue/GDP ratio at the same level.

Figure 8.3 depicts how much lifetime utility changes for each generation when the capital income tax rate is replaced with the consumption tax, compared to the baseline. A value greater than 1 in Fig. 8.3 indicates that the lifetime utility of the corresponding cohort increases as a result of the tax reform. The horizontal axis represents each cohort, with the first numbers from 1 to 60 denoting the present age groups in 2011 and the following numbers from 1 to 150 denoting the generations to be born in 2012 through 2161. According to Fig. 8.3, most of the current generations except those in their 20 and 30 s (real life age 40–59) are expected to experience a

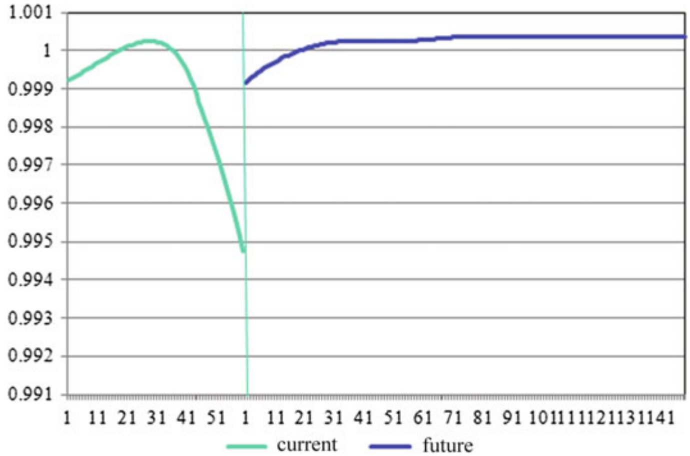


Fig. 8.3 Consumption tax hike against the baseline

decrease in lifetime utility when the capital income tax is replaced with consumption tax. In contrast, most of the future generations, except those to be born within 20 years or so, will become better off under the tax reform. Although not shown in the figure, the model also predicts that, in order to maintain the total tax revenue/GDP ratio constant in each period despite the abolition of the capital income tax, the consumption tax rate will have to rise to 16.8% on average.

The result that individual welfare may be improved in the long run by abolishing the capital income tax and raising the consumption tax agrees with conclusions of the conventional optimal taxation theory. However, according to Fig. 8.3, the welfare of future generations may increase only at the expense of most current generations. In particular, the current elderly generations are expected to be adversely affected by the consumption tax hike. The same point has been made by Nishiyama and Smetters (2005), who observed that the welfare of aged people may be sacrificed if capital income tax is replaced with consumption tax. The calibration results in Kim (2006), on the other hand, indicate that not only future generations but also most current generations would see their lifetime utility increase if the capital income tax rate were lowered by 10% points and the consumption tax were raised instead.

Figure 8.3 suggests that, whether the proposal to raise the consumption tax would be adopted will depend critically on the conflicts between the present and future generations. If political decisions are made by majority voting where only the current generations can cast votes, the influence of the elderly people will become increasingly important due to population aging, and it is unlikely that the proposal to raise the consumption tax will be adopted regardless of its long-term desirability.

Figure 8.4 shows the results of substituting capital income tax with labor income tax. As before, the curves in Fig. 8.4 indicate how the lifetime utility changes for each generation as a result of the tax reform, compared to the baseline. According to Fig. 8.4, most of the current generations, except those younger than 10 years (real

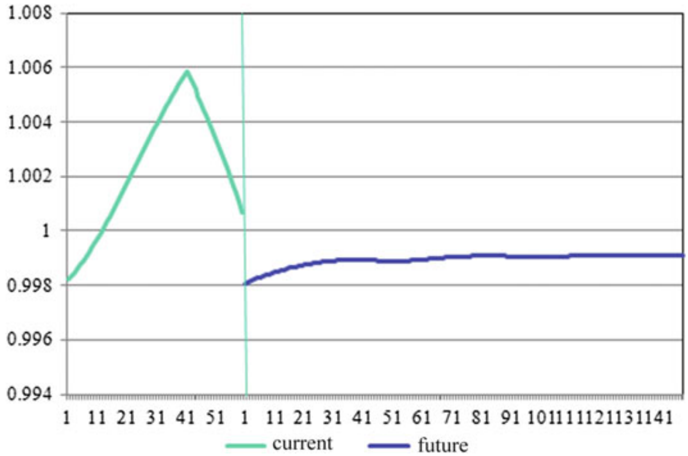


Fig. 8.4 Labor income tax hike against the baseline

life age younger than 30 years), would prefer a proposal for raising the labor income tax to the baseline, with all of the future generations preferring the baseline. The new labor income tax rate required to keep the total tax revenue/GDP ratio constant is estimated to be 23.5%. Figures 8.3 and 8.4 together indicate that the current generations would typically prefer a labor income tax hike, followed by the baseline, and then a consumption tax hike proposal. This result fits with intuitive expectations. For example, the current elderly people in their late stages of life would prefer the labor tax hike scenario to the baseline, as they rely less on labor income and have relatively high capital accumulation. Also, between a consumption tax hike and a labor income tax hike, the elderly people would prefer the latter, as the consumption tax will spread the burden more evenly over all age groups including the elderly. On the other hand, generations who would be born in the distant future would most prefer a proposal to raise the consumption tax, as they will be able to enjoy the full effects of higher capital stock and output increase that will arise from replacing capital income tax with consumption tax. While the preferences of the current and future generations can be sharply different as seen above, the political decision making process may only represent the current generations and result in adopting a tax scheme that goes against efficiency from a long-term perspective.

Scenario 2

While Scenario 1 examined revenue-neutral tax reforms, Scenario 2 considers tax reforms designed to raise more revenue. According to an empirical analysis of OECD member countries, for every 1% point increase in the share of population aged 65 or above, the total tax revenue/GDP ratio rises by 0.3–0.4% point (Hong 2013). And the population projection made by the National Statistical Office of Korea predicts that the share of people 65 years and older will rise from 11% in 2011 to 37% in 2050, a 26% point increase in 40 years. Using these two estimates, one can anticipate the tax revenue/GDP ratio to increase by 8–10% points over the 2011–2050 period as a result of population aging. Based on this estimate, we will assume in Scenario 2 that the nation's total tax burden (as a ratio of GDP) will rise by 0.2% points per year over the 40 years from 2011 to 2050 and then stay constant. Since the total tax revenue/GDP ratio is about 22% in the baseline model (with the capital income tax rate, consumption tax rate, and labor income tax rate given by 12, 10, and 18%, respectively), Scenario 2 will involve the ratio gradually increasing up to 30% by 2050.

In particular, for each of the three kinds of tax, we will consider an increase in the tax rate necessary to achieve the planned increase in tax revenue with the remaining two tax items kept unchanged. Once calibrations are completed for all three alternatives, we can compare how the lifetime utility of each cohort turns out under each alternative. In order to make comparison easier, the case in which the capital income tax rate is endogenously adjusted will be used as the benchmark. The lifetime utility of each cohort for the remaining two cases are expressed as a ratio relative to the benchmark value.

We first compare the benchmark case (proposal to increase the capital income tax rate) with the case with a consumption tax increase. The capital income tax rate and the consumption tax rate required under each proposal are about 35 and 25%, respectively, in the steady state. According to Fig. 8.5, most of the current generations favor the proposal to raise capital income tax rather than consumption tax, while most future generations would prefer the consumption tax hike proposal. Note that the outcome in Fig. 8.5 is very similar to that in Fig. 8.3. We believe this suggests that the preference of each cohort towards taxation composition is robust to reasonable variations in the details across different scenarios.

In Fig. 8.6, where the benchmark case is compared with the case with a labor income tax increase, most of the current generations are shown to favor the labor income tax hike proposal while future generations invariably prefer the benchmark proposal. Again, the results in Fig. 8.6 are almost identical to those in Fig. 8.4, suggesting robustness of each cohort's preference towards tax composition. The steady state labor income tax rate underlying Fig. 8.6 is about 30%.

2. Uncertainty Case

Now we turn to the model that incorporates uncertainty and conduct the same scenario analysis. As explained above, the solutions to the uncertainty model are approximate values obtained through grid search. For this reason, the calibration outcomes presented in this subsection are not perfectly smooth curves.

Baseline Results

As in the certainty model, we first plot individuals' asset accumulations and the interest rate series under the baseline tax rates. The existence of employment/income uncertainty implies that individuals may exhibit heterogeneity in terms of employment status and the size of asset holdings even within the same cohort. Even though individuals in the same cohort are born the same, the distribution of the individuals in the asset domain, a, becomes more dispersed with age, as the employment

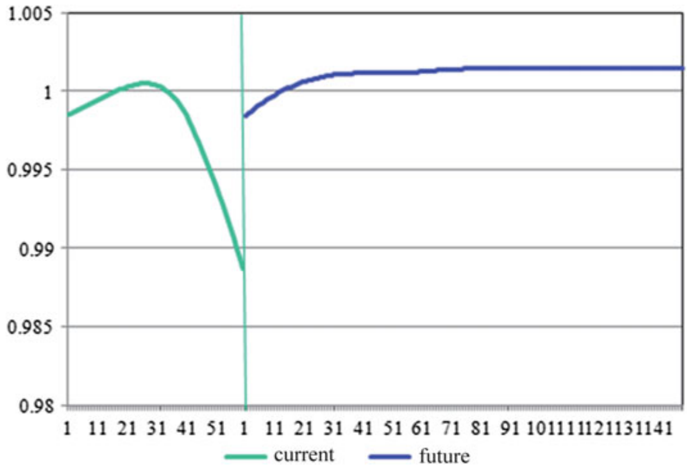


Fig. 8.5 Consumption tax hike versus capital income tax hike

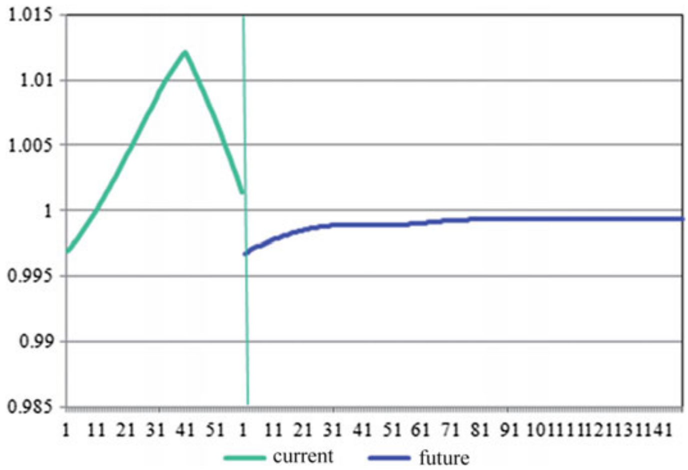


Fig. 8.6 Labor income tax hike versus capital income tax hike

status is realized in a probabilistic manner each period. The profile of asset accumulation over the life cycle as shown in Fig. 8.7 was obtained by taking such individual differences into account and computing the expected value of the probability distribution of asset holdings at each age, for the cohort whose age is 1 as of 2011. Figure 8.7 shows that the age pattern of asset holdings is an inverted U, similar to the one for the certainty case reported in Fig. 8.1. While the precautionary savings motive under employment/income uncertainty may induce individuals to increase asset holdings, the asset profile in Fig. 8.7, being the average for both the employed and the unemployed individuals, is slightly lower than that in Fig. 8.1. When Figs. 8.8

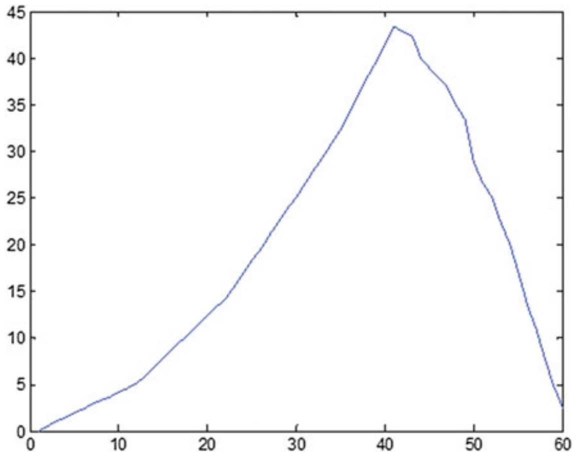


Fig. 8.7 Asset accumulation over lifetime: uncertainty model

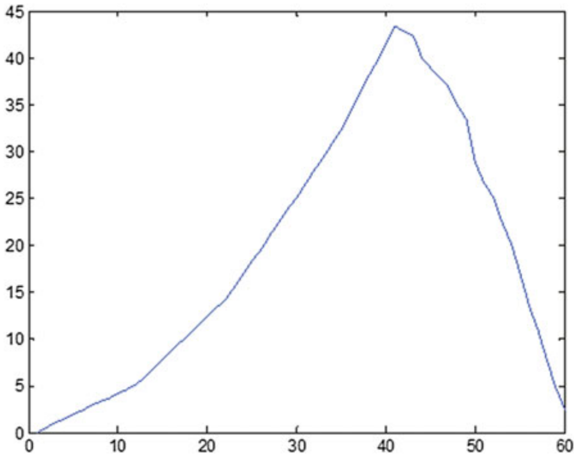


Fig. 8.8 Interest rate series: uncertainty model

and 8.2 are compared with each other, the interest rate in the uncertainty model turns out clearly lower. This indicates that the capital/labor ratio becomes higher under the uncertainty model as the stochastic incidence of unemployment lowers total labor by more than total capital.

Scenario 1

The intra-cohort individual heterogeneity, driven by unemployment risks, makes it more complicated to compare individuals' lifetime utilities across different tax reform proposals. In order to circumvent the difficulty that the lifetime expected utility of a current cohort cannot be given as a single value any more, we instead compute the proportion of people who prefer a certain reform plan for each cohort. This exercise, however, while conforming to our intention to depict the political decision making through majority voting, cannot show the strength of preference for each plan. For this reason, we also compute the average lifetime utility over all individuals within each cohort, assuming a utilitarian social welfare function. Of course, such intra-cohort heterogeneity in lifetime expected utility exist only for present generations. Individuals belonging to a future cohort are *ex ante* identical in every respect. Accordingly, the proportion of individuals preferring a certain proposal within a future cohort will always be either 0 or 1.

First, Fig. 8.9 shows the preference of each cohort towards the plan that adjusts the consumption tax rate endogenously to keep the total tax revenue/GDP ratio unchanged while reducing the capital income tax rate to zero, against the baseline. The first panel in the figure illustrates the proportion of individuals in each cohort that would prefer the proposal to raise the consumption tax rate, while the second panel represents the average lifetime expected utility of all individuals under the consumption tax hike proposal as a proportion to the corresponding value under the baseline. As before, the horizontal axis represents each cohort, with the first numbers from 1 to 60 denoting the present age groups in 2011 and the following numbers from 1 to 150 denoting future generations. According to the first panel, approximately 90% of individuals aged 10 and 20 s (real life age of 30 and 40 s) among the present generation would prefer the consumption tax hike proposal, while all other individuals in the present generation prefer the baseline. Among future generations, all would favor the consumption tax hike proposal except those who will be born within the first few years. The same pattern can be observed in the second panel, which shows that, while the lifetime expected utility will be lowered on average for most of the present generation under the consumption tax hike, most future generations would see their lifetime utility increase under the tax reform. Figure 8.10 shows the results of repeating the same analysis on the labor income tax hike proposal. As in Fig. 8.9, the first panel illustrates the share of population in each cohort that prefers the reform proposal, with the second panel comparing each cohort's average lifetime utility under the reform proposal against the baseline value. A notable result in Fig. 8.1 is that, in contrast to the outcomes of the certainty model reported in Fig. 8.4, most of the current generations and future generations are expected to prefer the labor income tax hike to the baseline when employment/income uncertainty is introduced. One may expect young individuals to oppose to a labor income tax hike as they rely heavily on labor income. However, Fig. 8.1 shows that future generations will likely choose a labor income tax hike combined with a zero capital tax than the baseline. The main reason for this is that, under the existence of employment/income uncertainty, the insurance role of the labor income tax becomes important (Nishiyama and Smetters 2005). Since unemployment benefits are exempt from labor income tax in Korea, a labor income tax hike can effectively increase the progressivity and insurance role of the labor income tax.

To summarize the results of Figs. 8.9 and 8.10: most of the present generations prefer a labor income tax hike, followed by a capital income tax hike, followed by a consumption tax hike proposal while most future generations would favor a consumption tax hike >labor income tax hike >capital income tax hike. These results are identical to the preference orders obtained in the certainty model except for the preference of future generations for the labor income tax hike proposal against the

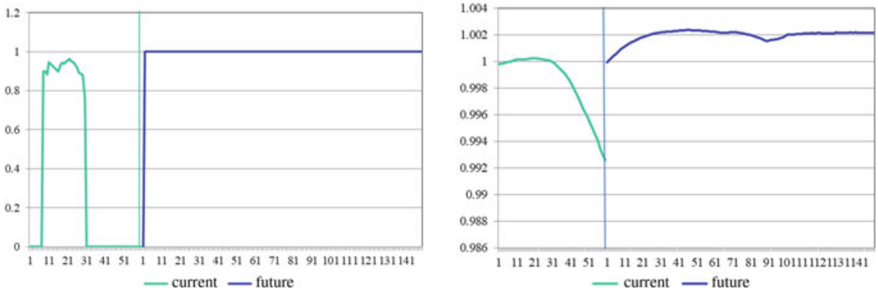


Fig. 8.9 Consumption tax hike against the baseline: uncertainty model

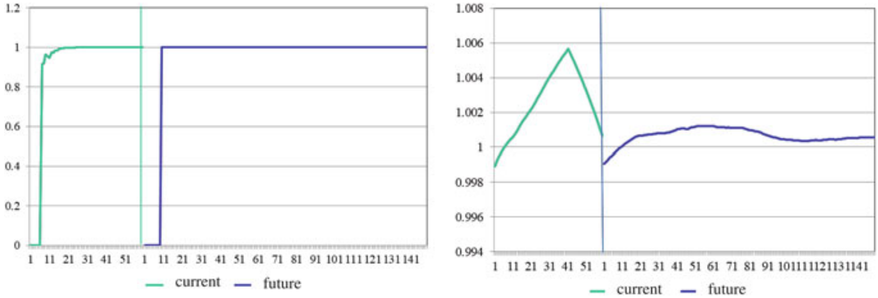


Fig. 8.10 Labor income tax hike against the baseline: uncertainty model

baseline. Future generations in the uncertainty model favor the labor income tax hike proposal due to the insurance role of the labor income tax.

Scenario 2

Now we turn to Scenario 2 under the uncertainty model. As in the certainty model, the case with endogenous capital income tax is set to be the benchmark against which the remaining two cases are compared. According to Fig. 8.11, all individuals in the present generation prefer the baseline over the labor income tax hike proposal, while those to be born in the distant future favor the opposite. Also, Fig. 8.12 shows that all individuals in the present generation prefer the labor income tax hike proposal over the baseline while most future generations favor the opposite. By combining the outcomes in Figs. 8.11 and 8.12, we can see that most of the present generations prefer in the order of labor income tax hike >capital income tax hike >consumption tax hike proposal, while most future generations prefer a consumption tax hike over a capital income tax hike over a labor income tax hike. These results are identical to the preference orders obtained under the certainty model.

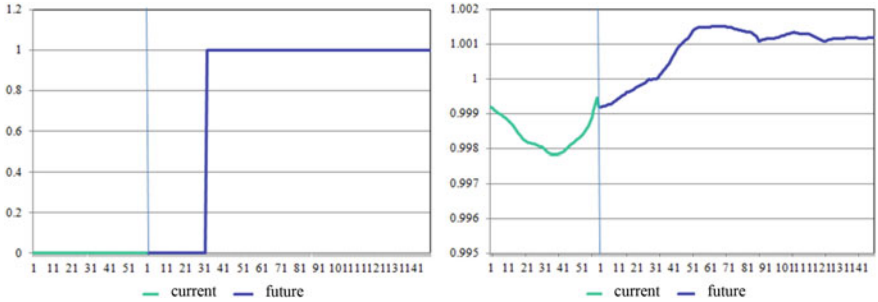


Fig. 8.11 Consumption tax hike versus capital income tax hike: uncertainty

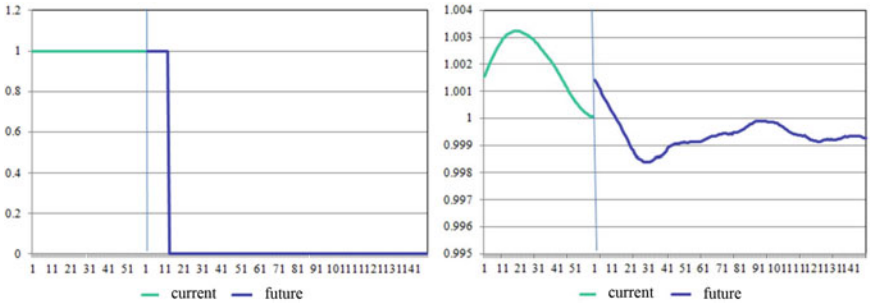


Fig. 8.12 Labor income tax hike versus capital income tax hike: uncertainty model

8.4 Conclusion

This study examines possible impacts of population aging on tax composition through calibration of a dynamic general equilibrium model with overlapping generations. Our calibration results confirm that the preferences of the present and future generations often conflict. In particular, it is likely that political decisions made through majority voting may favor the current elderly people who prefer a labor income tax hike, in disagreement with the welfare of future generations.

The general equilibrium analysis of this paper has several limitations. First, it only compares between a few tax reform proposals instead of searching for the globally optimal tax structure. The preference order for the reform proposals reported in this paper may not hold when more alternatives are taken into consideration. Second, this paper assumes that individual labor supply is given exogenously and thus cannot address possible tax distortions on labor supply. Third, while people's preferences for a tax reform plan will clearly depend on how the tax revenue is used by the government, this study has ignored government expenditure altogether for the sake of simplicity. For future studies, it will be important to improve the reality of the model by incorporating the expenditure side of the government budget.

References

(In Korean)

Chang. 2009. The analysis of pension reform and tax reform under population ageing. National Assembly Budget Office.
Cheon, Y., and S. Choi. 2006. The effects of tax reforms: an analysis using a computable general equilibrium model. In *Korea's fundamental tax reform toward 21st century*ed. by T. Kwak. Korea Economic Research Institute.
Choi, L., and B. Cheon. 2003. Fiscal implications of aging and inter-generational redistributive effects of fiscal policy. Korea Institute of Public Finance.
Hong, K. 2013. *Population aging and the determination of tax structure, mimeo*. Korea Institute of Public Finance.
Kim, S. 2006. A general equilibrium model of tax policy evaluation in Korea: tax incidence analysis. Korea Institute of Public Finance.
Kim, S., and W. Kim. 2007. The marginal efficiency cost of taxation in Korea: an econometric approach. Korea Institute of Public Finance.
Nam, J., G. Lyu, and H. Choi. 2005. The current status of job-insecure workers and the employment policy issues. Korea Labor Institute (In English).

(In English)

Aiyagari, R. 1995. Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy* 103 (6): 1158–1175.
Auerbach, A.J., and L.J. Kotlikoff. 1987. *Dynamic fiscal policy*. Cambridge: Cambridge University Press.
Chamley, C. 1986. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* 54: 607–622.
Conesa, J.C., S. Kitao, and D. Krueger. 2009. Taxing capital? Not a bad idea after all!. *American Economic Review* 99 (1): 25–48.
Fehr, H., and F. Kindermann. 2015. Taxing capital along the transition-not a bad idea after all? *Journal of Economic Dynamics and Control* 51: 64–77.
Imrohoroglu, S. 1998. A quantitative analysis of capital income taxation. *International Economic Review* 39 (2): 307–328.
Judd, K. 1985. Redistributive taxes in a simple perfect foresight model. *Journal of Public Economics* 28: 59–83.
Mankiw, G., and M. Weinzierl. 2006. Dynamic scoring: a back-of-the- envelope guide. *Journal of Public Economics* 90 (8–9): 1415–1433.
Mateos-Planas, X. 2010. Demographics and the politics of capital taxation in a life-cycle economy. *American Economic Review* 100 (1): 337–63.
Miles, D. 1999. Modelling the impact of demographic change upon the economy. *The Economic Journal* 109: 1–36.
Nishiyama, S., and K. Smetters. 2005. Consumption taxes and economic efficiency with idiosyncratic wage shocks. *Journal of Political Economy* 113 (5): 1088–1115.
Park, Y. 2012. Optimal taxation in a limited commitment economy. In PIER working paper archive 12-033, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.
Razin, A., E. Sadka, and P. Swagel. 2004. Capital income taxation under majority voting with aging population. *Review of World Economics* 140 (3): 476–495. (In English).

Chapter 9
Public Debt, Lobbying and Endogenous Redistribution in an Overlapping-Generations Model

Tsuyoshi Shinozaki, Isidoro Mazza and Minoru Kunizaki

9.1 Introduction

The recent slowdown of global economy contributed to ignite a sovereign debt crisis that has shaken several economies, and the European Economic and Monetary Union in particular. These events have helped to revamp the interests of scholars on fiscal reconstruction. Decisions concerning the amount, speed and approach (cut in public expenditures vs. tax increase) of containment and reduction of public debt have clear distributive impacts not only within but also across generations (think, for example, of the problem of sustainability of the pension system in aging societies). Such redistributive issues openly call for a political economic analysis.[1]

This paper investigates the redistributive impact of fiscal reconstruction across generations, in an overlapping generations model with short-lived governments. We assume that the incumbent government stays in power only for one period, during which it enacts policies caring about the welfare of its voters in that period. They are one young individual, who does not consume but invests for future consumption and contributes to public revenue, and one old, who just consumes. We also extend our political economic analysis by allowing the young to engage in lobbying by offering campaign contributions to the incumbent to affect the tax rate chosen by the latter.

[1]Drazen (2000) is an excellent reference for the political economic analysis of economic reforms.

T. Shinozaki (✉)
Tohoku Gakuin University, Sendai, Japan
e-mail: shinozaki@mail.tohoku-gakuin.ac.jp

I. Mazza
University of Catania, Catania, Italy

M. Kunizaki
Aichi University, Nagoya, Japan

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_9

The dynamic analysis applies the Markov perfect equilibrium concept, already used in similar political contexts by Krusell et al. (1997), Grossman and Helpman (1998), Ihori and Itaya (2001), Ono (2009), among others.

The main results of this study highlight that a short lived government may choose an inefficiently high debt and low tax rate, in a steady state. However, debt may also dynamically tend to diverge from the efficient level. The most surprising result is perhaps that the previous outcome can be reverted in case of lobbying by the young, when the incumbent has a bias for redistribution. In this case, as public debt accumulates, the government increases tax. If the government assigns a high weight to redistribution, the lobbying young may concede to a tax increase trading it off for higher transfers in next period.

The motivation of our study is close Ihori and Itaya (2001) but it extends their analysis in two directions. First, it accounts for the influence of lobbying on policymaking, which has become widely addressed in political economic studies especially after the influential contribution by Grossman and Helpman (1994). Secondly, it assumes a short-lived government in line with a common interpretation of governments as not (fully) benevolent agents with a short horizon, usually coinciding with next election term (see Grossman and Helpman 1998). Short term horizon for the incumbent gives the latter an incentive to accumulate an inefficiently high level of public debt that represents a burden for future generations and that may become economically unsustainable. For these aspects, our model is close to that adopted by Ono (2009) to examine the relationship between environmental lobbying and environmental stock.

The paper is organized as follows. Section 9.2 describes the model and derives the efficient policies as a benchmark. Section 9.3 shows a political economic analysis, first in absence, and then in presence of lobbying. Section 9.4 concludes the study with some final comments to this preliminary version.

9.2 The Model

We consider an overlapping generations economy in a small country. Each individual lives two periods of time, the young and the old periods, and both young and old generations exist in every period. Individuals are assumed to be identical within as well as across generations. A new generation is born in each period $t = 1, 2, 3, \ldots$ The size of each generation is assumed to be one.

In time t, the young supplies one unit of labor inelastically, earning a wage income w which given in a small country and paying a lump-sum income tax $\tau_t \geqslant 0$. As in Grossman and Helpman (1998) and Ono (2009), we assume that the young does not consume. The share of wage that the young contributes depends on the political environment. The young may also engage in exerting political influence by offering campaign contributions to the government. Thus, the budget constraint of the young is:$w - \tau_t - z_t = s_t$, where z_t and s_t are campaign contributions and savings, respectively. Savings earn a gross rate of return in the following period, which is $(1 + r)\, s_t$, where r is given. In the second period, the old does not work but consumes the sum of the rental income from savings and the income transfer, $\alpha\tau_t$. Therefore, the consumption is: $c_t = (1+r)(w - \tau_{t-1} - z_t) + \alpha\tau_t$, where $\alpha \in (0, 1)$ denotes the fraction of income transfer (or intergenerational redistribution) in tax revenue.

An individual in generation t obtains utility from private consumption and a public consumption good g in old period, and the utility function of the young period t is given by linear form as:

$$u_t(c_{t+1}, g_{t+1}) = c_{t+1} + mg_{t+1}$$

where $m > 0$ is a parameter that represents the interest for the public good.

In each period, government levies lump-sum tax on the young and allocates tax revenue among transfer to the old ($\alpha\tau_t$), bond redemption and the public good supply. The latter is assumed to obey a ceiling constraint as follows:

$$g^* = g_t + (1+r)\,b_t, \tag{9.1}$$

as in Ihori and Itaya (2001) and Ihori (2006), g^* and b_t are, respectively, an exogenous institutional ceiling (in order to keep public expenditure under control) and the amount of issued government bonds, which may be held also by foreign investors, respectively. Because of the ceiling, the government can increase the supply of public good if and only if it reduces government bonds.

From the above assumptions, the evolution of public debt, which changes over time according to above expenditure rule, is as follows:

$$b_{t+1} = g_t + (1+r)b_t - (1-\alpha)\,\tau_t \tag{9.2}$$

Optimal Allocation

We derive the optimal solution of the benevolent social planner. Consider a government who wants to maximize at the initial period (t =*1*). The social planner's objective function is given by the following discounted sum of all generations' utility$\sum_t \zeta^{t-1}(c_t + mg_t)$ (t =1,…,∞), where $\zeta \in (0, 1)$ is social discount factor. Given the initial size of public debt $b_1 > 0$, the government maximizes the objective function subject to the resource constraint and the evolution of public debt (2). Assuming that individuals cannot borrow, since $\tau \in [0, w]$, the evolution of public debt provides the following range of public debt level in period:$g^* - (1-\alpha)\,w \leqslant b_{t+1} \leqslant g^*$. Socially optimal allocations are found by solving the following problem:

$$\max_{\tau_t} \sum_{t=1}^{\infty} \zeta^{t-1}(c_t + mg_t)$$
$$s.t.\, c_t = (1+r)(w - \tau_{t-1}) + \alpha\tau_t,$$
$$b_{t+1} = g_t + (1+r)b_t - (1-\alpha)\tau_t,$$
$$g^* = g_t + (1+r)b_t,$$

$$b_1 > 0,$$

The solutions of this problem are summarized by the following

$$b_{t+1} = \begin{cases} g^* if \frac{1}{1-\alpha}\{\zeta(1+r)-\alpha\} - \zeta m(1+r) > 0 \\ \gamma if \frac{1}{1-\alpha}\{\zeta(1+r)-\alpha\} - \zeta m(1+r) = 0 \\ g^* - (1-\alpha)\, w if \frac{1}{1-\alpha}\{\zeta(1+r)-\alpha\} - \zeta m(1+r) < 0 \end{cases}$$

where $\gamma \in [g^* - (1-\alpha)\, w, g^*]$. In these solutions, $b_{t+1} = g^*$ is not feasible, since public debt level in steady state becomes negative, $b_{social,cor,\max} = -\frac{g}{r}$, where subscripts refer to the value of the debt in a upper corner solution to the above social welfare maximization problem.

Thus, we restrict attention to the alternative corner solution and to the interior solution. In the corner solution, when $(1-\alpha)\, w \geqslant g, b_{t+1} = g^* - (1-\alpha)\, w$ is feasible, the steady state solution becomes $b_{social,cor,\min} = \frac{(1-\alpha)w-g}{r}$.

Moreover, interior solution can have a positive value in steady state. In the interior solution, government imposes tax along the following condition

$$\zeta(1+r)\, m(1-\alpha) = \zeta(1+r) + \alpha$$

This equation recalls the familiar Samuelson rule for the optimal provision of public goods: the left hand side represents marginal utility from public goods, while the right hand side represents marginal utility from consumption. In the interior solution, the steady state public debt level is $b_{social,int} = \frac{(1-\alpha)\tau^*-g}{r}$.

9.3 A Political Economic Analysis

In this section, we allow the government to hold office for just one period and to be able to redistribute income between generations. Thus, the goal of government is to maximize support among the currently living voters. As in Grossman and Helpman (1998) and Ono (2009), we endow each government with the following objective function:

$$G_t = (c_t + mg_t) + (c_{t+1} + mg_{t+1})$$

and restrict our attention to Markov-perfect equilibrium, as analyzed by Krusell et al. (1997), in order to show the feedback mechanism between the bond level and the future redistribution policy. When supporting a candidate's policy and choosing savings, individuals form expectations about future transfer, τ_{t+1}. In a Markovian equilibrium, this variable depends on a fundamental state variable, b_{t+1}. Thus, the policy expected for period t depends only on the value of the state variable expected at that time, $\tau(b_{t+1})$. Thus, a government choosing a fiscal policy at time t, looks ahead to time $t+1$, since policies introduced at time t will affect the disposable income of the young, when he or she consumes in the next period. Thus, each government must forecast future political outcomes. If we do not constrain these forecasts somehow, there will be many policy trajectories that are self-fulfilling. Based on this reasoning, the government at time t solves the following problem:

$$\begin{aligned}
&\max_{\tau_t, b_{t+1}} (c_t + mg_t) + (c_{t+1} + mg_{t+1}) \\
&s.t.\; c_t = (1+r)(w - \tau_{t-1}) + \alpha\tau_t, \\
&\quad c_{t+1} = (1+r)\,(w - \tau_t) + \alpha\tau\,(b_{t+1}) \\
&\quad b_{t+1} = g_t + (1+r)b_t - (1-\alpha)\,\tau_t \\
&\quad g^* = g_t + (1+r)\,b_t
\end{aligned}$$

We can rewrite this program as follows:

$$\begin{aligned}
&\max_{b_{t+1}} (1+r)\left(w - \tau_{t-1}\right) + \alpha\tau_t + m\left\{g^* - (1+r)\,b_t\right\} + (1+r)\,(w - \tau_t) + \alpha\tau\left(b_{t+1}\right) \\
&\quad + \left\{g^* - (1+r)\,b_{t+1}\right\} \\
&s.t.\; g^* - (1-\alpha)\,w \leqslant b_{t+1} \leqslant g^*
\end{aligned}$$

Also note that $\tau_t = \frac{b_{t+1} - g^*}{\alpha - 1}$ from the evolution of public debt and the ceiling constraint. The solution to this problem satisfies the following Kuhn–Tucker conditions:

$$b_{t+1} = \begin{cases} g^* if \frac{1}{1-\alpha}\{1+r-\alpha\} - m\,(1+r) + \alpha\tau' > 0 \\ \gamma if \frac{1}{1-\alpha}\{1+r-\alpha\} - m\,(1+r) + \alpha\tau' = 0 \\ g^* - (1-\alpha)\,w if \frac{1}{1-\alpha}\{1+r-\alpha\} - m\,(1+r) + \alpha\tau' < 0 \end{cases}$$

In order to determine the political equilibrium, we impose the following assumption.

Assumption 9.1 *(Ono 2008):*

$$\tfrac{1}{1-\alpha}\{1+r-\alpha\} > m\,(1+r)\,.$$

A political equilibrium satisfies the Kuhn–Tucker conditions given by the corner solutions and the interior solution. Firstly, consider the corner solutions. There are two possible corner solutions $\tau = w$ and $\tau = 0$.

Moreover, we can restrict the corner solution to $\frac{1}{1-\alpha}\{1+r-\alpha\} - m\,(1+r) + \alpha\tau' > 0$ by applying Assumption 9.1. Thus, the evolution of public debt in this corner solution is given by $b_{t+1} = g^*$, that is $\tau = 0$, which is a Markovian political equilibrium, because of $\tau_{t+1} = \tau_t = 0$. In this case, however, public debt level *cannot* converge toward the steady state $b_{pol,cor} \geqslant 0$, where $b_{pol,cor}$ represents public debt level in a steady state. Thus, in the corner solution, the political equilibrium strategy brings divergence and therefore inefficiency.

Next, consider an interior solution where government imposes a tax on the young but leave them some disposable income. From the concept of Markovian equilibrium, the government in period t forecasts redistributional transfer in period $t+1$. We express this functional form $\tau_{t+1} = \tau(b_{t+1}) \in (0, w)$ which obeys the following first order condition:

$$\tau'(b_t) = -\frac{1}{\alpha}\left\{\frac{1+r-\alpha}{1-\alpha} - m(1+r)\right\}.$$

Solution to this equation will be a condition for equilibrium tax functions. The above differential equation has the general solution:

$$\tau(b_t) = A - \frac{1}{\alpha}\left\{\frac{1+r-\alpha}{1-\alpha} - m(1+r)\right\} b_t \tag{9.3}$$

where A represents arbitrary values of the parameter. Note that square bracket is positive. Equation (18.3) means that tax rate in current period is decreasing in public debt at that period. This implies that consumption (by the old) and public debt level are positively correlated in each period and that governments are not faced with a simple trade-off between consumption and public debt within a period. When public debt accumulates, government expenditure has to decrease due to ceiling condition. Thus, government has to compensate this disutility by increasing consumption.

The reasons occurring trade-off between tax and public debt can be summarized as follows. Short-lived government imposes tax and issues public debt by considering the following three effects: income effect, ceiling effect and Markovian tax effect. When public debt accumulates, firstly, government increases tax and redistribution in period t. As a result, income decreases ("income effect"). Secondly, the public good decrease by the ceiling constraint ("ceiling effect"). Thirdly, the incumbent expects the next period government to follow the same behavior. Thus, since next period tax level is expected to decrease, the redistribution to next old, who is a current young, decreases ("Markovian taxation effect"). Therefore, the government debt level of the interior solution may be higher than the efficient level and that may lead to divergence in the interior solution.

The path of public debt in a political economy in absence of lobbying is derived by substituting (18.3) into (18.2) as in

$$b_{t+1} = g^* - A(1-\alpha) + \frac{1-\alpha}{\alpha}\left\{\frac{1+r-\alpha}{1-\alpha} - m(1+r)\right\} b_t \tag{9.4}$$

This political equilibrium path with an interior solution is a sequence of taxes and public debt $(\tau_t, b_t)_{t=1}^{\infty}$, such that given $b_1 > 0$ and $A \in R$, τ_t and b_t are determined by (18.3), (18.4) and the ceiling constraint. These three equations are implicitly rewritten as $b_t = b(b_1, A, t)$ and $\tau_t = \tau(b_1, A, t)$, respectively.

The public debt may be or not be sustainable in interior equilibrium if it satisfies $\tau_t = \tau(b_1, A, t) \in (0, w)$ for all t. Given that $\tau_t = \tau(b_1, A, t)$ is linearly related to public debt, as expressed by (3), any explosive path of public debt means this path is not sustainable. A sustainable equilibrium path, if it exists, must be characterized by monotone convergence This condition of monotonic convergence is $0 < \frac{1-\alpha}{\alpha}\left\{\frac{1+r-\alpha}{1-\alpha} - m(1+r)\right\} < 1$. Thus, public debt convergence to

$$b_{pol,int} = \frac{A(1-\alpha) - g}{r + \frac{1-\alpha}{\alpha}\left\{\frac{1+r-\alpha}{1-\alpha} - m(1+r)\right\}}$$

where $b_{pol,int}$ denotes the steady-state level of political economic solution with an interior solution.

Proposition 9.1 *Suppose that a set of parameters satisfies* $\frac{1}{1-\alpha} < m < \frac{1+r-\alpha}{(1-\alpha)(1+r)}$. *There exists an interior, converging political equilibrium. The public debt in a steady state is given by* $b_{pol,int} = \frac{A(1-\alpha)-g}{r+\frac{1-\alpha}{\alpha}\left\{\frac{1+r-\alpha}{1-\alpha}-m(1+r)\right\}}$.

Characterization of Political Equilibrium

As in Ono (2009), we can compare the political equilibrium with the dynamically optimal allocation and examine the normative implications of the political equilibrium. In undertaking the above analysis, we can restrict our focus to the situation where the set of parameters satisfies the following two conditions: (1) the characterization of a dynamically optimal allocation in Proposition 9.1, (2) the condition for the existence of the corner solutions. Figure 9.1 represents these conditions and the corresponding characterization of the dynamically optimal allocation and political equilibrium.

The figure implies that the normative aspect of the political equilibrium depends on the social discount factor represented by ζ as well as the interest of public goods represented by m. To be more precise, we focus on the discount factor ζ and consider low and high cases in turn. First, consider the case of a high discount factor such that $\frac{1+r-2}{(1-)(1+\mathrm{r})} < \zeta < \frac{1+r-}{(1-)(1+\mathrm{r})}$. The political equilibrium allocation experiences over-accumulation of public debt because social planner imposes tax higher than short-lived government. Secondly, consider the case of a low discount factor $\zeta < \frac{1+r-2}{(1-)(1+\mathrm{r})}$. The political equilibrium allocation experiences under-accumulation of public debt. Thus, in general, the solution in the political economy does not match with the dynamically optimal allocation.

Sustainability of Public Debt in Political Equilibrium in Absence of Lobbying

Under above analysis, consider the sustainability of public debt in political equilibrium in absence of lobbying. When we derive the Markovian equilibrium solution, we find that there are two possible solutions. Firstly, in a corner solution, $\tau_t = 0$, all short-lived politician have to do is to decrease that tax rate, since the "income effect" is always larger than the "ceiling effect." And even if politician can impose tax in interior solution, there is decreasing pressure for tax rate in each period. This is the "trust effect" which means that each politician wants to get trust in office. Thus, on the one hand, *in short-run,* short-lived politician can maximize each current lived

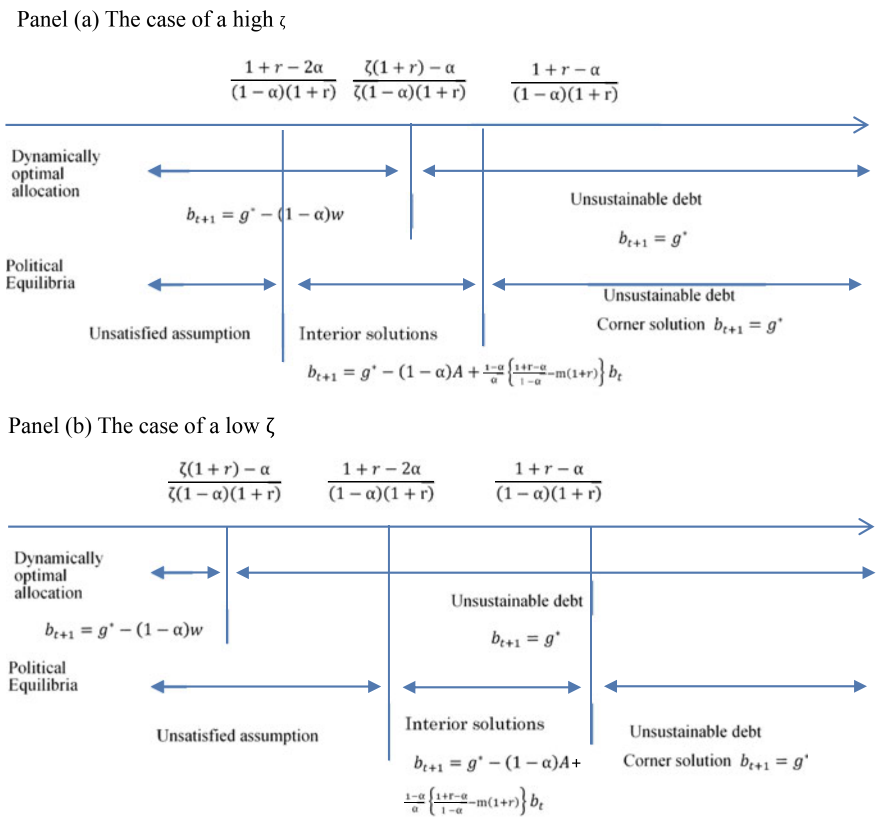


Fig. 9.1 Comparison of the dynamically optimal allocation and the political equilibria in the absence of lobbying

generation's utility, on the other hand, *in long-run*, since tax rate continue to decrease, public debt tends to diverge or over accumulate.[2]

9.4 Lobbying

In this stage, we examine how campaign contribution giving by young can change the taxation policy. Therefore, we assume that each short-lived government maximizes a political support index that consists of the utility levels of the two generations of voters, young and old, and the total amount of contribution; namely, in period t, the government seeks to maximize the following linear political support function:

[2]This result corresponds to empirical result of Buti and van den Noord (2003). They show that, when election, government expenditure tends to expand in developed country.

$$G_t = (c_t + mg_t) + (c_{t+1} + mg_{t+1}) + \varphi z_t \tag{9.5}$$

where z_t and φ denote the total amount of political gifts collected from the contributing lobby and the weight that the government attaches to a dollar of contribution relative to a dollar of additional consumption by old, respectively. We impose the following assumption that restricts the range of the parameter φ.

Assumption 9.2 (Grossman and Helpman 1998; Ono 2009)

$$\varphi > 1 + r$$

This assumption implies that the government values contributions more than the benefits accruing to either one of its constituents by spending an equal amount of resources. An increase in the contribution by one unit leads to an increase in utility from the contribution by φ units and to a decrease in utility by $1 + r$ units. Hence, without Assumption 9.2, politicians would never accept any gifts from lobbyist.

The lobby offers a tax contingent contribution schedule to the incumbent in order to induce the latter to choose a tax favorable to the lobby. When government is not lobbied, it would choose a tax which maximizes $(c_t + mg_t) + (c_{t+1} + mg_{t+1})$ and thus can always achieve a reservation level of political support equal to $G^*(b_t)$. Campaign contributions are made to compensate the incumbent for departing from the social welfare maximization; namely "the minimum contribution that induces the policy leaves the short-lived politician just indifferent between accepting and rejecting the lobby's offer" (Grossman and Helpman 1998, p. 1316). Unless following constraint is not satisfied, government does not receive campaign contribution:

$$\begin{aligned} G^*(b_t) = {} & ((1 + r)(w - \tau_{t-1} - z_{t-1}) + \alpha\tau_t + mg_t) \\ & + ((1 + r)(w - \tau_t - z_t) + \alpha\tau(b_{t+1}) + mg_{t+1}) + \varphi z_t . \end{aligned}$$

Under this participation constraint, the lobby in period t maximizes the utility of the young, $u_t = (1 + r)(w - \tau_t - z_t) + \alpha\tau(b_{t+1}) + mg_{t+1}$. The lobby's problem is given by:

$$\begin{aligned} & \max_{\tau_t} \; (1 + r)(w - \tau_t - z_t) + \alpha\tau(b_{t+1}) + mg_{t+1} \\ & s.t. \;\; b_{t+1} = g_t + (1 + r)b_t - (1 - \alpha)\tau_t, \\ & \qquad g^* = g_t + (1 + r)b_t \\ & and \\ & \qquad G^*(b_t) = ((1 + r)(w - \tau_{t-1} - z_{t-1}) + \alpha\tau_t + mg_t) \\ & + \quad ((1 + r)(w - \tau_t - z_t) + \alpha\tau(b_{t+1}) + mg_{t+1}) + \varphi z_t \end{aligned}$$

Corner Solution Under Lobbying

As for the political equilibrium in absence of lobbying, in the presence of lobbying there are two possible solutions for a political equilibrium: a corner solution and an interior solution. Firstly, consider the corner solution. Notice that $\frac{\partial z_t}{\partial \tau_t} =$

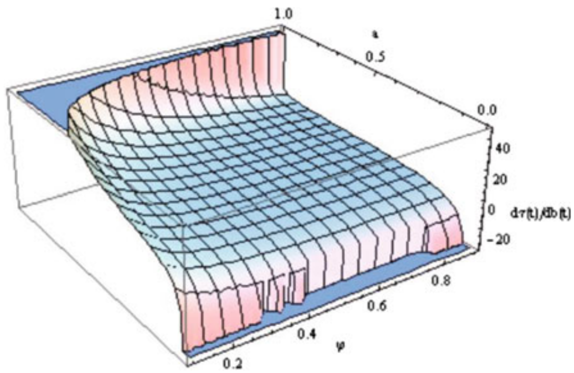


Fig. 9.2 The effect on tax of an increasing debt; $m = 0.5$, $r = 0.9$

$\frac{1}{\varphi-(1+r)}\left[1+r-\alpha-m\left(1+r\right)\left(1-\alpha\right)\right]>0$ from first-order condition implies $\frac{\partial u_t}{\partial \tau_t}$ $= (1+r)\left\{(1-\alpha)\,m-1-\frac{\partial z_t}{\partial \tau_t}\right\} < 0$. Since any increase in tax decreases the utility of the lobbying young, the solution is $\tau_t = 0$. In this case, because of $b_{t+1} = g^*$, the public debt diverges.

Interior Solution Under Lobbying

Considering the interior solution, the first order condition obeys following differential equation:

$$\tau'\left(b_t\right) = \frac{(1+r)\left[\varphi\left\{m\left(1-\alpha\right)-1\right\}+\alpha\right]}{\alpha\left(1-\alpha\right)\varphi}. \tag{9.6}$$

The solution to this equation is needed to derive the equilibrium tax function. The differential equation has the following general solution:

$$\tau\left(b_t\right) = B + \frac{(1+r)\left[\varphi\left\{m\left(1-\alpha\right)-1\right\}+\alpha\right]}{\alpha\left(1-\alpha\right)\varphi} b_t \tag{9.7}$$

where B represents arbitrary values. Equation (18.7) indicates that the relationship between tax rate and public debt depends on the size of interest of campaign contribution, φ, and on the size of the redistribution parameter, α, in contrast to the political equilibrium in absence of lobbying (18.3). Figure 9.2 shows the effect on tax of increase debt. Higher α and lower φ increase the tax rate.

In addition to the effects in the previous section, there is now a *campaign contribution effect*. In Fig. 9.2, whether tax is increasing or not depends on the values of α and φ. Consider the case of increasing tax. When the government has high interest for campaign contributions, the income effect overtakes the ceiling and the Markovian tax effects. However, since the lobby wants to compensate for this effect, it induces lower tax rate by increasing campaign contributions.

In contrast, in the case of higher weight to redistribution, the lobbyist can concede a current high tax burden in exchange of more advantageous redistribution next period.

Evolution of Public Debt in Presence of Lobbying

Since the corner solution cannot sustain economy in the long-run, we can concentrate on the interior solution. Substituting (18.7) into evolution of public debt (18.2), we can derive the path of accumulating public debt in presence of lobbying:

$$b_{t+1} = g^* - (1-\alpha)\,B - \frac{(1+r)\,[\varphi\,\{m\,(1-\alpha)-1\}+\alpha]}{\alpha\varphi} b_t \tag{9.8}$$

Given the initial condition $b_1 > 0$ and the parameter $B \in R$, an interior political equilibrium path $\{\tau_t, b_t\}_{t=1}^{\infty}$ is characterized by (18.7) and (18.8). The tax rate in interior equilibrium is feasible if it satisfies $\tau_t = \tau\,(b_1, B, t) \in (0, w)$ for all t. In contrast to a political equilibrium with absence of lobbying, a sustainable equilibrium path has two convergence patterns.

The Short-Run Effect

We consider the short-run effect in an accumulating process of debt. In order to show this effect, Differentiating (18.8) with respect to b_t yields:

$$\frac{\partial b_{t+1}}{\partial b_t} = -\frac{(1+r)\,[\varphi\,\{m\,(1-\alpha)-1\}+\alpha]}{\alpha\varphi} \quad \ldots\ldots\ldots.. \tag{9.9}$$

The sign of this depends on the magnitude of the interest of campaign contribution, φ, and redistribution parameter, . In the case of $0 < \frac{(1+r)[\varphi\{1-m(1-\alpha)\}-\alpha]}{\alpha\varphi} < 1$ that corresponds to $\tau'\,(b_t) > 0$public debt shows monotone convergence. However, when $-1 < \frac{(1+r)[\varphi\{1-m(1-\alpha)\}-\alpha]}{\alpha\varphi} < 0$, public debt shows cyclical convergence. These condition means that when r is high,$(1+r)\,[\varphi\,\{1-m\,(1-\alpha)\}-\alpha] > \alpha\varphi$, and public debt diverges.

The Long-Run Effect

When above stability conditions are satisfied, public debt in steady-state converges to

$$b_{lobby,int} = \frac{\{(1-\alpha)\,B - g\}\,\alpha\varphi}{\alpha\varphi r + (1+r)\,[\varphi\,\{1-(1-\alpha)\,m\}-\alpha]} \tag{9.10}$$

where $b_{lobby,int}$ denotes the interior solution of political economy in presence of lobbying. In order to have positive $b_{lobby,int}$, we need some conditions to be satisfied. In the case of monotone convergence, $(1-\alpha)\,B > g$ is a necessary condition.

Proposition 9.2 *Under Assumptions 9.1 and 9.2, suppose that a set of parameters satisfies:*

(i) $0 < (1+r)[\varphi\{m(1-\alpha)-1\}+\alpha] < \alpha\varphi$, *then there exists an interior, monotonic converging political equilibrium;*

(ii) $-\alpha\varphi < (1+r)[\varphi\{m(1-\alpha)-1\}+\alpha] < 0$, *then there exists an interior, cyclical converging political equilibrium.*

The steady-state public debt is given by $b_{lobby,int} = \frac{\{(1-\alpha)B-g\}\alpha\varphi}{\alpha\varphi r+(1+r)[\varphi\{1-(1-\alpha)m\}-\alpha]}$.

From the previous analysis, we derive few insights about the sustainability of public debt, in political equilibrium. In presence of lobbying, in a corner solution, $\tau_t = 0$: a short-lived politician can just decrease the tax rate, in addition to three effects in political economic equilibrium, there is a "campaign contribution effect". Looking at the interior solution, in contrast to the political equilibrium in absence of lobbying, now tax is not always a decreasing function of public debt in that period. In case of high interest of the incumbent for campaign contributions, as public debt accumulates, a short-lived government decreases tax. Thus, public debt converges to a higher level or diverges.

However, in the case of an incumbent's bias for redistribution, as public debt accumulates, the government increases tax. Thus, public debt converges to a lower level in contrast to what we could expect. The intuition is as follows. The lobbyist considers the tradeoff between higher taxes in current period and higher transfer in next period. Then, if the government assigns a high weight to redistribution, the lobbyist may concede to a current increasing tax because (s)he will benefit from a higher transfer in the next period.

To show the above intuition, we differentiate (18.8) with respect to φ:

$$\frac{db_{lobby,int}}{d\varphi} = \frac{\alpha^2\{g-(1-\alpha)B\}(1+r)}{[(m-1)(1+r)\varphi - \alpha\{(1+r)m\varphi-(1+r)+r\varphi\}]^2}$$

In the case of monotone convergence and in the case of cyclical convergence for $\alpha\varphi r + (1+r)[\varphi\{1-(1-\alpha)m\}-\alpha] > 0$, higher φ decreases public debt level. However in the case of cyclical convergence under $\alpha\varphi r + (1+r)[\varphi\{1-(1-\alpha)m\}-\alpha] < 0$, higher φ actually increases the public debt level.

9.5 Concluding Remarks

This study shows how the short term political goal of re-election of an incumbent government may have a perverse impact on debt accumulation, pushing the latter above the efficient level. Remarkably, lobbying may reverse this outcome. The reason is that, in our model, government cares of the transfers received by the lobbying subject, the young, as well as of social redistribution. If the government has a relative bias for redistribution, it can be 'trusted' about the fact that higher taxes today on the young imply larger benefits when the same individual retires. This redistributive bias, allows the lobbying subject to concede higher taxes (and lower debt) in exchange of future period benefits. This result seems consistent with the observation of economies with high redistribution and taxes and relatively contained public debt, such as the Scandinavian countries for example.

The analysis relies upon some restrictions that have been used to improve the tractability of the model. Future extensions should consider relaxing the small country's assumption, which nullifies the indirect effects on interest rate. Also wages have been taken as fixed, for simplicity. As for the political economic analysis, an important generalization would be to allow lobbying by the old. This extension would provide useful insights in societies where elders account for a large and growing share of population.

References

Buti, Mario, and Paul van den Noord. (2003). Discretionary fiscal policy and elections: the experience of the early years of EMU, OECD Economics Department Working Paper, No. 351.

Drazen, A. 2000. *Political economy in macroeconomics*. Princeton: Princeton University Press.

Grossman, G.M., and E. Helpman. 1994. Protection for sale. *American Economic Review* 84: 833–850.

Grossman, G.M., and E. Helpman. 1998. Intergenerational redistribution with short-lived governments. *The Economic Journal* 108: 1299–1329.

Ihori, T. 2006. Fiscal policy and fiscal reconstruction in Japan. *International Tax and Public Finance* 13: 489–508.

Ihori, T., and J. Itaya. 2001. A dynamic model of fiscal reconstruction. *European Journal of Political Economy* 17: 779–797.

Krusell, P., J. Rios-Rull, and V. Quadrini. 1997. Politico-economic equilibrium and economic growth. *Journal of Economic Dynamics and Control* 21: 243–272.

Ono, T. 2009. The political economy of environmental and social security policies: the role of environmental lobbying. *Economics of Governance* 10: 261–296.

Chapter 10
Intergenerational Living Arrangements and Labor Supply of Married Women

Hyunbae Chun, Olivia Hye Kim and Injae Lee

10.1 Introduction

Social norms and traditions matter in the pattern of intergenerational living arrangements. The nuclear family is the rule in Western countries, whereas the multigenerational household has been commonplace in East Asian countries such as Korea, China, and Japan. In these countries, it has been customary for married children, especially the eldest sons and their wives, to live with their parents. This custom arises from deep-rooted Confucian ideology that emphasizes the importance of maintaining the integrity of the generational lineage (Takagi and Silverstein 2006). Although the percentage of multigenerational households has declined precipitously, traditional family values and practices are still prevalent in various aspects of social life.

Intergenerational living arrangements in East Asian countries may have a significant impact on the labor supply of married women. Some aspects are of particular interest. On one hand, the aged parent plays the role of a built-in childcare system in the traditional family structure (Freedman et al. 1982). Living with parents can thus be a substitute for childcare and other household services. Coresiding parents-in-law may lessen a married woman's burden of household work and thereby allow her more time for market work. On the other hand, the intergenerational household also entails informal care of elderly parents (Ogawa and Ermisch 1996). As substitutes for weak public support of the elderly, married women have served as the main providers of informal elderly care services for their husbands' parents. If married women spend a substantial amount of time caring for their parents-in-laws, coresidence may re-

H. Chun (✉)
Sogang University, Seoul, South Korea
e-mail: hchun@sogang.ac.kr

O. H. Kim
Seoul National University, Seoul, South Korea

I. Lee
Incheon National University, Incheon, South Korea

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_10

duce the labor supply of married women. Therefore, the effect of coresidence with parents-in-law on the labor supply of married women is ambiguous.

Over the past three decades, the labor force participation rate of married women has risen rapidly in East Asian countries. However, a substantial percentage of married women is still concentrated in non-market work as full-time homemakers. The employment-to-population ratios among women aged 15–64 in Korea has been below the OECD average ratio until now. This gap can be largely attributed to the low employment rate of married women. Researchers and policymakers in the East Asian region have questioned whether traditional living arrangements may affect the labor force participation of married women (Sung and Chah 2001; Sasaki 2002; Mano and Yamamura 2011).

This study explores this question empirically. Using data from the seventh wave of the Korean Labor and Income Panel Study (KLIPS), we estimate the impact of coresidence with the husband's parents on the labor supply of married women in Korea. We use an instrumental variable estimation to identify the causal effect of intergenerational living arrangements on the labor supply of married women. Because coresidence is an option that married women may choose, their labor supply and coresidence with their husbands' parents may be determined simultaneously. Empirical models that treat each decision as exogenous would give biased estimates of the effect.[1]

To account for this endogeneity problem, we use the husband's birth order as an instrumental variable (IV) for living arrangements. Intergenerational household structure is greatly affected by cultural preferences. As mentioned above, in East Asian countries such as Korea, where traditional family values and practices still persist, the cultural preference for living arrangements is that the eldest son or a son of lower birth order (a son born relatively early in the birth order) lives with his parents. This cultural prescription of intergenerational living arrangements allows us to use the husband's birth order as an instrument for coresidence with the husband's parents.

The main finding of this study is the absence of any causal effect of coresidence with the husband's parents on the labor force participation of married women. The positive effect in OLS estimation disappears when we control for the endogeneity of the coresidence decision. Our instrumental variable estimation results suggest that this positive effect can be explained by a sorting process, in which married women oriented toward market work are more likely to coreside with their parents-in-laws.

The study is organized as follows. Section 10.2 describes the data used in this study. Section 10.3 presents the birth order instrument to control for the endogeneity of intergenerational living arrangements in the labor supply decision of married women

[1]Related US studies have mainly focused on the relationship between parental care and employment decisions of adult daughters. Early studies were generally based on models of labor market participation and parental care that treated each decision as exogenous (Stone and Short 1990; White-Means 1992). Recent research has recognized the joint nature of these decisions and used instrumental variables to address the endogeneity problem (Wolf and Soldo 1994; Ettner 1995, 1996; Pezzin and Schone 1999). The results from these studies provide inconclusive evidence of labor supply in response to competing demands for women's time.

and analyzes its validity. Section 10.4 presents the estimation results. Section 10.5 provides conclusions.

10.2 Data

This study uses data from the Korean Labor and Income Panel Study (KLIPS). The KLIPS is a longitudinal survey of a representative sample of Korean households. The original cohorts at the start of the survey in 1998 were 13,321 individuals from 5,000 households. The KLIPS is designed to follow the basic structure of the Panel Study of Income Dynamics (PSID) and is supplemented with questionnaires related to labor market activity from the National Longitudinal Survey (NLS). The KLIPS collects information on the sampled households and their individual members aged 15 or older.

We use information on 4,592 households and 11,543 individual household members collected in the seventh wave of the KLIPS in 2004.[2] The KLIPS is conducted annually to collect information on the basic characteristics, income, and expenditures of the sampled households and on the personal characteristics and labor market activities of individual household members. In addition to a wide array of demographic and socioeconomic variables, the KLIPS also provides detailed information on the family structure and intergenerational relationships of household members.

Our sample consists of married women aged 25 to 55 with a spouse present. Thus, respondents who are divorced, separated, or widowed are excluded from the sample. We set age 25 as a lower bound to account for the high rate of female college enrollment and its impact on marriage decisions. A sample restricted to women at younger ages might tend to omit those who postpone marriage until completing their education. Our choice of age 55 as an upper bound corresponds to labor market detachment for working women in Korea. Working women aged 56 or above make up only 16% of all female workers in 2004. Using a wider span of ages (for example, age 20 to 60) for our sample neither significantly increases the sample size, nor affects the main results of this paper.

To identify types of living arrangements, we need information on the family relationships within households. Like the Current Population Survey (CPS) data in the United States, the KLIPS data are constructed in such a way that multiple families, related or unrelated, may be intermingled in a household unit. Therefore, exact matching of the husband and wife as well as married children and parents is required in order to construct a sample. We first use household and spousal identification information to find all possible matches of a husband and wife. Then, using information on family relationships, we identify whether a married son lives with his parents.

This study focuses on the relationship between labor supply of married women and coresidence with their husbands' parents. We exclude married women living with

[2]Refer to footnote 5 for the reason why we use the mid-2000 data rather than recent ones in this paper.

Table 10.1 Descriptive statistics: married women aged 25–55

Variable	Total	Working	Non-working
Age	40.204 (7.768)	41.549 (7.301)	38.798 (7.994)
Education			
Middle school or lower	0.256 (0.437)	0.319 (0.466)	0.191 (0.393)
High school	0.476 (0.500)	0.446 (0.497)	0.507 (0.500)
College or higher	0.268 (0.443)	0.235 (0.424)	0.303 (0.460)
Log of annual household income (excluding wife's earnings)	7.710 (0.822)	7.589 (0.866)	7.836 (0.752)
Home ownership (= 1 if owns home; = 0 otherwise)	0.634 (0.482)	0.647 (0.478)	0.620 (0.486)
Self-employed husband (= 1 if self-employed; = 0 otherwise)	0.200 (0.400)	0.249 (0.433)	0.149 (0.356)
Number of children			
Total number of children	1.153 (0.947)	1.023 (0.941)	1.290 (0.934)
Number of children aged 6 or below	0.408 (0.672)	0.232 (0.520)	0.592 (0.759)
Coresidence with husband's parents (= 1 if coresiding; = 0 otherwise)	0.102 (0.303)	0.124 (0.329)	0.079 (0.270)
Sample size	2197	1123	1074

Notes Data are expressed as means (standard deviations)

their own parents. Traditionally, most extended families are patrilineally composed in Korea. Our sample also shows that most coresiding married women live with their husbands' parents. Among the 2,217 matches of husband and wife, only 20 married women coreside with their own parents (0.90% of the total sample and 8.20% of the coresidence sample).[3]

After imposing restrictions and omitting married women with missing relevant information, we obtain 2,197 sample observations. Table 10.1 presents descriptive statistics of the sample according to the labor market status of married women. In the sample, 51.1% of married women participate in the labor market, whereas 48.9% do not work outside the home. Married women who live with their husbands' parents make up 10.2% of the sample. The remaining 89.8% live independently.

Married women who participate in the labor market are more likely to coreside with their husbands' parents than are their non-working counterparts. Of working married women, 12.4% coreside with their husbands' parents, whereas 7.9% of non-working married women coreside with their husbands' parents.

[3]Including married women who live with their own parents in the sample does not change the main results of this paper.

Table 10.1 also shows some differences in economic and socio-demographic characteristics between families in which married women are full-time homemakers and those in which married women participate in the labor market. First, homemakers are on average younger than working married women, and they are more likely to have more children, especially children aged between 0 and 6 years. This corresponds to the well-known pattern of female labor supply across age in Korea. Female labor force participation tends to drop among married women around childbearing ages, and rises among women in their late 30s and mid-40s when the burdens of childbirth and childcare lessen (Lee et al. 2008). Second, married women who participate in the labor market tend to have lower educational attainment and slightly lower total annual household income. They are more likely to have self-employed husbands and to own their home.

10.3 Empirical Framework

Labor Supply and Living Arrangements

To investigate the effect of coresidence with the husband's parents on the labor supply of married women, we estimate the following labor supply equation:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{10.1}$$

where Y_i is a measure of labor supply for a married woman i, X_i is a dummy variable equal to 1 if a married woman lives with her husband's parents, and ε_i is the error term. The coefficient of interest is β_1. If coresidence with husband's parents increases the labor supply of married women, then $\beta_1 > 0$.

In estimating Eq. (10.1), we face an endogeneity problem. Coresidence with the husband's parents would arguably be endogenous to the labor supply of married women. Unobservable attributes of married women that affect coresidence might be associated with their unmeasured propensity to participate in the labor market. This correlation between the error term in labor supply and the coresidence variable makes the OLS estimate of β_1 biased.

For example, married women who want to work in the labor market may be more likely to reside with their husbands' parents if coresiding parents can provide services for home production such as childcare. In this case, one might find a positive relationship, even in the absence of a causal effect of coresidence on labor supply. Alternatively, the opposite may be true. If married women who have strong tastes for non-market work are also more inclined to coreside with their husbands' parents, one might find a negative correlation between labor supply and coresidence. In both of these cases, one would find a spurious relationship between labor supply and coresidence. The direction of the bias cannot be determined a priori. OLS estimates can be biased upward or downward depending on whether the error term is positively or negatively correlated with the coresidence variable.

To address the endogeneity problem of living arrangements, we use an instrumental variable approach. The first-stage coresidence equation is expressed as

$$X_i = \gamma_0 + \gamma_1 Z_i + \nu_i \tag{10.2}$$

where Z_i denotes an instrument for living arrangements and ν_i is the error term. An ideal instrument should be correlated with living arrangements but uncorrelated with all other factors that determine labor supply. We suggest husband's birth order as an instrument and describe its validity in the next section.

Husband's Birth Order as an Instrument

Considerable differences in terms of family ties exist between Western countries and East Asian countries. The social norm in Western Europe and the United States has always been for the child to leave home before or for marriage (Aquilino 1990; Reher 1998).[4] The traditional East Asian family has been described as a unit with strong patriarchal authority and patrilineal descent, in which one of the sons marries and continues to live with his parents while the other sons and daughters marry and leave the family unit (Logan et al. 1998).

Intergenerational family relationships involve many and sometimes contradictory elements. However, Logan et al. (1998) point out two findings from research on coresidence in the United States and other Western countries. First, coresidence is childcentered. The usual direction of assistance is from parents to children. Although some parental characteristics, such as physical disability and disease, increase the likelihood of living with children (Wolf and Soldo 1988), parents' needs do not predict coresidence as well as do children's needs. Another central aspect of the Western pattern is its gender neutrality. Family theorists usually argue that daughters are assigned the key role in intergenerational support. But many studies report mixed effects of child's gender on coresidence (Wolf and Soldo 1988; Ward et al. 1992).

In East Asian countries, the traditional pattern of coresidence is parent centered (Logan et al. 1998). This implies that parental needs and preferences are more important than child needs, fulfilling the deeply ingrained moral mandate that children should be caregivers for their elderly parents. The traditional pattern is also highly gendered, with a strong preference for living with sons rather than daughters (Asis et al. 1995).[5] Coresidence depends partly on the number of children, but the number of sons is more important than the number of daughters. Coresidence with a daughter is more likely in the absence of a son. In addition to preferences with respect to the

[4]Even in Western countries, patterns in intergenerational living arrangements show some differences. Young adults in North American and Northern European countries leave home before marriage, whereas young adults in Southern European countries stay at home until marriage (Manacorda and Moretti 2006; Giuliano 2007). Nonetheless, most married children in Western countries do not live with their parents.

[5]Son preference in childbearing is also prevalent in China (Arnold and Zhaoxiang 1986) and Korea (Park and Cho 1995). However, son preference in Korea has weakened during the past three decades and the sex ratio at birth has reached the natural level after 2010 (Choi and Hwang 2015). It is the main reason why this paper uses the sample in the mid-2000s.

gender of the coresiding child, cultural prescriptions may also exist with respect to the position of the son within the sibling set. Traditionally, coresidence with a son of a higher birth order (a son born later in the birth order) is less likely, and importance is placed on the eldest son (Asis et al. 1995).

The shift away from the traditional family structure in East Asian countries would involve a movement toward a child-centered and a more gender-neutral pattern. Studies on the levels of and trends in coresidence in various East Asian countries suggest both a close bond between generations and a clear trend away from established patterns in this region. New familial relationships emerge as traditional values and expectations give way to changing economic circumstances. However, even if coresidence patterns shift at a rapid rate, people's choices about when and with whom to reside are still influenced by traditional norms and values (Takagi and Silverstein 2006).[6]

The mixture of traditional norms and emerging values in the pattern of intergenerational coresidence is well documented in our data. Only 10.2% of the married children in our sample reside with their parents. According to the Population and Housing Census in Korea, three generation households or parents and married children households accounted for about 17.1% of total households in 1980, but had declined to 7.9% in 2005.[7] This drastic change clearly shows a shift away from the traditional coresidence pattern. A family in which married children live with their parents is no longer the representative family in Korea.

However, traditional norms and values still have a strong effect on the pattern of coresidence. First, most married couples who live with parents reside with the husband's parents. Of the 246 sampled couples living with parents, 226 couples live with the husband's parents and 20 couples live with the wife's parents. This indicates a strong cultural preference that parents live with sons rather than with daughters. Another prevalent traditional norm and value in coresidence pattern is the importance of birth order. A son coresiding with his parents is more likely to be the eldest son or a son of a low birth order.

Recent studies show that family decisions about intergenerational living arrangements are affected by the birth order of the husband (Ogawa and Ermisch 1996; Sasaki 2002; Wakabayashi and Horioka 2006). The exogeneity of husband's birth order and its relationship to intergenerational living arrangements allow us to use the husband's birth order as an instrument for coresidence. Using the husband's birth order, we can construct several instruments for coresidence with the husband's parents. These instruments are a dummy variable indicating whether the husband is the eldest son; or the husband's birth order *among siblings*.

Among these three possible instruments, the first may not be free from the problem of endogeneity. The eldest son is more likely to adhere to the traditional family norms

[6]Asai et al. (2015) show that childcare availability has no effect on maternal employment because childcare has rapidly shifted from informal to formal ones.

[7]Because the Population and Housing Census of Korea includes single-member households, the percentage of coresidence is slightly lower than that in our KLIPS sample. If we exclude single-member households from the Census, the coresidence rate is 9.8%, which is very close to the rate in our sample.

Table 10.2 Husband's birth order and probability of coresidence with husband's parents

Husband's birth order	Total		Non-Coresidence	Coresidence	Probability of coresidence (%)
	Frequency	Relative frequency (%)	Frequency	Frequency	
1	639	29.09	540	99	15.49
2	473	21.53	419	54	11.42
3	374	17.02	339	35	9.36
4	289	13.15	275	14	4.84
5	215	9.79	199	16	7.44
6	105	4.78	102	3	2.86
7	72	3.28	69	3	4.17
8 or above	30	1.37	30	0	0.00
Sample size	2197		1973	224	

and values than are his brothers of higher birth order. If positive assortative mating occurs in the marriage market, a woman who marries an eldest son is more likely to share the same values as her husband and to perform traditional roles within her household. Therefore, the dummy for the eldest son can be correlated with the labor supply of married women. One obvious solution is to exclude eldest sons from the sample and to use the husband's birth order among brothers. In this case, we restrict the sample to sons of second or higher birth order. However, most families in our sample have fewer than four brothers. When eldest sons are excluded, the husband's birth order among brothers does not have enough variation and thus turns out to be a weak instrument. Thus our choice in this paper is to *exclude eldest sons* from the sample and then to use the husband's birth order *among siblings* as an instrument.

Table 10.2 shows the relationship between the husband's birth order among siblings and the probability of coresidence in the sample. The higher the husband's birth order is, the lower the probability of coresidence becomes. The probability of coresidence for the eldest sons is 15.49%, which is the highest among the sons. The coresidence rate declines from 11.42% for sons of second birth order to 4.84% for sons of fourth birth order. Given that eldest sons account for about 29% of the sample and second to fourth birth order sons together account for about 52%, the monotonically decreasing pattern is observed for more than 80% of the whole sample. Furthermore, a strong negative relationship between coresidence and birth order is also evident in the sample that excludes eldest sons. This clearly indicates that the variation in the probability of coresidence associated with birth order is not confined to a small fraction of the sample. Thus, Table 10.2 suggests that all of the husband's birth order variables mentioned above can be relevant instruments. Statistical tests of the relevance of these instruments, described in the next section, also confirm this conclusion.

Table 10.3 Correlations between husband's birth order and characteristics of married women

Variable	Whole Sample		Sample without eldest sons
	(1) Dummy for the eldest son	(2) Husband's birth order	(3) Husband's birth order
Age	0.002 (0.931)	0.005 (0.831)	0.009 (0.710)
Education			
Middle school or lower	0.008 (0.723)	−0.011 (0.621)	−0.009 (0.725)
High school	−0.048** (0.024)	0.050** (0.020)	0.028 (0.269)
College or higher	0.047** (0.028)	−0.046** (0.033)	−0.023 (0.360)
Log of annual household income (excluding wife's earnings)	−0.050** (0.019)	0.027 (0.204)	−0.011 (0.678)
Home ownership (= 1 if owns home; = 0 otherwise)	0.002 (0.934)	−0.021 (0.315)	−0.032 (0.204)
Self-employed husband (= 1 if self-employed; = 0 otherwise)	−0.040* (0.061)	0.031 (0.150)	0.007 (0.797)
Number of children			
Total number of children	−0.018 (0.399)	0.031 (0.150)	0.030 (0.242)
Number of children aged 6 or below	0.027 (0.199)	−0.042** (0.048)	−0.038 (0.131)
Sample size	2197	2197	1558

Notes p-values are reported in parentheses
*Significant at the 10% level
**Significant at the 5% level
***Significant at the 1% level

To be a valid instrument, an instrument should be exogenous to the labor supply equation. However, the exogeneity of a single instrument cannot be statistically tested. As in the study of Angrist and Evans (1998), we check whether husband's birth order is randomly assigned for characteristics that can affect the labor supply of married women. These characteristics are included as omitted variables in error terms for the univariate labor supply regression. Table 10.3 shows the correlation between the husband's birth order and the determinants of the labor supply of married women.

As expected, a dummy variable for the eldest son is correlated with the determinants of labor supply of married women. Column (1) of Table 10.3 shows that the eldest son dummy has statistically significant correlations with the variables related to the wife's educational attainment. An eldest son is less likely to marry a high school graduate and is more likely to marry a woman with a college or higher degree. The results also indicate that the eldest son dummy is negatively correlated with household income at the 5% significance level. Column (2) of Table 10.3 shows the correlations between the husband's birth order among siblings and characteristics related to the wife's labor supply. In addition to the wife's educational attainment, the husband's birth order among siblings is also negatively correlated with the number of children aged 6 or below at the 5% significance level. These statistically significant correlations in the whole sample imply that the factors that affect a woman's decision to marry an eldest son or a son of a lower birth order can also have an effect on her decision to work in the labor market.

A woman who is married to the eldest son may be more likely to share the same traditional norms and values as her husband. To the extent that such values are manifested in coresidence, her unobservable propensity to coreside may be correlated with the propensity to participate in the labor market. This suggests another possible instrument for coresidence: the wife's birth order. Interestingly, a woman with a lower birth order is more likely to coreside with her husband's parents, and this correlation is statistically significant at the 5% level with and without controlling for her husband's birth order. Thus, the wife's birth order can also serve as an instrument for coresidence with the husband's parents. However, our focus is on estimating the wife's labor supply, and the wife's birth order may be more likely than the husband's birth order to be correlated with individual characteristics that may affect her decision on labor supply. In fact, the wife's birth order is correlated with her age and education. Thus, we do not use the wife's birth order as an instrument.

The correlations in the sample that excludes the eldest sons are shown in column (3) of Table 10.3. Surprisingly, no correlation exists between the husband's birth order among siblings and the determinants of the wife's labor supply in this sample. In particular, the wife's educational attainment, household income, and number of children show no statistically significant association with the husband's birth order. A valid instrument should be correlated with the coresidence variable, but not with the determinants of married women's labor supply. To obtain exogenous variations in living arrangements, we therefore exclude women married to eldest sons. The next section formally tests the relevance and exogeneity of our proposed instrument.

10.4 Results

Main Results

Table 10.4 presents estimates of the coresidence effect on the labor force participation of married women. The OLS estimates of the coresidence dummy variable in columns (1) and (4) are 0.122 and 0.144, respectively, for the samples with and without eldest sons. Both estimates are statistically significant at the 1% level. These estimates indicate that the labor force participation rate of married women living with the husband's parents is 12.2% or 14.4% higher than that of non-coresiding married

Table 10.4 Effect of coresidence with husband's parents on the labor force participation of married women: OLS versus IV estimation results

	Whole sample			Sample without eldest sons	
	(1) OLS	(2) IV Dummy for the eldest son	(3) IV Husband's birth order	(4) OLS	(5) IV Husband's birth order
Dependent variable: Labor force participation of married women					
Coresidence with husband's parents	0.122*** (0.034)	0.543* (0.327)	−0.012 (0.268)	0.144*** (0.045)	−0.738 (0.498)
Dependent variable in the first-stage regression: Coresidence with husband's parents					
Husband's birth order		0.075*** (0.014)	−0.022*** (0.004)		−0.017*** (0.004)
Weak IV test:		22.10	50.37		21.21
Specification test H_0: OLS versus H_1: IV		1.830 [0.176]	0.260 [0.613]		3.820* [0.051]
Sample size	2197	2197	2197	1558	1558

Notes The sample used in columns (4) and (5) excludes eldest sons. The instrument is the eldest son dummy in column (2) and the husband's birth order among siblings in columns (3) and (5). Numbers in parentheses are heteroskedasticity-consistent standard errors. Numbers in the weak IV test are F-statistics for the instrument in the first-stage regressions. Numbers in the specification test are F-statistics and those in brackets are p-values. Coefficient estimates for intercepts are omitted
*Significant at the 10% level
**Significant at the 5% level
***Significant at the 1% level

women, depending on the sample used. The OLS findings are also consistent with those of Sung and Chah (2001) and Sasaki (2002), who estimated the effect of coresidence on the labor supply of married women in Korea and Japan, respectively.

We use the three instruments in columns (2), (3), and (5): a dummy for the eldest son, the husband's birth order among siblings, and the husband's birth order among siblings in the sample that excludes eldest sons. First, we check the relevance condition for these instruments. Instruments that explain too little of the variation in the coresidence variable would be weak instruments. If instruments are weak, the normal distribution is a poor approximation to the sampling distribution of the IV estimator, even if the sample size is large. We perform an instrument relevance test for the husband's birth order variables. In the first-stage coresidence regression, the coefficient estimates of the husband's birth order variables in columns (2), (3), and (5) are 0.075, −0.022, and −0.017, respectively. All three estimates are statistically significant at the 1% level, confirming a negative correlation between the husband's birth order and the coresidence variable. The lower the husband's birth order is, the greater the likelihood is that he and his wife will live with his parents. F-statistics for the instruments in the first-stage regressions are 22.10, 50.37, and 21.21, respectively, which are greater than the approximate cut-off value of 10 for weak instruments suggested by Stock and Yogo (2005). Thus, the husband's birth order is not a weak IV, regardless of which instrument is used.

We also perform a specification test for the null hypothesis that the OLS estimator is consistent and efficient. Because we allow heteroskedasticity-consistent standard errors, we use the auxiliary regression method instead of the Hausman test. If the coresidence variable is endogenous, the two estimators of OLS and IV would appear to be significantly different. In contrast, the IV estimator would be less efficient than OLS if the coresidence variable is exogenous. Under the null hypothesis the Hausman test, the OLS estimator would be consistent and efficient, whereas the IV estimator would be consistent under the alternative hypothesis.

The F-statistic [p-value] for the instrument of the eldest son dummy is 1.830 [0.176], as reported in column (2) of Table 10.4, and thus the null hypothesis cannot be rejected at the 10% significance level. The result indicates that the eldest son dummy is not an exogenous instrument. Consistently, the eldest son IV estimate for the coresidence variable is positive and statistically significant at the 10% level, which is qualitatively the same as the OLS result in column (1). These results are consistent with the findings of Sasaki (2002), who used the eldest son dummy as an instrument for living arrangements in Japan.

In columns (3) and (5) of Table 10.4, the F-statistics [p-values] are 0.260 [0.613] and 3.820 [0.051], respectively, which indicates that the husband's birth order instrument in the whole sample might not be exogenous.[8] The result in column (3) suggests that unobservable characteristics of women married to eldest sons are different from those of other married women, and that these characteristics affect women's labor supply. When eldest sons are excluded from the sample, the husband's birth order IV estimate in column (5) becomes statistically significantly different from the OLS estimate, and we can therefore reject the null hypothesis. Thus, we focus on the husband's birth order IV in the sample without eldest sons to avoid possible endogeneity.

In column (5), our IV estimate of the coresidence effect on the labor force participation of married women is -0.783, which is not statistically different from zero. This suggests that coresidence with the husband's parents has no causal effect on the labor force participation of married women. The insignificant IV estimator indicates an endogeneity problem associated with intergenerational living arrangements. Unobservable attributes of married women that affect coresidence appear to be positively correlated with their unmeasured propensity to participate in the labor market. Thus, the OLS estimates are biased upward.

The main results in Table 10.4 are also maintained when we include additional variables in the regression model. Table 10.5 presents the multiple regression results. As expected, the labor force participation rate of married women increases with age, at a decreasing rate. The impact of schooling on labor force participation does not show a monotonically increasing pattern. Labor force participation rates of high school graduates are lower than those in any other group. The estimates of other variables have the expected signs and are statistically significant. Higher household

[8] The inclusion of additional control variables used in Table 10.5 does not change results in columns (2) and (3).

Table 10.5 Effect of coresidence with husband's parents on the labor force participation of married women: multivariate estimation results

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Age	0.81*** (0.017)	0.88*** (0.018)	0.95*** (0.016)	0.100*** (0.018)	0.077*** (0.019)	0.080*** (0.021)
Age squared	−0.001*** (0.000)	−0.001*** (0.000)	−0.001*** (0.000)	−0.001*** (0.000)	−0.001*** (0.000)	−0.001*** (0.000)
Education						
High school			−0.64* (0.034)	−0.075** (0.037)	−0.049 (0.034)	−0.063* (0.037)
College or higher			−0.018 (0.041)	−0.033 (0.046)	−0.010 (0.041)	−0.026 (0.045)
Log of annual household income (excluding wife's earnings)			−0.91*** (0.017)	−0.098*** (0.018)	−0.086*** (0.017)	−0.093*** (0.018)
Home ownership (= 1 if owns home; = 0 otherwise)			−0.027 (0.027)	0.025 (0.047)	−0.023 (0.027)	0.026 (0.046)
Self-employed husband (= 1 if self-employed; = 0 otherwise)			0.106*** (0.044)	0.119*** (0.034)	0.104*** (0.030)	0.117*** (0.033)
Number of children						
Total number of children					−0.045** (0.018)	−0.035* (0.020)
Number of children aged 6 or below					−0.156*** (0.024)	−0.144*** (0.028)
Coresidence with husband's parents (= 1 if coresiding; = 0 otherwise)	0.136*** (0.044)	−0.548 (0.461)	0.129*** (0.044)	−0.525 (0.466)	0.151*** (0.044)	−0.472 (0.458)
Sample size	1558	1558	1558	1558	1558	1558

Notes The sample excludes eldest sons. The instrument is the husband's birth order among siblings. Numbers in parentheses are heteroskedasticity-consistent standard errors. Coefficient estimates for intercepts are omitted

*Significant at the 10% level

**Significant at the 5% level

***Significant at the 1% level

income, having more children, and having younger children have negative effects on the labor supply of married women. A married woman with a self-employed husband has a higher labor force participation rate, which suggests that she has an option to work as a family worker in her husband's business. These results are consistent with the findings from the literature on female labor supply.

Overall coresidence estimates, shown in columns (4) and (5) of Table 10.4 and in Table 10.5, are very similar in magnitude and significance level. The estimated coresidence effects on labor force participation in Table 10.5, when coresidence is taken as exogenous, are similar to those in Table 10.4. Coresidence with the husband's parents has a significant positive association with the labor force participation of married women in columns (1), (3), and (5) in OLS specifications. However, this effect becomes statistically insignificant when coresidence is treated as endogenous. The two-stage least squares (2SLS) estimates of the coresidence variable in columns (2), (4), and (6) of Table 10.5 are all statistically insignificant. In summary, coresidence with the husband's parents does not have any causal effect on the labor force participation of married women, even when controlling for other explanatory variables that may affect their labor supply.

Subsample Results

Previous studies have consistently found that parental health status is an important determinant of coresidence (Stone and Short 1990; White-Means 1992; Ettner 1995, 1996; Pezzin and Schone 1999). Since a married woman's use of time greatly depends on the health of her coresiding parents-in-law, it is necessary to control for this to obtain unbiased estimation results. Other things being equal, poor health of the husband's parents is more likely to reduce a married women's time for market work. Without examining the direction of such time transfer within the family, we cannot precisely identify the effects of coresidence on the labor supply of married women. For example, the insignificant findings with respect to the IV estimators might result from the fact that positive effects of coresidence on labor supply are canceled out by uncontrolled negative effects of poor health of coresiding parents.

The relevant health status variable in our data is a subjective health indicator, which asks respondents to indicate their overall health status using a 5-point scale: 'very poor,' 'poor,' 'fair,' 'good,' and 'excellent.' Unfortunately for statistical analysis, most of the respondents (70%) record their health status as 'fair.' Furthermore, this information can be obtained only for the husband's parents who are coresiding with the married women in the sample. The data do not provide any individual information on the husband's non-coresident parents. Therefore, we can use neither the subjective health indicator nor the ages of the husband's parents to control for the effect of parental health.

We account for the parental health effect in an indirect way. We divide the sample into two groups based on the age of the married women: married women aged under 40 and married women aged 40 or above. If the age of married women is positively correlated with the age of their husbands' parents and the age of the husbands' parents is negatively correlated with their health status, these two subsamples may show different patterns of intra-family time transfer. These differences may result in different effects of coresidence on the labor force participation. Married women aged 25–39 in particular are less likely to spend their time on their husbands' parents than married women aged 40–55. Thus, the positive coresidence effect would be more conspicuous for married women under age 40. However, the upper rows of Table 10.6 show that the results from the two subsamples are virtually the same. In both subsamples, the coresidence variable is positively associated with labor force participation of married women, but this association disappears with correction for the endogeneity of coresidence.

Another way to account for the countervailing effects of childcare and parental care effects of coresidence on the labor supply of married women is to divide the sample into two groups based on whether a couple has children aged 6 or below. If a couple has younger children and coresides with parents, it is more likely that care services flow from parents to children rather than from children to parents. Then, two subsamples of married women may show different patterns of labor supply. In the sample with younger children, the effect of childcare services provided by the husband's parents may dominate the negative effect on labor supply of parental care. The opposite may be true in the other sample. This prediction is confirmed in OLS

Table 10.6 Effect of coresidence with husband's parents on the labor force participation of married women: subsample estimation results

	Without controls		With controls	
	(1) OLS	(2) IV	(3) OLS	(4) 2SLS
Married women's age				
Aged 25–39	0.120* (0.086)	−0.614 (0.603)	0174** (0.071)	−0.506 (0.634)
Sample size	738	738	738	738
Aged 40–55	0.153*** (0.057)	−0.610 (0.749)	0.152*** (0.055)	−0.299 (0.623)
Sample size	820	820	820	820
Children aged 6 or below				
At least one child aged 6 or below	0.202** (0.096)	−0.654 (0.739)	0.273*** (0.099)	−0.512 (0.717)
Sample size	342	342	342	342
No child aged 6 or below (including no child)	0.131*** (0.050)	−0.556 (0.600)	0.119** (0.048)	−0.409 (0.539)
Sample size	1216	1216	1216	1216

Notes The sample excludes eldest sons. The instrument in columns (2) and (4) is the husband's birth order among siblings. Numbers in parentheses are heteroskedasticity-consistent standard errors. The control variables included in columns (3) and (4) are the same as those in columns (5) and (6) of Table 10.5. Coefficient estimates for intercepts and controls are omitted

*Significant at the 10% level

**Significant at the 5% level

***Significant at the 1% level

estimates. The OLS estimate of the coresidence variable is positive in both samples and is much higher in the sample with children aged 6 or below than in the sample without children in this age range. However, the insignificant IV estimates in both samples indicate that the OLS findings are spurious.

Our main result that coresidence with the husband's parents does not have a causal effect on the labor supply of married women holds for various subsamples. If coresidence with the husband's parents affects married women's time allocation, we may expect different coefficients of the coresidence variable for the subsamples divided according to the age of married women and the number of younger children. However, our IV estimation results indicate otherwise. The insignificant IV estimates do not result from the positive (childcare) effects of coresidence being canceled out by the negative (elder care) effects of coresidence.

Robustness Checks

To check the robustness of the estimates, we modify our empirical model and estimate the effect of coresidence on the labor supply of married women using several additional specifications. The estimation results reported in Table 10.7 indicate that our main results are robust to model specifications.

Columns (1) and (2) in Table 10.7 present the estimation results from univariate probit and simultaneous bivariate probit models, respectively. Because our measure of labor supply is a dummy variable indicating whether married women participated in the labor market, it is possible to use the probit estimation. Given that the coresidence dependent variable is dichotomous and included as an independent variable in the labor force participation equation, this bivariate probit model is recursive or simultaneous. As shown by Greene (1998), unlike a linear simultaneous model, the bivariate probit model can be estimated as if no simultaneity problem exists. The simple probit estimates of coresidence with parents are positive and statistically significant at the 1% level. However, as expected, the bivariate probit estimates are not statistically different from zero.

Another measure of labor supply is working hours. Because a substantial proportion of married women work part time, the dummy for labor force participation may not properly represent the intensity of the labor market activity of married women. Columns (3) and (4) in Table 10.7 present estimation results from the Tobit model and the Tobit model with the IV, for which we use the weekly working hours of married women as a dependent variable. The results from the Tobit estimation mimic our main results. Although the Tobit estimates of coresidence with parents are positive and statistically significant, none of the Tobit-with-IV estimates is statistically significant. Taken together, the results in Table 10.7 provide further evidence that coresidence with the husband's parents does not have a causal effect on the labor supply of married women.

Table 10.7 Robustness checks: probit and tobit

	(1) Probit	(2) Bivariate probit	(3) Tobit	(4) Tobit with IV
Age	0.212*** (0.057)	0.213*** (0.057)	7.744*** (2.260)	7.997*** (2.278)
Age squared	−0.003*** (0.001)	−0.003*** (0.001)	−0.097*** (0.028)	−0.100*** (0.028)
Education				
High school	−0.133 (0.093)	−0.141 (0.093)	−5.651 (3.597)	−6.193* (3.619)
College or higher	−0.013 (0.112)	−0.023 (0.113)	−3.694 (4.417)	4.128 (4.442)
Log of annual household income (excluding wife's earnings)	−0.239*** (0.047)	−0.241*** (0.047)	−9.340*** (1.782)	9.573*** (1.795)
Home ownership (= 1 if owns home; = 0 otherwise)	−0.061 (0.074)	−0.028 (0.088)	−3.691 (2.924)	−2.209 (2.918)
Self-employed husband (= 1 if self-employed; = 0 otherwise)	0.284*** (0.083)	0.291 *** (0.083)	13.909*** (3.195)	14.174*** (3.216)
Number of children				
Total number of children	−0.127** (0.049)	−0.120** (0.051)	−4.991*** (1.937)	4.780** (1.948)
Number of children aged 6 or below	−0.443*** (0.071)	−0.431*** (0.074)	−20.012*** (2.933)	−19.506** (2.941)
Coresidence with husband's parents (= 1 if coresiding; = 0 otherwise)	0.427*** (0.126)	0.009 (0.640)	18.169*** (4.714)	25.962 (47.105)
Log-likelihood	−983.58	−1389.61	−4625.49	−4632.70
Sample size	1558	1558	1558	1558

Notes The sample excludes eldest sons. The instrument in columns (2) and (4) is the husband's birth order among siblings. Numbers in parentheses are standard errors. Coefficient estimates for intercepts are omitted. The bivariate probit model in column (2) jointly estimates both labor force participation and coresidence equations. The tobit models in columns (3) and (4) use the weekly working hours of married women as a dependent variable

*Significant at the 10% level

**Significant at the 5% level

***Significant at the 1% level

10.5 Conclusion

This paper has explored the hypothesis that intergenerational living arrangements may affect the labor supply of married women. To correct for the potential endogeneity of coresidence, we construct a new instrumental variable for coresidence with the husband's parents and estimate the causal effect of coresidence on the labor supply of married women.

An interesting finding in this study is that a woman married to the eldest son or a son of a low birth order has a higher propensity to live with her husband's parents. This suggests that deep-rooted family values and practices still persist in contemporary Korea. These cultural prescriptions, along with son preference in intergenerational living arrangements, allow us to use the husband's birth order as an instrument for coresidence with the husband's parents.

Our empirical results cast doubt on the argument that coresidence with parents affects the labor supply of married women. The positive effect of coresidence on labor force participation found in OLS estimation disappears after controlling for the endogeneity of coresidence. When coresidence is treated as endogenous, the IV estimates of the coresidence variable are statistically insignificant, regardless of subsamples and model specifications. This suggests the existence of a sorting process in which a married woman who tends to participate in the labor market is more likely to reside with her husband's parents.

An important policy implication emerges from our empirical findings. It appears that a policy intended to increase the labor supply of married women by promoting intergenerational living arrangements could be ineffective. For example, our results suggest that a small tax break for married couples living with their parents cannot substantially increase the labor supply of married women because coresidence with the husband's parents does not have a causal effect on the labor supply of married women. Rebuilding the traditional family structure is frequently proposed as a policy instrument for inducing more married women to participate in the labor market in East Asian countries. However, our results suggest that this is not a viable policy option.

This paper suggests both a strong influence of traditional family norms on intergenerational living arrangements and a clear trend away from traditional family structure in Korea. While new familial relationships emerge as traditional values and expectations give way to changing economic circumstances, people's choices about when and with whom to reside are still influenced by traditional norms and practices. It remains to be verified whether the results from Korean married women have general applicability to other countries with different cultural preferences for intergenerational living arrangements.

References

Angrist, Joshua D., and William N. Evans. 1998. Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review* 88: 450–77.

Arnold, Fred, and Liu Zhaoxiang. 1986. Sex preference, fertility, and family planning in China. *Population and Development Review* 12: 221–46.

Aquilino, William S. 1990. The likelihood of parent-child coresidence: Effects of family structure and parental characteristics. *Journal of Marriage and the Family* 52: 405–19.

Asis, Maruja Milagros B., Lita Domingo, John Knodel, and Kalyani Mehta. 1995. Living arrangements in four Asian countries: A comparative perspective. *Journal of Cross-Cultural Gerontology* 10: 145–162.

Asai, Yukiko, Ryo Kambayashi, and Shintaro Yamaguchi. 2015. Childcare availability, household structure, and maternal employment. *Journal of the Japanese and International Economies* 38: 172–192.

Choi, Eleanor Jawon, and Jisoo Hwang. 2015. Child gender and parental inputs: No more son preference in Korea? *American Economic Review* 105: 638–643.

Ettner, Susan. 1995. The impact of parent care on female labor supply decisions. *Demography* 32: 63–80.

Ettner, Susan. 1996. The opportunity costs of elder care. *Journal of Human Resources* 31: 189–205.

Freedman, Ronald, Ming-Cheng Chang, and Te-Hsiung Sun. 1982. Household composition, extended kinship and reproduction in Taiwan: 1973–1980. *Population Studies* 36: 395–411.

Giuliano, Paola. 2007. Living arrangements in western Europe: Does cultural origin matter? *Journal of the European Economic Association* 5: 927–52.

Greene, William H. 1998. Gender economics courses in liberal arts colleges: Further results. *Journal of Economic Education* 29: 291–300.

Lee, Bun Song, Soomyung Jang, and Jayanta Sarkar. 2008. Women's labor force participation and marriage: The case of Korea. *Journal of Asian Economics* 19: 138–54.

Logan, John R., Fuqin Bian, and Yanjie Bian. 1998. Tradition and change in the urban Chinese family: The case of living arrangements. *Social Forces* 76: 851–82.

Manacorda, Macro, and Enrico Moretti. 2006. Why do most Italian youths live with their parents? Intergenerational transfers and household structure. *Journal of the European Economic Association* 4: 800–29.

Mano, Yukichi, and Eiji Yamamura. 2011. Effects of husband's education and family structure on labor force participation and married Japanese women's earnings. *Japanese Economy* 38: 71–91.

Ogawa, Naohiro, and John F. Ermisch. 1996. Family structure, home time demands, and the employment patterns of Japanese married couples. *Journal of Labor Economics* 14: 677–702.

Park, Chai Bin, and Nam-Hoon Cho. 1995. Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea. *Population and Development Review* 21: 59–84.

Pezzin, Liliana, and Barbara Schone. 1999. Intergenerational household formation, female labor supply and informal caregiving: A bargaining approach. *Journal of Human Resources* 34: 475–503.

Reher, David S. 1998. Family ties in western Europe: Persistent contrasts. *Population and Development Review* 24: 203–34.

Sasaki, Masaru. 2002. The causal effect of family structure on labor force participation among Japanese married women. *Journal of Human Resources* 37: 429–40.

Stock, James H., and Motohiro Yogo. 2005. Testing for weak instruments in linear IV regression. In *Identification and inference in econometric models: Essays in honor of Thomas J. Rothenberg*, ed. Andrews, Donald W.K., and James H. Stock, 80–108. Cambridge: Cambridge University Press.

Stone, Robyn I., and Pamela F. Short. 1990. The competing demands for employment and informal caregiving to disabled elderly. *Medical Care* 28: 513–26.

Sung, Jaimie, and Eun Young Chah. 2001. Co-residence and its effect on the labor supply of married women. *Journal of Korean Labor Economic Association* 24: 97–124.

Takagi, Emiko, and Merril Silverstein. 2006. Intergenerational coresidence of the Japanese elderly: Are cultural norms proactive or reactive? *Research on Aging* 28: 473–92.
Wakabayashi, Midori, and Charles Horioka. 2006. Is the Eldest Son Different? The Residential Choice of Siblings in Japan, NBER Working Paper No. 12655.
Ward, Russell, John Logan, and Glenna Spitze. 1992. The influence of parent and child needs on coresidence in middle and later life. *Journal of Marriage and the Family* 54: 209–21.
White-Means, Shelly I. 1992. Allocation of labor to informal home health production: Health care for frail elderly, if time permits. *Journal of Consumer Affairs* 26: 69–89.
Wolf, Douglas A., and Beth J. Soldo. 1994. Married women's allocation of time to employment and care of elderly parents. *Journal of Human Resources* 29: 1259–76.
Wolf, Douglas A., and Beth J. Soldo. 1988. Household composition choices of older unmarried women. *Demography* 25: 387–403.

Part III
Public Policy and Political Economics

Chapter 11
An Employment Model of Social Enterprises and the Effects of Government Subsidy

Hyunwoo Hong and Biung-Ghi Ju

Abstract We consider an employment model of social enterprises which maximize the weighted average of profit and social value generated by employing disadvantaged workers. The social value consists of the direct welfare improvement of the disadvantaged employees and the positive ripple effect so generated. In the Cournot oligopoly model, we show that social enterprises can contribute to improving social welfare (the sum of the standard social surplus and the social value). We propose a subsidy scheme for social enterprises, the subsidy amount of which is proportional to the social performance (measured by social value). We show that the subsidy can play a positive role in providing incentives for social enterprises to produce more social values and in enhancing social welfare. Finally, we provide an empirical analysis on the effects of the government subsidy on economic and social performances of social enterprises using the case of Korea during 2014-2017. We use the voluntary disclosure data of social enterprises certified by Korean Government. Only the subsidy for social insurance fees has a positive effect on the social performance while the subsidy for personnel expenses, the major part of the government expenditure, has neither positive nor negative effect. On the other hand, subsidies for personnel expenses and social insurance fees have positive effects on net income. Our results suggest a proposal for redesigning the current subsidy scheme in order to improve its effectiveness.

[1]See more detailed explanation on social enterprises in the web page of Korea Social Enterprise Promotion Agency http://www.socialenterprise.or.kr.

[2] Also are 128 social service provision types, 120 local community contribution types, 189 mixed types, and 239 other types. The "other types" refers to the cases where it is difficult to quantify whether or not the social purpose of the social enterprise is realized.

The original version of this chapter was revised: The acknowledgement has been added. The correction to this chapter can be found at https://doi.org/10.1007/978-981-13-7036-6_22

H. Hong · B.-G. Ju (✉)
Seoul National University, Seoul, South Korea
e-mail: bgju@snu.ac.kr

H. Hong
e-mail: wonderwoo@snu.ac.kr

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_11

11.1 Introduction

A social enterprise is an intermediate form between for-profit and non-profit enterprises, and it conducts ordinary business activities such as production and sales of goods and services while pursuing other social purposes. Unlike commercial companies seeking only the profit for shareholders, social enterprises pursue social purposes as their main objective, providing social services and offering jobs for the disadvantaged.[1] The Social Enterprise Promotion Act and the Enforcement Ordinance in Korea categorize the types of social enterprises as employment type (employment model of social enterprises), social service provision type, mixed type (combination of employment type and social service provision type), local community contribution type, and other types. According to the Korea Social Enterprise Promotion Agency, as of November, 2018, a total of 2,089 social enterprises are certified by the Korean government and active. Among them, 1,413 employment types accounts for about 68% of the total.[2]

In this chapter, we provide a model of employment type social enterprises that occupy the largest portion of certified social enterprises in Korea.[3] We analyze the effects of social enterprise activities on distribution, efficiency and social welfare, and the impact of the subsidy policy subsidizing social enterprises based on their social performance.

According to Alter (2007), employment type social enterprises offer job opportunities for the disabled, the homeless, disadvantaged teenagers, criminals, etc. They are referred to as work integration type social enterprises by Davister et al. (2004), who view their main purpose is to achieve professional integration of the disadvantaged and the ordinary individuals through employment. Davister et al. (2004) categorize the forms of work integration into four types. First is the transitional occupation that provides vocational training or work experience so that disadvantaged workers can get involved in the labor market. Second is the creation of permanent self-financed jobs that relies on subsidies to compensate for the lack of productivity of disadvantaged workers in the early stage, but creates jobs that are self-funded without resorting to subsidies through productivity improvement of disadvantaged workers. Third is the professional integration with permanent subsidies by public authorities. Fourth is the socialization through a productive activity whose main purpose is the (re)socialization of the target groups through social contact, respect for rules, etc. The Social Enterprise Promotion Act in Korea encompasses several definitions of employment type social enterprises, and defines it as an organization that provides jobs to disadvantaged workers and realizes, through the employment, its social purpose.

In our model, disadvantaged workers have lower productivity than regular workers. Employment type social enterprises produce social value by hiring disadvantaged

[3] Different models need to be constructed to deal with the other types of social enterprises since their social purposes are different from the one we consider here. Hence the social values created by the other types also differ considerably. For instance, Hong and Ju (2016) provide a model of social enterprises that provide social services for the disadvantaged.

workers and paying higher wages than their productivity. The most important element of social value is the welfare improvement of individual workers. This welfare improvement will be measured by a multiple of the gap between the wage provided by social enterprises and their productivity (market wage). In addition, employment and welfare enhancement of disadvantaged workers generate positive ripple effects such as job integration and socialization through production activities as in Davister et al. (2004), which are assumed to be incorporated in what we call here as the secondary social value.

There are very few theoretical studies on the effects of social enterprises in Economics. Based on the case of Grameen Yogurt, Chu (2015) compares two types of social enterprises, one is the social enterprises that donate corporate profits to disadvantaged groups and the other is the social enterprises that maximize the utility of disadvantaged groups. He shows that the latter is more effective in terms of producing social values. The result is established in the monopoly market and does not take into account overall social welfare, which limits the applicability of his conclusion. Nevertheless, Chu (2015) is closely related with our investigation since he defines social value in terms of welfare change in disadvantaged groups and incorporates it in the objective of social enterprises. Hong and Ju (2016) investigate social service provision type social enterprises using the same framework as we use in this chapter. They show that the social welfare of the oligopolistic market is higher when commercial and social enterprises coexist than when only commercial enterprises compete. Credit constraints of disadvantaged groups play an important role in this result. Social enterprises resolve the credit constraints by directly reallocating social services. In our investigation, employment type social enterprises offer a higher wage to the disadvantaged than in labor market, which helps relaxing the credit constraints of the workers.

As far as we know, no prior study attempted economic analysis of employment type social enterprises. We provide a formal definition of social value that employment type social enterprises produce pursuing their social goals. We analyze the decision making of these social enterprises and examine whether social enterprises and the subsidy policy supporting them help improving social welfare. As in Hong and Ju (2016), it is assumed that the objective of a social enterprise is to maximize its performance that reflects both profit and social value. We analyze the (Cournot) oligopoly market where social enterprises compete with for-profit enterprises and derive the following results. First, employing disadvantaged workers in credit constraints and paying them a higher wage than their market wage yield net social surplus since social value so generated is higher than the increased wage cost. Second, if the fixed cost gap between the for-profit and the social enterprises is not too large, social enterprises help improving social welfare. Finally, monetary compensation (grants) for their social performance (measured by social value) can be used to improve social performance of social enterprises and ultimately social welfare.

We apply our theoretical framework to Voluntary Management Disclosure Data established during 2014–2017 by Korean Government and analyze the effect of government subsidy using the fixed effect regression model. Our regression result reports that only the subsidy for paying social insurance fees has a positive effect on the social performance of social enterprises. The subsidy for personnel expenses (labor cost) that accounts for the largest portion in the government subsidy has no significant effect on the social performance, while it has a positive effect on net profit.

To make the government subsidy to be more effective in promoting social performance, it is necessary to redesign the current subsidy scheme. More precisely, if comprehensive labor support is provided in proportional to the social value social enterprises generate as we design in our theoretical model, the subsidy will have a positive impact on improving social outcomes according to our theoretical results.

The impact of government subsidy for social enterprises has been studied by a number of previous Korean researchers (Kim and Lee 2012; Kim 2014, 2015; Hur and Yang 2015; Hong and Kim 2016, etc.). Most of the existing studies measure social performance in terms of the number of disadvantaged workers employed by social enterprises or the proportion of disadvantaged workers, which is the major difference from our empirical analysis. Our measure of social performance is in terms of the social value in our theoretical framework, that is, the improved well-being of disadvantaged workers employed by social enterprises over their well-being at their market wage. In addition, while most existing studies use only one year's data, we use four years' data and the fixed effect regression model, which allows us to handle firm-specific factors.

Kim and Lee (2012) show that government subsidies have no effect on economic performance (EROI). Hur and Yang (2015) show that the subsidies for personnel expenses and business development activities have no effect on employment and that the subsidy to supplement the payment for specialists positively affects the employment of the disadvantaged. Hong and Kim (2016) show that government subsidies have no effect on net profit, which is in contrast with our result. Also, the negative relation between the subsidy for paying social insurance fees and the rate of disadvantaged worker employment by Kim (2014) is in contrast with our result. The positive effect of external subsidies (including both government subsidies and private subsidies) on the disadvantaged worker employment reported by Kim (2015) is not confirmed in our result.

The rest of this chapter is organized as follows. Section 11.2 presents the theoretical model. Section 11.3 presents the equilibrium analysis of the model. Section 11.4 analyzes the effects of government subsidy for social enterprises. Section 11.5 provides some numerical examples. Section 11.6 provides the empirical analysis using Voluntary Management Disclosure Data of Korean social enterprises during 2014–2017. Section 11.7 concludes this chapter.

11.2 Model

11.2.1 Basic Human Needs, Credit Constraint and Social Value of Employment

Basic Human Needs and Credit Constraint

We consider an economy with workers who need to satisfy their basic human needs as well as other secondary needs. Let z be the good for basic human needs (e.g., basic goods or services provided in the market), called the *basic good*, and M the numeraire good, that is, the expenditure for all the other goods. All workers, for the sake of simplicity, have an identical quasi-linear utility from consuming the two goods $v(z) + M$, where $v(z)$ represents the benefit from the basic good, or basic human needs satisfaction. Let l be the amount of labor supply and $e(l)$ the disutility of labor supply which is assumed to be additively separable from the utility from the two consumption goods. Then the utility function of each worker is given by

$$U(z, l, M) = v(z) - e(l) + M.$$

Workers' utility function satisfies classical conditions of monotonicity and convexity.

Assumption 1 (*Monotonicity and Convexity*) $v'(z) > 0, e'(l) > 0, v''(z) < 0, e''(l) < 0$.

Maslow's hierarchy of needs (Maslow 1943, 1954) assumes the most basic level of needs that must be met prior to be motivated to desire the secondary or higher level needs in the hierarchy. Tay and Diener (2011) provides an empirical support for the human tendency of satisfying the basic needs prior to satisfying the secondary needs. Maslow's theory is supported also by John Rawls (1999) when he proposes the index of primary goods as the objective standard of interpersonal comparison of well-being and also by Amartia Sen (1987) and Martha Nussbaum (1992) in their proposal of basic human functioning. We incorporate the priority of basic needs in worker's utility by assuming that if the amount of the basic good diminishes indefinitely, its marginal benefit increases infinitely.

Assumption 2 (*Basic Human Needs*) $\lim_{z \to 0} v'(z) = \infty$.

Next we assume that borrowing money ($M < 0$) is not allowed. This reflects capital market imperfection and credit constraint.

Assumption 3 (*Credit Constraint*) $M \geq 0$.

There are two types of workers, *regular* workers and *disadvantaged* workers in the labor market. A worker's wage is determined by her marginal productivity. The wage of disadvantaged workers is lower than that of regular workers since disadvantaged workers have lower productivity than regular workers owing to various reasons such as lack of skills and physical defects.

Let w be the wage per labor unit of a regular worker and $\lambda \in [0, 1)$ the marginal rate of technical substitution of disadvantaged workers' labor for regular workers' labor. Then the wage per labor unit of a disadvantaged worker is λw. If a disadvantaged worker supplies l labor units, her labor income is λwl. Given p the price of basic good, the utility maximization problem of a disadvantaged worker is given by

$$\max_{z,l,M} U(z, l, M) = v(z) - e(l) + M,$$

$$\text{subject to}$$

$$pz + M \leq \lambda wl,$$
$$0 \leq M, z,$$
$$0 \leq l \leq \bar{l},$$

where $\bar{l}$ is the maximum level of labor supply.

We will examine two cases of binding credit constraint. The first case is when a disadvantaged worker's marginal productivity is so low that even if the worker supplies labor as much as possible and spends all her income on basic goods, the marginal benefit of basic goods is higher than the market price, that is, $v'(\lambda w\bar{l}/p) > p$. If she can borrow money in the capital market (i.e., $M < 0$), she will consume basic goods more. However, borrowing is not allowed (Assumption 2). In such a case, we will say that she is *under credit constraint*.

The second case is when a disadvantaged worker does not supply enough labor to support her basic need (this occurs when the disutility of labor is sufficiently high). Suppose that if a disadvantaged worker supplies maximum labor, she can consume the basic good as much as she wants ($v'(\lambda w\bar{l}/p) \leq p$). In this case, whether her credit constraint is binding or not depends on the level of labor supply. Let $H(\cdot)$ be an inverse function of $v'(\cdot)$ and $I_m (\equiv pH(p))$ the threshold income level such that $v'(I_m/p) = p$. The threshold labor l_m is the level of labor supply which is needed to earn the threshold labor income I_m (i.e., $l_m = I_m/\lambda w = pH(p)/\lambda w$). If one's marginal disutility of the threshold labor is lower than her wage ($e'(l_m) \leq \lambda w$), she will earn a higher income than I_m, and vice versa. Therefore when one's marginal disutility of threshold labor is larger than her wage (i.e., $e'(l_m) > \lambda w$), her credit constraint will be binding.

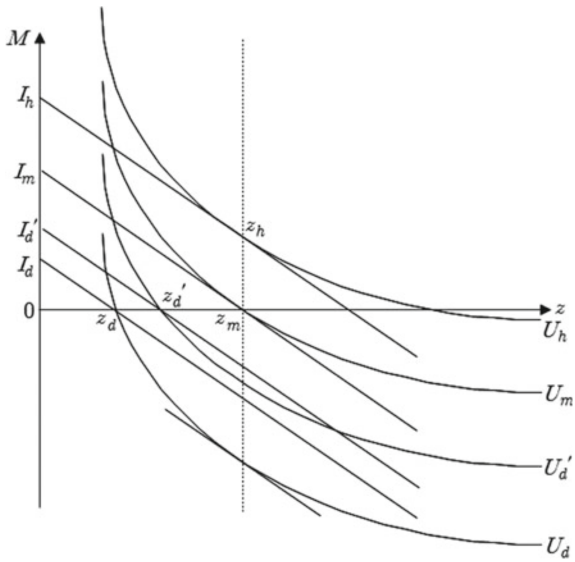
In what follows, for simplicity, we will only consider the first case and consider a range of wage offers for disadvantaged workers given by $[\lambda w, \bar{w}]$ with $\bar{w} < w$ such that *for all wage $w_d \in [\lambda w, \bar{w}]$, disadvantaged workers supply maximum labor $\bar{l}$ and their credit constraint is binding, that is,*

$$v'(w_d\bar{l}/p) > p.$$

Clearly, this implies $v'(\lambda w\bar{l}/p) > p$.

In Fig. 11.1, a disadvantaged worker whose income I_d is lower than the threshold income I_m spends all her income on the basic good and consume z_d, nevertheless. Note that the marginal benefit of the disadvantaged worker at z_d is larger than the price

Fig. 11.1 Labor income and credit constraint



p (i.e., $v'(z_d) > p$). If the disadvantaged worker is paid w_d which is below $\overline{w}$ and higher than her market wage λw, her income increases from $\lambda w\bar{l}(= I_d)$ to $w_d\bar{l}(= I_d')$. As her basic good consumption increases from $z_d(= I_d/p)$ to $z_d'(= I_d'/p)$, her utility also increases from U_d to U_d'. Note that

$$U_d' - U_d = v(z_d') - v(z_d) = \int_{z_d}^{z_{d'}} v'(x)dx > \int_{z_d}^{z_d'} pdx = I_d' - I_d.$$

Hence the utility of the disadvantaged worker increases more than the extra wage cost. Therefore, it is socially desirable that social enterprises employ disadvantaged workers under credit constraint and pay them a higher wage than their market wage.

The incremental utility of the disadvantaged worker under credit constraint can be rewritten as

$$\triangle U_d = v(w_d\bar{l}/p) - v(\lambda w\bar{l}/p) \equiv (1 + \theta(\lambda))(w_d - \lambda w)\bar{l}, \tag{11.1}$$

where $\theta(\lambda) \equiv \Delta U_d/(w_d - \lambda w)\bar{l} - 1$ represents the level of binding credit constraint. If there is no binding credit constraint, $\theta(\lambda) = 0$; otherwise, $\theta(\lambda) > 0$.

Social Value of Employment

Haugh (2006) classified performance of social enterprises into economic, social, and environmental dimensions. In each dimension, social impact is divided into direct and indirect impacts. For example, in the dimension of the economic performance, the number of jobs created is a direct impact and skill improvement in the local population and greater employment opportunities in other organizations are indirect impacts. In the dimension of the social performance, the supply of services to the community, improved access to services, higher quality of life are direct impacts and social trust, social integration of disadvantaged workers are indirect impacts. Reduction of unrecycled wastes is a direct environmental impact and more attractive regional environment and contribution to sustainability agenda are indirect environmental impacts.

Building on Haugh (2006), social values are defined by incorporating both direct and indirect impacts on the disadvantaged and the society. Direct impact on the disadvantaged is measured by the welfare improvement of the disadvantaged workers employed by social enterprises, which is called as the *primary social value*. Indirect impact is the positive ripple effects of employing disadvantaged workers and improving their welfare, which is called as the *secondary social value*. The primary social value can be regarded as the direct impact of social performance in Haugh (2006) and the secondary social value can be regarded as the indirect impact of social performance. The primary social value is measured by the change in the well-being of disadvantaged workers employed by social enterprises. Since it is difficult to measure the secondary social value, we assume that the secondary social value arises in proportion to the primary social value and the degree is $\eta \geq 0$.[4]

Denote labor supply of regular workers and that of disadvantaged workers by L_r and L_d, respectively. If regular and disadvantaged workers are perfectly substitutable with the rate of λ, as we assume throughout the chapter, the effective labor L is $L \equiv L_r + \lambda L_d$.

Throughout the paper, capital inputs are fixed and all firms have the identical production function, denoted by $f(L)$, which depends on effective labor as well as capital. Note that the inverse of f, $g \equiv f^{-1}$ is the conditional effective labor demand function.

As we will show later, social enterprises achieve the effective labor by employing only disadvantaged workers. Then their conditional effective labor demand function is $L_d(x) = g(x)/\lambda$. Then the (short-run) variable cost (labor cost) function of for-profit enterprises is given by $c_P(x, w) = wg(x)$. When w_d is the wage that a social enterprise pays to her disadvantaged employees, then her (short-run) variable cost (labor cost) function is given by $c_S(x, w_d) = w_d g(x)/\lambda$. To summarize,

- w: wage rate paid to regular workers by for-profit enterprises

---

[4]Social values are distinct from positive externalities because social values are explicitly taken into account in social enterprises' decision-making. Nevertheless, the reduction of social costs (e.g., the increase in crime rate, the spread of the social unrest, and so on) may not be fully reflected in the decision making of social enterprises and so it also exhibits positive externality, which can be the basis of supporting social enterprises.

- λw: wage rate paid to disadvantaged workers by for-profit enterprises
- w_d: wage rate paid to disadvantaged workers by social enterprises
- $g(\cdot) \equiv f^{-1}(\cdot)$
- $L_r(x) \equiv g(x)$: conditional labor demand for regular workers
- $L_d(x) \equiv g(x)/\lambda$: conditional labor demand for disadvantaged workers
- $c_P(x, w) \equiv wL_r(x) = wg(x)$: (short-run) variable cost of each for-profit enterprise
- $c_S(x, w_d) \equiv w_d L_d(x) = w_d g(x)/\lambda$: (short-run) variable cost of each social enterprise

Using (11.1), the *primary social value*, that is, the welfare improvement, per unit of labor, of a disadvantaged worker employed by social enterprises can be expressed as

$$(1 + \theta)(w_d - \lambda w).$$

We assume that the wage paid to disadvantaged workers by social enterprises cannot exceed the wage paid to regular workers by for-profit enterprises, that is, $w_d \in [\lambda w, \bar{w}]$ and $\bar{w} \leq w$.

Reflecting *the secondary social value*, which is assumed to be generated in proportional to the primary social value in the rate of $\eta > 0$, the (total) *social value* that a social enterprise produces per unit of labor is given by

$$(1 + \eta)(1 + \theta)(w_d - \lambda w).$$

Let $\gamma \equiv (1 + \eta)(1 + \theta) - 1$ be the coefficient that reflects both credit constraint and the secondary social value, i.e., $1 + \gamma = (1 + \eta)(1 + \theta)$. The social value produced by a social enterprise which employs L_d units of labor from disadvantaged workers is given by

$$SV = (1 + \gamma)(w_d - \lambda w)L_d. \tag{11.2}$$

The social value can also be approached from the viewpoint of an enterprise's cost. Consider a social enterprise which pursues social goal in a rational manner through either employing disadvantaged workers or a donation. Then her employment decision and cost function reveal at least partially how much social value she produces. A social enterprise which employs disadvantaged workers instead of regular workers must endure higher variable cost than for-profit enterprises. The reason a rational social enterprise is willing to pay this higher cost is because the social value so generated is at least as large as the extra variable cost. If the extra variable cost is higher than the social value, then she would rather hire regular workers and donate the saved cost to disadvantaged workers. In other words, the social value produced by hiring disadvantaged workers offsets the cost disadvantage.

Lemma 1 *If a rational social enterprise pursuing both profit and social value hires disadvantaged workers at wage w_d and produces x, then social value generated by hiring disadvantaged workers is at least as large as the extra variable cost*

$c_S(x, w_d) - c_P(x, w)$ of hiring disadvantaged workers. In fact, social value and cost gap satisfy

$$SV = (1 + \gamma)(c_S(x, w_d) - c_P(x, w)). \tag{11.3}$$

Proof Suppose that a social enterprise hires disadvantaged workers at the wage w_d and produces x. Suppose that the social value so generated is lower than $c_S(x, w_d) - c_P(x, w)$. Then when the firm hires regular workers instead, she can save $c_S(x, w_d) - c_P(x, w)$ and can donate this saved cost to the disadvantaged, which will generate greater social value due to their credit constraints. This contradicts the rationality of the firm.

To prove the second part, note that for all x, w_d, and w, $c_P(x, w) = wL_r(x) = wg(x)$ and $c_S(x, w_d) = w_d L_d(x) = w_d g(x)/\lambda$. Hence, $c_S(x, w_d) - c_P(x, w) = (w_d - \lambda w)\frac{g(x)}{\lambda} = (w_d - \lambda w)L_d(x)$. Using this relation and (11.2), we obtain (11.3).

□

11.2.2 Social Enterprises

Organizations' payoffs vary according to their goals or missions. For example, for-profit enterprises pursue profit maximization, whereas public firms concern not only profit but social welfare addressing the issue of market failure. Matsumura (1998) and Matsumura and Kanda (2005) model partially privatized public firms by assuming that they maximize the weighted average of profit and social welfare, i.e., $U = (1 - a)\pi + aSW$ where the weight on social welfare $a \in [0, 1]$ represents the degree of the privatization. Kaneda and Matsui (2003) and Kato and Tomaru (2007) assume that the payoff of private enterprises which seek for profit and some other objectives such as revenue is $U = (1 - \theta)\pi + \theta F$, where F represents non-profit objectives and $\theta \in [0, 1)$ the weight on F.

We assume that social enterprises pursue both profit and social goal by maximizing the weighted average of profit and social value, which is given by

$$U_S = (1 - \alpha)\pi + \alpha SV,$$

where $\alpha \in (0, 1)$ is the weight which the social enterprises put on their social goal.[5] We exclude two extreme cases of α with either 0 or 1, namely non-profit organization or for-profit enterprise.

We now consider the employment decision by a social enterprise. Assume that the quantity of production by the firm and the output price are given by x and p. The

[5]Similar to our definition of social value is the goal of social enterprises in Chu (2015), where social enterprises are assumed to maximize the surplus of the disadvantaged. In Cho and Lee (2017), the objective function of social enterprises is the sum of consumer surplus and social value created in proportion to production of social enterprises.

firm needs to decide how to mix regular and disadvantaged workers as well as the wage for the latter, L_r, L_d, w_d. Using (11.2), the firm solves

$$\max_{L_r, L_d, w_d} \quad (1-\alpha)(px - wL_r - w_d L_d) + \alpha(1+\gamma)(w_d - \lambda w)L_d$$
$$subject\ to \quad L_r + \lambda L_d = g(x), \quad \lambda w \leq w_d \leq \bar{w}.$$

When a social enterprise puts sufficiently large weight on the social goal, she will hire only disadvantaged workers and provide the highest wage in the range of her wage offer as shown by the next lemma.

Lemma 2 *If the weight on social value is larger than* $1/(2+\gamma)$*, the social enterprise employs only disadvantaged workers and sets their wage at the highest level* $\bar{w}$*.*

Proof If x is given, the payoff maximization problem is equal to the following cost minimization:

$$\min_{L_r, L_d, w_d} \quad (1-\alpha)(wL_r + w_d L_d) - \alpha(1+\gamma)(w_d - \lambda w)L_d$$
$$subject\ to \quad L_r + \lambda L_d = g(x),\ \lambda w \leq w_d \leq \bar{w}.$$

The objective function can be rearranged into $(1-\alpha)wg(x) + (w_d - \lambda w)((2+\gamma)\alpha - 1)L_d$. Therefore, if $\alpha \geq \frac{1}{2+\gamma}$, then the solution is obtained when $w_d = \overline{w}$, $L_r = 0$, and $L_d = \frac{g(x)}{\lambda}$. □

Employing a disadvantaged worker is more costly for raising profit but helps producing social value. When $\alpha \geq 1/(2+\gamma)$, the latter positive effect on the payoff of the social enterprise exceeds the former negative effect. This net surplus of replacing a regular worker with a disadvantaged worker increases as the wage for the disadvantaged worker increases up to the highest wage level that social enterprises can offer.

11.2.3 Mixed Oligopoly

We consider a (Cournot) oligopoly model with homogeneous goods where both social and for-profit enterprises compete by choosing quantities of their output. The market demand function is given by $p = p(X)$, which is twice continuously differentiable with $p'(X) < 0$. There are n identical for-profit enterprises $i \in \{1, \ldots, n\}$ and m identical social enterprises $j \in \{1, \ldots, m\}$. All for-profit enterprises have the same cost function $C_P(x) = c_P(x) + F_P$ composed of variable and fixed costs. All social enterprises also have the same cost function $C_S(x) = c_S(x) + F_S$, and they also have the same weight α on social value. These cost functions are assumed to satisfy: $c'_S(x) > c'_P(x) > 0$, $c''_S(x) > c''_P(x) > 0$, and $F_S \geq F_P$.

Denote the output of for-profit enterprise i by $x_{P,i}$ and the output of social enterprise j by $x_{S,j}$. We also use the following notation $X = \sum_{i=1}^{n} x_{P,i} + \sum_{j=1}^{m} x_{S,j}$, $X_{-P,i} = X - x_{P,i}$, and $X_{-S,j} = X - x_{S,j}$.

For-profit enterprises maximize their profit by choosing the quantity of supply:

$$\max_{x_{P,i}} p\left(x_{P,i} + X_{-P,i}\right) x_{P,i} - wL_r\left(x_{P,i}\right).$$

Assume that the solution is unique and denote the solution function, or reaction function by $R_{P,i}\left(X_{-P,i}\right)$.

A social enterprise maximizes the weighted average of profit and social value. In particular, social enterprise j decides output $x_{S,j}$, employment of regular and disadvantaged workers L_r, L_d, and wage for disadvantaged workers w_d by solving:

$$\max_{x_{S,j}, L_r, L_d, w_d} (1-\alpha)\left(p\left(x_{S,j} + X_{-S,j}\right) x_{S,j} - wL_r - w_d L_d\right) + \alpha(1+\gamma)\left(w_d - \lambda w\right) L_d,$$

$$\text{subject to } L_r + \lambda L_d = g(x_{S,j}),\ \lambda w \leq w_d \leq \bar{w}.$$

Assume that the solution is unique and denote the solution function, or reaction function by $R_{S,j}(X_{-S,j})$.

Since all products are homogeneous (perfectly substitutable), the derivatives of reaction functions have negative values.

$$R_i'(X_{-P,i}) = -\frac{p'' x_{P,i} + p'}{p'' x_{P,i} + 2p' - c_P''} < 0. \tag{11.4}$$

$$R'(X_{-S,j}) = -\frac{p'' x_{S,j} + p'}{p'' x_{S,j} + 2p' - c_S'' + \{\alpha(1+\gamma)(c_S'' - c_P'')\}/(1-\alpha)} < 0. \tag{11.5}$$

Finally, we assume the second-order conditions of firms' problems are satisfied. From (11.4) and (11.5), we obtain

$$p'' x_{P,i} + p' < 0;\ p'' x_{S,j} + p' < 0. \tag{11.6}$$

Social evaluation of the oligopoly market outcome is based on overall social welfare that is the sum of consumer surplus, producer surplus, and social value. When the market outcome gives total output X composed of individual outputs by firms $\left((x_{P,i})_{i=1,\ldots,n}, (x_{S,j})_{j=1,\ldots,m}\right)$, social welfare SW is given follows:

$$SW = \int_0^X p(x)dx - \sum_{i=1}^{n} C_P(x_{P,i}) - \sum_{j=1}^{m} C_S(x_{S,j}) + \sum_{j=1}^{m} SV(x_{S,j}). \tag{11.7}$$

11.3 Social Welfare in the Mixed Oligopoly

To investigate the effect of social enterprises on social welfare, we compare Cournot equilibria in the mixed oligopoly with both for-profit and social enterprises and in the oligopoly with only for-profit enterprises. For a fair comparison, we assume that in both cases, the total number of firms equals $n + m$. We show that social enterprises can contribute for improving social welfare, that is, the social welfare in the mixed oligopoly is higher than the social welfare in the oligopoly.

Proposition 1 *When the fixed cost gap $F_S - F_P$ is small enough, social enterprises can improve social welfare. In other words, for some weight $\alpha \geq 1/(2+\gamma)$, the mixed oligopoly equilibrium achieves a higher social welfare than the oligopoly equilibrium with the same number of firms. When the fixed cost gap is not small enough, whether social enterprises can improve social welfare or not depends on the size of the fixed cost gap, the degree of credit constraint of disadvantaged workers θ, and the size of the secondary social value η.*

Proof First consider the oligopoly of $(n+m)$ for-profit enterprises. Let x_E and SW_E denote the output of each for-profit enterprise and the social welfare in the equilibrium, respectively. Then we obtain:

$$p'x_E + p - c'_P(x_E) = 0, \tag{11.8}$$

$$SW_E = \int_0^{(n+m)x_E} p(x)dx - (n+m)C_P(x_E). \tag{11.9}$$

Next consider the mixed oligopoly with n for-profit enterprises and m social enterprises. Let x_P, x_S, and $SW_{P,S(\alpha)}$ denote the output of each for-profit enterprise, the output of each social enterprise, and the social welfare in the equilibrium, respectively. Using the first-order conditions, we get:

$$p'x_P + p - c'_P(x_P) = 0, \tag{11.10}$$

$$(1-\alpha)[p'x_S + p - c'_S(x_S)] + \alpha(1+\gamma)[c'_S(x_S) - c'_p(x_S)] = 0, \tag{11.11}$$

$$SW_{P,S(\alpha)} = \int_0^{nx_P+mx_S} p(x)dx - nC_P(x_P) - mC_S(x_S) + m(1+\gamma)[c_S(x_S) - c_P(x_S)]. \tag{11.12}$$

Total differentiation of (11.10) and (11.11) gives

$$\frac{dx_P}{d\alpha} = -\frac{L}{LM-KN}\frac{SV'(x_S)}{(1-\alpha)^2}, \tag{11.13}$$

$$\frac{dx_S}{d\alpha} = -\frac{K}{LM-KN}\frac{SV'(x_S)}{(1-\alpha)^2}, \tag{11.14}$$

where $K \equiv np''x_P + (n+1)p' - c''_P(x_P), L = m(p''x_P + p'), M \equiv n(p''x_S + p')$, and $N \equiv mp''x_S + (m+1)p' - c''_S(x_S) + \frac{\alpha(1+\gamma)}{1-\alpha}[c''_S(x_S) - c''_P(x_S)]$. From the assumption of strategic substitutes and the second-order conditions, $K < 0, L < 0, M < 0$, and $N < 0$. If $\alpha_1 = \frac{1}{2+\gamma}$, then $x_E = x_P = x_S, LM|_{\alpha=\alpha_1} = mn(p''x_E + p')^2$, and $KN|_{\alpha=\alpha_1} = (n(p''x_E + p') + p' - c''_P(x_E))(m(p''x_E + p') + p' - c''_S(x_E))$. $LM - KN|_{\alpha=\alpha_1} < 0$ gives the followings.

$$\frac{dx_P}{d\alpha}|_{\alpha=\alpha_1} < 0, \tag{11.15}$$

$$\frac{dx_S}{d\alpha}|_{\alpha=\alpha_1} > 0. \tag{11.16}$$

We know that $SW_{P,S(\alpha_1)} - SW_E = m\gamma[c_S(x_E) - c_P(x_E)] - m(F_S - F_P)$. Divide two cases.

Case (i) $F_P = F_S$.

In this case, if $\gamma > 0$, $SW_{P,S(\alpha_1)} - SW_E > 0$. If $\gamma = 0$, $SW_{P,S(\alpha_1)} - SW_E = 0$. Note $\frac{dSW_{P,S(\alpha_1)}}{d\alpha} = (n\frac{dx_P}{d\alpha}|_{\alpha=\alpha_1} + m\frac{dx_S}{d\alpha}|_{\alpha=\alpha_1})[p - c'_P(x_E)] + m\frac{dx_S}{d\alpha}|_{\alpha=\alpha_1}\gamma[c'_S(x_E) - c'_P(x_E)]$. From (11.13)–(11.16), $n\frac{dx_P}{d\alpha}|_{\alpha=\alpha_1} + m\frac{dx_S}{d\alpha}|_{\alpha=\alpha_1} > 0$. Therefore $\frac{dSW_{P,S(\alpha_1)}}{d\alpha} > 0$. For some $\alpha > \alpha_1$, despite $\gamma = 0$, the social welfare in the mixed oligopoly is higher than the social welfare in the oligopoly.

Case (ii) $F_P < F_S$.

If $\gamma = 0$, $SW_{P,S(\alpha_1)} - SW_E < 0$. If $\gamma > 0$, $SW_{P,S(\alpha_1)} - SW_E = m\gamma[c_S(x_E) - c_P(x_E)] - m(F_S - F_P)$, which is positive if $(F_S - F_P)$ is sufficiently small. □

In particular, consider the case where $F_P = F_S$ and $\gamma = 0$. If $\alpha < 1/2$, social enterprises decrease social welfare. If $\alpha = 1/2$, social enterprises do not change social welfare. Therefore, only if α is larger than $1/2$, social enterprises may improve social welfare. However, if α is too large, they spend so much for producing social value and sacrifice profits that they may disimprove social welfare.

11.4 Social Value Subsidy

We have shown that when disadvantaged individuals suffer from insufficient satisfaction of basic needs due to credit constraint, social enterprises can play an important role for relaxing the credit constraint, which helps improving the social welfare. However, it is necessary for social enterprises to have a proper weight on social value and produce it at an appropriate level. If not, social enterprises may also deteriorate the social welfare by increasing production cost too much. In such a case, there is a need for policy means that enables social enterprises to behave optimally. Social enterprises can also face negative profit and are financially unsustainable. Public policy may also help making social enterprises financially sustainable.

We consider a subsidy scheme that provides financial support for firms based on social value they produce. Let the subsidy be paid in proportion to social value. We call this subsidy scheme as social value subsidy, SVS, which is given by

$$Social\ Value\ Subsidy = \beta \cdot SV$$

where β is the compensation rate on social value.

Social value subsidy may be provided only to social enterprises or to both for-profit and social enterprises.[6] In the following, we analyze how the social value subsidy affects firms' decision.

Under the social value subsidy, for-profit enterprises have a chance to generate profits by employing disadvantaged workers with a higher wage than market and receiving subsidies. Their profit maximization problem, then, is given by

$$\begin{aligned} \max_{x_P, L_r, L_d, w_d} \quad & [p(x_P + \overline{X})x_P - wL_r - w_d L_d] + \beta(1+\gamma)(w_d - \lambda w)L_d \\ subject\ to \quad & L_r + \lambda L_d = g(x_P),\ \lambda w \le w_d \le \bar{w}. \end{aligned}$$

Given output level x_p, the firm needs to decide how to mix two types of workers to produce x_P. If the compensation rate β on social value is larger than $1/(1+\gamma)$, hiring only disadvantaged workers and paying them the maximum wage (i.e., $w_d = \bar{w}$, $L_r = 0$, and $L_d = g(x_P)/\lambda$) is optimal. If β is less than $1/(1+\gamma)$, hiring only regular workers or hiring disadvantaged workers at their market wage is optimal, that is, $w_d = \lambda w$ or $L_d = 0$. If β is equal to $1/(1+\gamma)$, hiring regular workers and hiring disadvantaged workers with any wage between λw and $\bar{w}$ are indifferent. So the compensation rate on social value should be higher than $1/(1+\gamma)$ in order to induce for-profit enterprises to produce social value.

Social value subsidy increases the profits of social enterprises. Moreover, it provides a greater incentive of producing social value. Note that the objective function of social enterprises under the subsidy is given by

$$U_S = (1-\alpha)(\pi_S + \beta SV_S) + \alpha SV_S = (1-\alpha)\pi_S + [(1-\alpha)\beta + \alpha]SV_S.$$

If $\beta > 0$, social value subsidy makes social enterprises behave as if they have a higher weight on social value than their original weight.

Our main result in this section concerns the case that the social value subsidy is provided only to social enterprises. Let α^* be the optimal weight on social value maximizing the social welfare of the mixed oligopoly equilibrium. Then by Proposition 1, α^* has a value between $1/(2+\gamma)$ and 1. When social enterprises have different weight on social value from α^*, social value subsidy can help to provide the optimal incentive for them to produce social value.

[6] In Korea, the government provides various subsidies for only social enterprises certified by the Korean government based on the Social Enterprise Promotions Act.

Proposition 2 *Assume that social value subsidy is provided only to social enterprises. If the compensation rate on social value β is given by*

$$\beta^* = \frac{\alpha^*}{1-\alpha^*} - \frac{\alpha}{1-\alpha},$$

then under the subsidy, the social welfare in the mixed oligopoly with social enterprises with weight α on social value reaches the maximum social welfare in the mixed oligopoly with social enterprises with the optimal weight α^*.

Proof With the subsidy given by βSV, the first-order condition of social enterprises is given by

$$(1-\alpha)(p'x_S + p - c'_S) + [(1-\alpha)\beta + \alpha](1+\gamma)(c'_S - c'_P) = 0. \tag{11.17}$$

The optimal compensation rate β^* is β that makes the equation obtained from putting α^* into (11.11) be the same as (11.17). That is,

$$\frac{(1-\alpha)\beta^* + \alpha}{1-\alpha} = \frac{\alpha^*}{1-\alpha^*},$$

which gives the result. □

11.5 Numerical Examples

In this section, we explain our main results, using a duopoly market where a demand curve is linear and cost functions are quadratic forms. The inverse demand curve is $p(X) = -X + 100$. The cost functions of for-profit and social enterprise are $C_P(x) = x^2$ and $C_S(x) = 2x^2 + F_S$, respectively. F_S is 0 if there is no fixed cost gap. Otherwise, $F_S = 50$.

The equilibrium of the duopoly market with only for-profit enterprises is as follows.

$$x_P^* = 20,\ X^* = 40,\ SW^* = 2400.$$

The equilibrium of the mixed duopoly market with both for-profit and social enterprises is as follows.

$$x_P^{**} = \frac{100(-5+\alpha(7+2\gamma))}{-23+\alpha(31+8\gamma)},\ x_S^{**} = \frac{300(-1+\alpha)}{-23+\alpha(31+8\gamma)},\ X^{**} = 25 + \frac{225(-1+\alpha)}{-23+\alpha(31+8\gamma)}.$$

As the weight on social value becomes larger, the output of the for-profit enterprise becomes smaller ($dx_P^{**}/d\alpha < 0$) and the output of the social enterprise becomes larger ($dx_S^{**}/d\alpha > 0$), which increases the total output in the market ($dX^{**}/d\alpha > 0$).

The social welfare in the equilibrium of the mixed duopoly market is dependent on α and γ.

Figure 11.2 shows the case in which there is no binding credit constraint and no secondary social value (i.e., $\gamma = 0$). If there is no fixed cost gap ($F_S = 0$), the social enterprise with a proper weight on social value can attain higher social welfare than in the duopoly market. However, in the case of the fixed cost gap ($F_S = 50$), the social enterprise cannot improve social welfare regardless of her weight on social value.

Figure 11.3 shows the case in which there are binding credit constraint and secondary social value (i.e., $\gamma > 0$). Unlike Fig. 11.2, there is a range of α where the thick curve which depicts social welfare when $F_S = 50$ is above the dotted horizontal line which depicts the social welfare of pure duopoly. Even if there is a fixed cost gap, a social enterprise can contribute to improving social welfare when credit constraint and secondary social value are taken into account.

If the government (or non-governmental organizations) transfers the same amount of money as the increased amount of wage by social enterprises to the disadvantaged workers, the primary social value the government can produce will be the same. However, secondary social values (such as their improved productivity due to learning effects and expanding employment opportunities for homogeneous private enterprises) cannot be achieved. Hence the overall social goal can be better achieved by social enterprises than by the government.

Finally, we will show the effects of social value subsidy. Suppose that SVS is provided only to the social enterprise. Then the equilibrium is changed as follows.

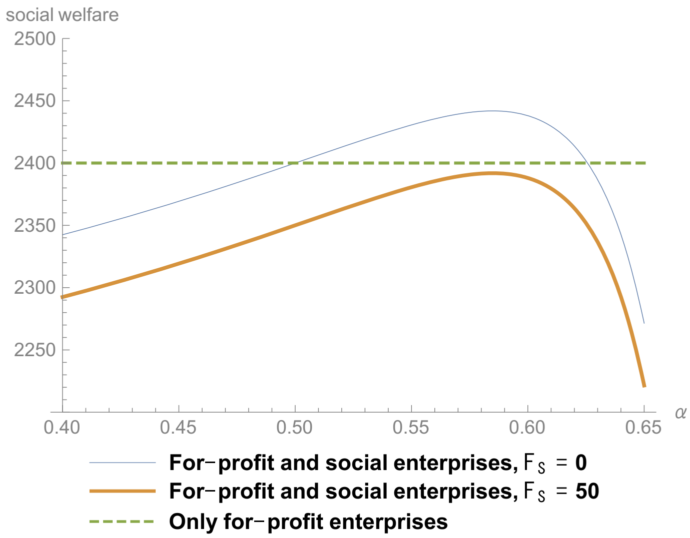


Fig. 11.2 Social welfare when $\gamma = 0$

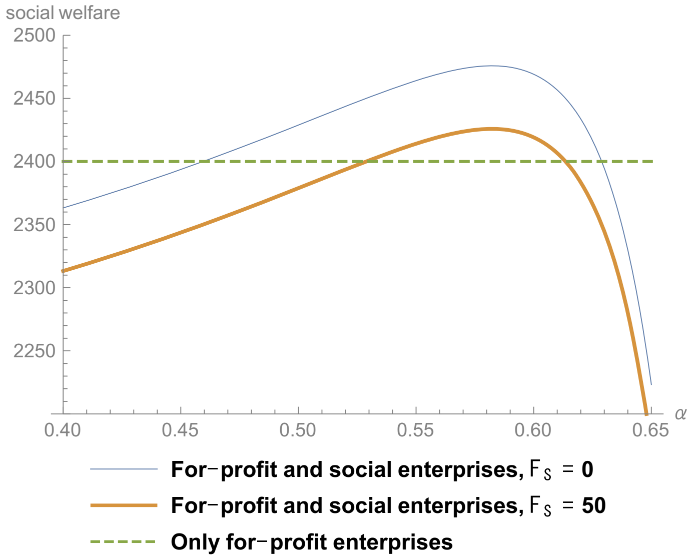


Fig. 11.3 Social welfare when $\gamma = 0.05$

$$x_P^{***} = 25 + \frac{75(-1+\alpha)}{23 - 8\beta(1+\gamma) + \alpha(-31 - 8\gamma + 8\beta(1+\gamma))},$$

$$x_S^{***} = -\frac{300(-1+\alpha)}{23 - 8\beta(1+\gamma) + \alpha(-31 - 8\gamma + 8\beta(1+\gamma))},$$

$$X^{***} = 25 - \frac{225(-1+\alpha)}{23 - 8\beta(1+\gamma) + \alpha(-31 - 8\gamma + 8\beta(1+\gamma))}.$$

As β increases, the output of the for-profit enterprise decreases ($dx_P^{***}/d\beta < 0$) and the output of the social enterprise increases ($dx_S^{***}/d\beta > 0$). The total output in the market also increases ($dX^{***}/d\beta > 0$). As we mentioned, SVS makes the social enterprise behave as if she puts a higher weight on social value than she does originally. Therefore, SVS can be used as a policy means to change the behavior of the social enterprise, which can result in improving social welfare.

Consider a for-profit enterprise and a social enterprise whose weight on social value is 0.45 in the market (Fig. 11.4). If there is no SVS ($\beta = 0$), social welfare is lower than 2400. By introducing SVS with compensation rate 0.5 ($\beta = 0.5$), social welfare can increase over 2400. In fact, social welfare can be maximized when the compensation rate is set at the socially optimal level (β^*).

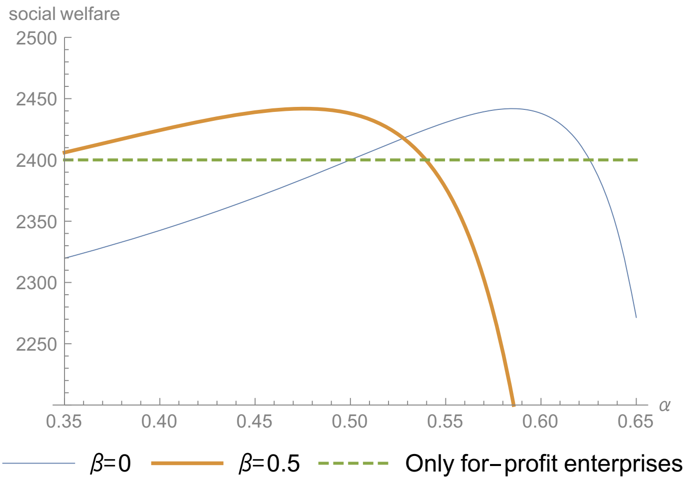


Fig. 11.4 Social value subsidy and social welfare

11.6 Empirical Analysis on Social Enterprises in Korea

In Korea, the Social Enterprise Promotions Act was enacted in January, 2007 and enforced in July, 2007. To receive government supports, social enterprises must be certified by this act. The government provides various supports such as financial supports (e.g., subsidy for personnel expenses, subsidy to supplement the payment for specialists, subsidies for social insurance fees and business development activities), management consulting, preferential public purchase of products, and tax benefits. Some of the certified social enterprises provide voluntary management disclosures (hereafter, VMD) which include data about public and private supports, and their economic and social performances of the previous year.[7] In total, 116 social enterprises in 2014, 218 social enterprises in 2015, 270 social enterprises in 2016, and 367 social enterprises in 2017 participated in VMD.

To analyze the relationship between government subsidies and social enterprises' performances, we used VMD data in 2014–2017. In previous sections, we only dealt with an employment model of social enterprises. Therefore, we also restrict our empirical analysis to employment type social enterprises. As a result, 67 social enterprises in 2013, 150 social enterprises in 2014, 185 social enterprises in 2015, and 259 social enterprises in 2016 are analyzed. In the VMD data, we use operating profit and net income as economic performance variables. Since there is no social performance measure in monetary unit in the VMD data, we construct one using

[7] Social performances includes the number of vulnerable workers employed, their working hour and wage, the number of beneficiaries of social services, the reinvestment of profits on social purposes, and so on.

the minimum wage,[8] the number of disadvantaged workers, the average wage and working hours of disadvantaged workers. Based on our definition of social value, our social performance measure is given by $(w_d - w_*) \times n$, where w_d is the average wage of disadvantaged employees, w_* is the minimum wage, and n is the number of disadvantaged employees. Table 11.1 shows the descriptive statistics of employment type social enterprises used in our analysis.

We analyze the relation between government subsidies and social enterprises' performances. In our theoretical model, social enterprises' weight on social value is an important factor determining their performances. However, we cannot obtain the weights in the VMD data. Assuming the weights on social value do not vary over time, the fixed effect regression model allows us to do without them. The following regression equation is to be estimated:

$$y_{it} = \beta' X_{it} + \gamma' Z_{it} + \alpha_i + e_{it},$$

where y_{it} denotes an economic (or social) performance of social enterprise i in year t; X_{it} is a vector of government financial supports such as subsidy for personnel expenses, subsidy to supplement the payment for specialists, subsidies for social insurance fees and business development activities; Z_{it} is a vector of other factors of social enterprises including total subsidy from private sectors, being commercial companies or not (dummy variable), and sales; α_i denotes the time-invariant unobserved effect of the weight on social value; and e_{it} is a random error term. Table 11.2 shows the relation between government subsidies and social enterprises' performances.

Our empirical result shows that most government subsidies except subsidy for social insurance fees have no significant effects on social performance. On the other hand, some government subsidies are found to have significant effects on economic performances. To be specific, subsidy for personnel expenses, subsidy to supplement the payment for specialists, and other subsidy have negative effects on operating profit, while subsidies for personnel expenses and social insurance fees have positive effects on net income.

If the main purpose of government subsidies is to induce social enterprises to produce more social value, it is necessary to redesign the subsidy scheme as proposed in Sect. 11.4. However, if the main purpose is to guarantee sustainability of social enterprises, the government policy can be recognized to be valid because subsidies for personnel expenses and social insurance fees are shown to have a significant positive impact on net income. It should be noted that these results cannot be applied to all social enterprises certified in Korea, since we analyzed only certified employment type social enterprises that have participated in VMD.

[8]The income increase of disadvantaged workers, that is, social value produced by social enterprises should be measured using the market wage of disadvantaged workers. Since we could not obtain their market wage (i.e., their productivity), we inevitably used the minimum wage. Also we did not take into account credit constraint of disadvantaged workers and secondary social value. Therefore social value used in our analysis can be interpreted as the most conservative estimation of true social values.

Table 11.1 Descriptive statistics of employment type social enterprises in the VMD data

Year		2013	2014	2015	2016
Observations (N)		67	150	185	259
Commercial companies (%)		68.66	69.33	74.59	71.43
Regular worker	Number of employees	15.09	9.93	12.75	9.51
	Wage (kW)	1608.70	1589.02	1675	1765.75
	Work hours per week	37.36	34.37	37.47	36.20
Disadvantaged worker	Number of employees	28.84	17.47	23.17	18.05
	Wage (kW)	1218.21	1277.03	1362	1434.19
	Work hours per week	36.40	37.16	36.70	35.58
Economic performance (kW)	Sales	2071100	1214369	1717505	1468712
	Operating profit	−91498	−69855	−43612	14195
	Net income	58201	45011	62897	44331
Social performance (kW)		8583	3110	11449	4072
Governmental subsidy (kW)	Personnel expenses	65443	52011	46617	35393
	Specialists	10041	6212	7023	9396
	Business development	16016	15768	14577	11777
	Social insurance fees	9352	6599	10187	5346
	Others	64734	50140	38732	45820
Private sector subsidy (kW)	Corporate	2042	1572	2354	3297
	Parent corporate	3102	1396	777	1118
	General donation	2197	3360	1239	1338
	Others	14173	13942	16721	11548

*kW: thousand Korean won (approximately equivalent to 0.9 USD)

11.7 Conclusion

We have investigated an employment model of social enterprises. These firms pursue welfare improvement of disadvantaged workers by employing them and paying them higher wage than their market wage. It is assumed that social enterprises maximize the weighted average of profit and social value.

Table 11.2 Government subsides and social enterprises' performances

	Social value	Operating profit	Net income
Subsidy for personnel expenses	−0.0071 (0.0098)	−0.4869*** (0.1420)	0.2152** (0.1189)
Subsidy for specialists	0.0243 (0.0365)	0.8381 (0.5269)	−0.0544 (0.4413)
Subsidy for business development	0.0023 (0.0225)	−0.8010** (0.3244)	−0.1879 (0.2718)
Subsidy for social insurance fees	0.0325* (0.0169)	−0.0839 (0.2442)	0.7613*** (0.2045)
Other subsidy	−0.0031 (0.0061)	−0.5207*** (0.0875)	0.0825 (0.0733)
Private donation	0.0025 (0.0139)	−0.1548 (0.2006)	0.0679 (0.1680)
Commercial company (dummy)	819.27 (3145.72)	−74821.18 (45418.51)	−28348.8 (38044.85)
Sales	0.0010 (0.0006)	0.0843*** (0.0091)	0.0708*** (0.0076)
Constant	4163.29 (2533.61)	−53732.55 (36580.67)	−55342.72 (30641.9)
Observations (N)	691	691	691
Groups	396	396	396
R-squared	0.2773	0.2736	0.3834
sigma_α	62102.39	232545.84	172285.86
sigma_e	6328.86	91377.29	76542.45
rho	0.9897	0.8662	0.8352

Note ( ) standard errors. significance levels, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Paying the disadvantaged workers higher wage than their market wage improves their welfare and such welfare improvement is larger than the wage increase because they have the shortage of satisfying their basic needs due to credit constraint. This is one channel in which social enterprises can improve social welfare. Social enterprises also mitigate under-production problem in the oligopoly, which also results in social welfare improvement.

Social welfare in the mixed oligopoly with both for-profit and social enterprises varies depending on the weight on social value. There exist some weights that admit a higher social welfare in the mixed oligopoly than in the oligopoly with only for-profit enterprises. The optimal weight on social value is the weight that maximizes social welfare in the mixed oligopoly. Nevertheless, if social enterprises have too high or too low weight on social value, social welfare may be affected negatively, even lower than in the oligopoly. We show that social value subsidy can play a positive role in improving social welfare by providing an optimal incentive of producing social value.

Social value subsidy can be applied not only to the employment type social enterprises but to the other types of social enterprises such as social service provision or local community contribution types. A comprehensive analysis on the effect of social value subsidy is left for future investigation. Our numerical examples provide the possibility of financially sustainable social enterprises and their role of welfare improvement upon the oligopoly. Nevertheless, we leave the general investigation on the financial sustainability of social enterprises for future investigation.

We have considered only the quantity competition oligopoly. Nevertheless, the same results are expected in the price competition oligopoly (Bertrand model). This is because our results can be applied in the oligopoly with arbitrary number of firms, the limit of which is the perfectly competitive market. The Bertrand price competition yields the same efficient allocation as in the competitive market. Nevertheless, social enterprises will not be sustainable since their profits are below zero. Social value subsidy can be used in this case to improve financial sustainability of social enterprises as well as social welfare.

Although there may be various ways to support for social enterprises, we have provided the most basic subsidy scheme where social value is compensated at a fixed rate. We have shown that it can have a positive impact on social value and social welfare. In Korea, government supports for social enterprises are designed somewhat differently from our scheme. Using Korean data and a fixed effect regression model, we analyze the relation between government subsidies and economic and social performances of social enterprises certified by the Korean government under the Social Enterprises Promotion Act. We find that only subsidy for social insurance fees has a positive impact on the social performance of social enterprises. Subsidy for personnel expenses, the major part of the government subsidy, has neither positive nor negative impact. We also show that subsidies for personnel expenses and social insurance fees have a positive impact on net income of social enterprises. Our theoretical and empirical results suggest that the current scheme of governmental financial supports should be redesigned in the way we suggest here.

A different approach can be adopted to model social enterprises. Examples are social enterprises that maximize their social goal under the budget constraint (non-negative profit) as in Chu (2015), social enterprises that maximize profit at first and use their profit to acquire social goal, and social enterprises that maximize the well-being of their members as a co-operative, etc. Cho and Lee (2017) assume that social enterprises maximize the sum of consumer surplus and social value which is proportional to production, under a budget constraint.

Finally, we assumed that all social enterprises share an identical weight on social value. In reality, social enterprises make decisions based on the participation of various stakeholders. Therefore, the weight given to social values may vary from one social enterprise to another. Dealing with heterogeneous social enterprises and the competition among them are also left for future study.

Acknowledgements This chapter is based on our article "An Employment Model of Social Enterprises and the Effects of Government Subsidy", *Journal of Korean Economic Analysis*, Vol. 23(3), Dec., 2017 (in Korean). The empirical analysis in the published article has been updated with more recent data. We are grateful to Sung-Jin Cho, Wonki Cho, Jinwoo Kim and Woojin Lee for many valuable comments and suggestions. Biung-Ghi Ju acknowledges financial supports from Center for Distributive Justice, Institute of Economic Research at Seoul National University and from Korean Research Foundation (NRF-2016S1A3A2924944). Hyunwoo Hong acknowledges financial support from Korean Research Foundation (NRF-2016H1A2A1906832).

References

Alter, K. 2007. Social Enterprise Typology, Virtue Ventures.

Cho, S.M., and S.H. Lee. 2017. Subsidization policy on the social enterprise for the underprivileged. *Korean Economic Review* 3 (1): 153–178.

Chu, C.Y. 2015. Cyrus, Warren Buffet versus Muhammad Yunus. *Journal of Institutional and Theoretical Economics* 171 (4): 696–708.

Davister, C., J. Defourny, and O. Grégoire. 2004. Work integration social enterprises in the European Union: an overview of existing models. *RECMA-Revue Internationale de l'Economie Sociale* 293: 24–50.

Haugh, H. 2006. Social enterprise: Beyond economic outcomes and individual returns. Social entrepreneurship, 180–205. UK: Palgrave Macmillan.

Hong, H.W., and B.G. Ju. 2016. An economic analysis of social enterprises: Social service provision type. *Korean Journal of Public Finance* 9 (1): 87–112. (in Korean).

Hong, H.S., and Y.K. Kim. 2016. A study on the social and economic performance of social enterprise. *Journal of Finance and Accounting Information* 16 (1): 1–29. (in Korean).

Hur, M.H., and G.S. Yang. 2015. Exploring influential factors of social enterprise on job creation. *Korean Journal of Public Administration* 24 (3): 121–146. (in Korean).

Kaneda, M., and A. Matsui. 2003. Do profit maximizers maximize profit? *Divergence of objective and result in oligopoly*. Unpublished manuscript.

Kato, K., and Y. Tomaru. 2007. Mixed oligopoly, privatization, subsidization, and the order of firms' moves: Several types of objectives. *Economics Letters* 96 (2): 287–292.

Kim, J.I. 2014. A study on the performance difference depend on the characteristics and support type of social enterprise: Focusing on the social enterprise notifying management information autonomously. *Journal of Korean Social Welfare Administration* 16 (2): 181–212. (in Korean).

Kim, H.C. 2015. A study on performance analysis of social enterprises. *Korean Journal of Business Administration* 28 (7): 1797–1812. (in Korean).

Kim, J.H., and J.K. Lee. 2012. The analysis of economic and social performances of social enterprises. *Korean Public Administration Quarterly* 24 (4): 1037–1063. (in Korean).

Maslow, A. 1954. *Motivation and personality*. New York: Harper.

Maslow, A. 1943. A theory of human motivation. *Psychological Review* 50 (4): 370–396.

Matsumura, T. 1998. Partial privatization in mixed duopoly. *Journal of Public Economics* 70: 473–483.

Matsumura, T., and O. Kanda. 2005. Mixed oligopoly at free entry markets. *Journal of Economics* 84: 27–48.

Nussbaum, M. 1992. Human functioning and social justice. In defense of Aristotelian essentialism. *Political Theory* 20 (2): 202–246.

Rawls, J. 1971, 1999. *A theory of justice*. Cambridge: Harvard University Press.

Sen, A. 1987. The standard of living. In *The standard of living: The tanner lectures on human values*, ed. Sen, Muellbauer, Kanbur, Hart, and Williams. Cambridge: Cambridge University Press.

Tay, L., and E. Diener. 2011. Needs and subjective well-being around the world. *Journal of Personality and Social Psychology* 101 (2): 354–365.

White, Mark D. 1996. Mixed oligopoly, privatization and subsidization. *Economics Letters* 53: 189–195.

Chapter 12
Congestion, Technical Returns, and the Minimum Efficient Scales of Local Public Expenditures: An Empirical Analysis for Japanese Cities

Masayoshi Hayashi

12.1 Introduction

Most industrial countries have restructured their local government systems through a series of subnational government mergers. In so doing, policy makers typically anticipate that merging fiscal units would exploit economies of population scale in public service production (Blom-Hansen et al. 2016). Taking advantage of the data obtained from the developments of municipal mergers in various countries, the recent literature has been examining whether municipal mergers reduced per capita local government expenditures. However, such studies do not reveal if economies of population scale are present. Following Duncombe and Yinger (1993), economies of population scale exist if per capita cost is decreasing for the local population, *provided that all other variables in the local government cost function are maintained constant*. Such "other" variables typically include the level of public service, levels of input prices (e.g., wages of local government employees), and local characteristics that affect the technology of public service delivery (e.g., surface area and demographics). Obviously, a municipal merger would change all these variables in the merging municipalities by construction. Contrary to the claim made in the recent literature, therefore, we cannot easily identify economies of population scale by examining how per capita local government expenditure changes after mergers.

To identify the existence of economies of scale in public production, we need to estimate the structural parameters of production technology and consumption technology of local public services, which is what we do in this study. Our theoretical framework borrows from the standard model (e.g., Bradford et al. 1969; Brueckner 1981; Duncombe and Yinger 1993), which distinguishes between the direct outputs produced by a government and the level of the public services actually consumed by citizens. Indeed, since Hirsch (1959, 1965) raised the issue of possible economies of

M. Hayashi (✉)
The University of Tokyo, Tokyo, Japan
e-mail: hayashim@e.u-tokyo.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_12

scale by estimating the “U-shaped” per capita local public expenditures, the literature has focused its effort on the measurement of returns to scale in local public production as well as the identification of the population size at the bottom of the U-shape, or the “minimum efficient scale” (MES). However, less consideration has been given to relating the factors that affect the returns to scale to the construction of an indicator for the efficient scales of local public production. In this study, we therefore examine the factors that account for the “U-shaped” per capita local public expenditures and construct an indicator that quantifies a deviation from the MES of a local population for public expenditures.

We develop our discussion as follows. In Sect. 12.2, our analysis of the U-shaped per capita expenditures highlights two important concepts: (i) technical elasticity of scale, which characterizes production technology for the direct output, and (ii) congestion elasticity, which characterizes consumption technology for the public services. In particular, we articulate that population-induced changes in per capita cost are related to the relative magnitude between the two elasticities, and derive an indicator for efficiency scale that quantifies the distance of a local population from its MES, an indicator which turns out to be a simple ratio of the congestion elasticity and the technical scale elasticity. In Sect. 12.3, we describe the procedure to implement our analysis to obtain the efficient scale scores with the data of Japanese cities. In Sect. 12.4, we list the estimation results and rank Japanese cities according to the scores obtained from the results. Section 12.5 then concludes with some cautionary note.[1]

12.2 Theoretical Framework

We employ the standard model of local public production (e.g., Bradford et al. 1969; Brueckner 1981; Duncombe and Yinger 1993). This model distinguishes between the direct outputs produced by a government and the public service level consumed by citizens, which implies that the public production-consumption process can be decomposed into two stages.

The first stage is analogous to the standard theory of the firm. As in private production, a local government produces a direct output g, with a technology:

$$g = g(\boldsymbol{x}) \tag{12.1}$$

where $\boldsymbol{x}$ is a vector of factor inputs. The expenditure (or cost) function is thus derived as:

$$c = c(g, \boldsymbol{w}) \equiv \{\min_{\boldsymbol{x}} \boldsymbol{w} \cdot \boldsymbol{x} | g(\boldsymbol{x}) = g\} \tag{12.2}$$

[1] This is a updated version of Hayashi (2002a, b), which has been cited by a number of Korean studies and documents issued by international organizations, but have been unavailable as a discussion paper in English version (Hayashi 2002a). While the text in this version has been edited, the analysis and the estimation results have been kept intact as those estimated in the original discussion paper.

where $\boldsymbol{w}$ a vector of factor input prices. The standard properties apply to functions (12.1) and (12.2).

The second stage transforms the direct output g into its service level z, which is of interest to citizen-consumers. Note that this transformation may be influenced by the number of consumers (i.e., population) n and other local characteristics or environmental factors $\boldsymbol{a}$. This relation, known as the "crowding function" or "congestion function," is expressed as:

$$z = z(g, n, \boldsymbol{a}) \tag{12.3}$$

where $\partial z/\partial g > 0$ and $\partial z/\partial n < 0$. From this, we derive the following inverse function with respect to g:

$$g = \gamma(z, n, \boldsymbol{a}) \tag{12.4}$$

This shows the level of direct production that is necessary to keep the consumed service level z for a given set of population n and local characteristics $\boldsymbol{a}$. To see the effects of population on the public expenditures, we substitute (12.4) into (12.2) to obtain:

$$c = c(\gamma(x, n, \boldsymbol{a}), \boldsymbol{a}). \tag{12.5}$$

Returns to scale are defined in terms of changes in average cost when the level of an "output" varies. As taxonomized in Duncombe and Yinger (1993), there are three dimensions for returns to scale in local public production: (i) technical returns to scale over g, (ii) returns to population scale over n, and (iii) returns to quality scale over z. The "U-shaped" per capita expenditure is associated with the second concept and the partial derivative which is characterized as:

$$\frac{\partial[c\gamma(z, n, \boldsymbol{a})]/n)}{\partial n} = \frac{c/n}{n} \cdot \left(\frac{\eta_n^g}{\eta_c^g} - 1\right) \tag{12.6}$$

where

$$\eta_n^g \equiv \frac{\partial \gamma}{\partial n} \cdot \frac{n}{g}, \text{ and} \tag{12.7}$$

$$\eta_c^g \equiv \frac{\partial g}{\partial c} \cdot \frac{c}{g}. \tag{12.8}$$

This shows that the direction of population-induced changes in per capita cost is related to the relative magnitude between the two elasticities, η_n^g and η_c^g.

The first elasticity (12.7) is a *congestion elasticity* characterized by the consumption technology (12.3)–(12.4). It measures the elasticity of g with respect to n, keeping z and $\boldsymbol{a}$ constant, and shows how much direct product needs to be increased in order to keep the service level as before, when population increases. On the other hand, the very definition that congestion refers to a negative impact of a population increase on the level of public service may imply

$$\eta_n^z \equiv -\frac{\partial z}{\partial n} \cdot \frac{n}{z} \tag{12.9}$$

which is the elasticity of z with respect to n (in absolute value), keeping g and $\boldsymbol{a}$ constant. However, (12.7) outperforms (12.9) in that the former is metric-free while the latter is not (Reiter and Weichenrieder 1999). Note that as only ordering is required when z is evaluated by citizen-consumers, any variable that is a positive monotonic transformation of z can validly replace z. However, (12.9) is invariant *only* to multiplicative changes of z. As such, when we employ (12.9) rather than (12.7), the value of the elasticity will arbitrarily differ depending on the monotonic transformation.

The second elasticity (12.8) is technical elasticity of scale, which is associated with the production technology (12.1). Increasing technical returns to scale (or technical economies of scale) imply that $\partial(c/g)/\partial g < 0$. A natural expression for its elasticity will then be given as $\eta_g - [\partial(c/g)/\partial g] \cdot [g/(c/g)]$, which can be related to (12.8) as:

$$\eta_g = 1 - \frac{1}{\eta_c^g}. \tag{12.10}$$

This shows that the values of η_c^g and η_g are monotone, allowing us to regard the former as a valid measure for technical economies of scale.

Equation (12.6), in fact, is a different restatement of what is derived in Duncombe and Yinger (1993).[2] Our articulation, however, may be more instructive. First, we may easily relate changes in per capita expenditures to the relative magnitudes between η_n^g and η_c^g. Specifically, per capita cost is decreasing (or returns to population scale are increasing) when technical elasticity of scale is greater than congestion elasticity ($\eta_c^g > \eta_n^g$). Furthermore, per capita cost is increasing (or returns to population scale are decreasing), when technical elasticity of scale is less than congestion elasticity ($\eta_c^g < \eta_n^g$). While a typical explanation in the literature relates U-shaped per-capita local expenditures to some returns to scale and congestion, Eq. (12.6) indeed formalizes the fact that the U-shaped curve hinges on the balance between the two measures, η_n^g and η_c^g.

Second, our expression helps construct an indicator that quantifies a deviation of a local population from its MES for local public expenditures. By multiplying both sides of Eq. (12.6) by *n*/(*c*/*n*), we obtain an elasticity expression for the population returns to scale:

$$\frac{\partial c/n}{\partial n} \cdot \frac{n}{c/n} = \frac{\eta_n^g}{\eta_c^g} - 1. \tag{12.11}$$

We may then define an efficiency indicator for a local population as:

---

[2]Duncombe and Yinger (1993) use an elasticity expression that is analogous to Eq. (12.10). Their expression is $[\partial(c/n)/\partial n] \cdot n/(c/n) = \eta_c^g \cdot \eta_n^g - 1$, where the elasticity for technical returns to scale is defined differently as $\eta_c^g \equiv (\partial c/\partial g) \cdot (g/c)$, a reciprocal of our measure.

$$S(n) \equiv \frac{\eta_n^g}{\eta_c^g}. \tag{12.12}$$

We then see from Eq. (12.6) that n^*, such that $S(n^*) = 1$ (i.e., $\eta_n^g = \eta_c^g$), is the MES, provided that (i) both η_n^g and η_c^g are not constant, (ii) η_n^g is non-decreasing in n, and (iii) η_c^g is non-increasing in n. We also see from Eq. (12.6) that a local population is too large (small) if $S(n) > (<) 1$. Thus, we may treat $S(n)$ as an indicator that quantifies the deviation of a local population size from its MES.

Notice that the MESs differ across jurisdictions in so far as the local factors that influence the two elasticities also differ. Those factors generally consist of the variables that appear in Eq. (12.5). However, they do not necessarily include all of the variables, which depend on the specifications of the production technology (12.1).

12.3 Empirical Implementation

On the basis of the foregoing analysis, we estimate the urban public production structure in Japan. With the estimates obtained, we calculate the efficient indicators for the cities whose data are used for the estimation and rank them according to the resultant values of the calculated scores.

Specifications

To implement the foregoing analysis, we need to estimate the technical elasticity of scale and congestion elasticity of population. To do so, we have to specify the cost function (12.2) and the congestion function (12.3). Let us start with the congestion function. A popular specification for the congestion function is $z = n^{-\alpha} g$ (Borcherding and Deacon 1972; Bergstrom and Goodman 1973). This, in effect, assumes that congestion decreases at the margin, and its elasticities are constant at α. However, congestion may or may not accelerate with the intensity of use, and its elasticities should not be constant. In addition, as discussed in the previous section, factors other than population may influence congestion.

We thus specify our congestion function as:

$$z(g, n, \mathbf{a}) = g \cdot n^{-\left(\lambda_0 + \lambda_n \ln n + \sum_j \lambda_j a_j\right)} \cdot e^{-\sum_j \delta_j a_j} \tag{12.13}$$

where a_j represents environment factors (i.e., is an element of vector $\boldsymbol{a}$), and λ_j and δ_j are parameters to be estimated. First, this specification of the congestion function is a proportional metric: $z(g, n, \boldsymbol{a}) = g \cdot \xi(n, \boldsymbol{a})$. Equation (12.3) yields $\eta_n^g = \eta_n^z / \eta_g^z$ where $\eta_g^z \equiv (\partial z / \partial g)(g/z)$. Since $\eta_g^z = 1$ if $z(g, n, \boldsymbol{a}) = g \cdot \xi(n, \boldsymbol{a})$, it is immediate that the two congestion elasticities (12.7) and (12.9) coincide (cf., Reiter and Weichenrieder 1999). Second, we can test whether congestion elasticity is increasing or decreasing. Given specification (12.13), the congestion elasticity is given as $\eta_n^g = \lambda_0 + 2\lambda_n \ln n + \sum_j \lambda_j a_j$, which enables us to test whether congestion is increasing or decreasing by examining the sign of λ_n (e.g., Hayes and Slottje 1987).

Second, we allow the environmental factors to affect the elasticity through $\sum_j \lambda_j a_j$ (e.g., Hayes 1986; Duncombe and Yinger 1993). Note that this also allows for the possibility of a "camaraderie" effect (Edwards 1990), since if some of the λ_0s and λ_js are negative, sharing congestible goods may actually increase benefits to each user for a given combination of values of the environmental factors.[3] Finally, we include $e^{-\sum_j \delta_j a_j}$ in (12.13) so that the environmental factors multiplicatively influence the level of public services.

Following the convention of the literature, we aggregate factor inputs into labor and capital, which is reflected as two factor prices, w for labor and r for capital, in the cost function. For the specification of the cost function (12.2), we first assume the translog form:

$$\begin{aligned}\ln c = \beta_0 + \beta_g \cdot \ln g + \beta_w \cdot \ln w + \beta_r \cdot \ln r \\ + \beta_{gg} \cdot (\ln g)^2 + \beta_{ww} \cdot (\ln w)^2 + \beta_{rr} \cdot (\ln r)^2 \\ + \beta_{gw} \cdot \ln g \ln w + \beta_g \cdot \ln g \ln r + \beta_{wr} \cdot \ln w \ln r\end{aligned} \tag{12.14}$$

where the βs are parameters to be estimated. Note that the price of the capital r could be regarded as uniform across regions within a country in a given period (e.g., Kitchen 1976; Stevens 1978). As such, we can forego obtaining the data for r since our data are cross-sectional. This means that some of the coefficients are not identifiable, and the feasible specification is reduced to

$$\ln c = A_0 + A_w \ln w + \beta_{ww} (\ln w)^2 + \left(A_g + \beta_{gw} \ln w\right) \cdot \ln g + \beta_{gg} (\ln g)^2 \tag{12.15}$$

where coefficients A_0, A_g, and $A_w$correspond respectively to $\beta_0 + \beta_r \cdot \ln r + \beta_{rr} \cdot (\ln r)^2$, $\beta_g + \beta_{gr} \cdot (\ln r$, and $\beta_w + \beta_{wr} \cdot (\ln r$in the original specification.

We next obtain the log for (12.13) and rearrange the resultant terms to obtain:

$$\ln g = \ln z + \left(\lambda_0 + \lambda_n \ln n + \sum_j \lambda_j a_j\right) \cdot \ln n + \sum_j \delta_j a_j. \tag{12.13$'$}$$

By substituting this into (12.17), indexing its variables with subscript i, and adding error term u_i, we finally obtain the following statistical specification for the cost function:

$$\begin{aligned}\ln c = A_0 + A_w \ln w + \beta_{ww} \cdot (\ln w)^2 \\ + \left(A_g + \beta_{gw} \ln w\right) \cdot \left(\ln z + \left(\lambda_0 + \lambda_n \ln n + \sum_j \lambda_j a_j\right) \cdot \ln n + \sum_j \delta_j a_j\right) \\ + \beta_{gg} \left(\ln z + \left(\lambda_0 + \lambda_n \ln n + \sum_j \lambda_j a_j\right) \cdot \ln n + \sum_j \delta_j a_j\right)^2 + u_i.\end{aligned} \tag{12.16a}$$

[3] While Edwards (1990) considers the question of increasing congestion and that of camaraderie effects by setting up *separate* specifications, our *single* specification allows for both possibilities.

We can test the forms of production technology by imposing coefficient restrictions. Specifically, we test the Cobb–Douglas form against the translog specification with restriction $\beta_{ww} = \beta_{gw} = \beta_{gg} = 0$. If the restrictions are valid, the specification will be:

$$\ln c = B_0 + B_w \ln w + \\ + \beta_g \cdot \left(\ln z + \left(\lambda_0 + \lambda_n \ln n + \sum\nolimits_j \lambda_j a_j\right) \cdot \ln n + \sum\nolimits_j \delta_j a_j\right) + u_i \tag{12.16b}$$

where $B_0 = \beta_0 + \beta_r \cdot \ln r$.

12.3.1 Data

We need data for total cost c_i, population n_i, wages for local public officials w_i, public service consumed z_i, and a set of environmental factors $\boldsymbol{a}$. The data we utilize are obtained from a cross section of Japanese cities in 1995. Although there are more than 600 cities in Japan in that year, the availability of the variables used for the estimation reduces the sample size down to 572. The data descriptions and their sources are listed in Table 12.1. The details of each variable are explained as follows.

Total cost and population: We examine the cost for overall public services provided by city governments. The total cost c_i is naturally represented by the total expenditures of a city government, listed in the Ministry of Internal Affairs and Communication (MIC) (1997). Our data for population n_i are drawn from the Statistics Bureau, Management and Coordination Agency (1997).

Wages: The price of public labor w_i is obtained as average wages, that is, the compensation of public employees divided by the number of public employees (MIC 1997). This implicitly assumes that the average working hours are identical across the cities under examination, which may be justified given the fact that labor conditions for local public employees seem to be similar across municipalities.

Public service outputs: Obtaining the data for the consumed level of total public service z_i is not straightforward. We use the "total score of public services" provided by Nihon Keizai Shinbunsya (1998). This score quantifies the overall quality level of Japanese local government outputs, by considering the total of 24 public service categories, which include fees for charged public services, per capita capacities of welfare and education facilities, and per capita levels of infrastructures like roads, parks, water supply, and sewage facilities. The data for each category are transformed to standardized values. If the standardized value is less than 20, only 1 point is given; 20–40, 2 points; 40–60, 3 points; 60–80, 4 points; and 80 and over, 5 points. The final score for a jurisdiction is the total of these points scored in the 24 categories. While this score may not be a good measure for the actual quality of the public services, we believe it should give us a reasonable evaluation of local public services. In addition, it is the only available data for public service quality.

Table 12.1 Data description

Variables	Descriptions	Sources
c	Total public expenditures	a
n	Population	b
w	Average wages for public employees	a
z	Total score of public services	d
a_1	Log of total area of a city	b
a_2	Ratio of "distant-island area"	e
a_3	Ratio of "heavy-snow area"	e
a_4	Ratio of forested area	e
a_5	Ratio of farm land	e
a_6	Ratio of DID population	b
a_7	Ratio of DID	b
a_8	Ratio of daytime population	b
a_9	Ratio of 15 years or younger	b
a_{10}	Ratio of 65 years or older	b
a_{11}	Log of per capita income	f
a_{12}	Ratio of manufacturing labor	b
a_{13}	Ratio of service labor	b
D_1	Designated cities (dummy)	–
D_2	Special wards (dummy)	–
D_3	Great Kobe earthquake (dummy)	–

Sources a: Ministry of Internal Affairs and Communication; b Statistic Bureau, Management and Coordination Agency; c: National Land Agency; d: Nihon Keizai Shinbunsya; e: Ministry of Agriculture, Forestry and Fishery; f: Ministry of Finance. Note that DID stands for "Densely Inhabited Districts"

Environmental factors: The environmental factors fall into four groups: physical factors, urban factors, demographic factors, and economic factors. We try to capture the physical environment by considering the total land area in log form (a_1), as well as the following shares of sub-areas: "distant-island" area (a_2), "heavy-snowfall" area (a_3), forested area (a_4), and farm lands (a_5). The "distant-island" and "heavy-snowfall" areas are subdivisions in a jurisdiction designated as such by national laws, which should be a convenient set of indicators for harsh natural environments. Note that the *number* of the subdivisions is used to calculate shares a_2 and a_3. For example, if a city has five subdivisions and one of them is designated as "distant-island," then the share is 1/5. For a_4 and a_5, on the other hand, we simply use *areas* to compute shares, that is, percentages of forested area and farm land in the total area of a city. As urban factors, we consider three elements. They are the proportion of population in Densely Inhabited Districts (DID) population (a_6), the share of the areas that DIDs cover (a_7), and the ratio of daytime population to nighttime population (a_8).

For demographic factors, we only consider percentages of the young, that is, those aged less than 16 years (a_9), and the elderly, that is, those aged above 64 years (a_{10}). Finally, economic factors are controlled by the log of per capita income (a_{11}), and the proportions of workers engaged in the manufacturing (a_{12}) and service sectors (a_{13}).

Additional dummies: Although not explicitly written in (12.16a) or (12.16b), we add three dummy variables that affect the intercept of the regression function (A_0 or B_0). Two dummies are intended to capture the different administrative roles assumed by two special types of Japanese cities. One, (D_1), is for a group of 12 *designated cities* that have a longer list of expenditure authorities than ordinary cities do. Another, (D_2), is for 23 *special wards (ku)* in Tokyo, which delegates some of their expenditure responsibilities to the Tokyo metropolitan government.[4] The last is the dummy for the Great Kobe Earthquake in January 1995 (D_3). While our fiscal data are for the 1995 fiscal year (April 1995–March 1996) and do not cover the time of the earthquake, it is plausible that the damages were so serious that local public expenditures in the inflicted areas in the following fiscal year increased owing to the special needs caused by the disaster. We thus include a dummy variable that indicates such inflicted areas.

12.3.2 Specification Tests

Since we use cross section data from a variety of cities, we may legitimately assume that the error term u_i follow some pattern of heteroskedasticity. However, as there are several variables to consider in our nonlinearly specified regression functions, determining a specific pattern of the skedastic function is a difficult task. We therefore assume that the pattern of the heteroskedasticity is unknown, and employ a method that allow for unknown patterns of heteroskedasticity. Before presenting the final estimates, we perform specification tests to choose between (12.16a) and (12.16b) to determine the type of production technology (i.e., translog or Cobb–Douglas functions). The null hypothesis is that the technology is a Cobb–Douglas function ($\beta_{ww} = \beta_{gw} = \beta_{gg} = 0$). If the null hypothesis is not rejected, we opt for (12.16b). Otherwise, (12.16b) is chosen. Our test is based upon the Lagrange multiplier (LM) principle. To construct relevant LM test statistics, we utilize the Gauss-Newton regression (GNR) that allows for heteroskedasticity, that is, the heteroskedasticity-robust Gauss-Newton regression (HRGNR).[5] The HRGNR shows that the Cobb-Douglas technology is not rejected at the standard significance levels with a P value

[4]For example, while police and refuse collection are responsibilities for cities, towns, and villages, they are taken care of by the Tokyo metropolitan government in the special wards (note that refuse collection has now become a responsibility of the special districts, starting in 2000).

[5]See Davidson and MacKinnon (1993, pp. 399–402) for a textbook explication.

of 0.115. We thus base our discussion on (12.16b), whose estimates are obtained by the method of non-linear least squares, as discussed in the following section.[6]

12.4 Results

The estimation results for (12.16b) are listed in Table 12.2, where the coefficients for D_1, D_2, and D_3 are indicated as b_1, b_2, and b_3, respectively. Note that P values are based upon the heteroskedasticity consistent covariance matrix estimator (HC-CME). The estimates for the production technology, namely B_0, β_w, and β_g, have the expected signs and are all statistically significant. Also, the latter two estimates are within the range suggested by the theory. Note that since we have opted for the Cobb-Douglas technology, the values for technical elasticity of scale η_c^g are constant and equal to the reciprocal of η_g. With $\beta_g = .328$, the value of the elasticity is $\eta_c^g = 1/\beta_g = 3.049$, which implies that $\partial(c/g)/\partial g < 0$ by Eq. (12.10). That is, the technical returns to scale are increasing and technical economies of scale exist in the Japanese local public production.

The estimate for λ_n is positive and statistically significant at the significance levels larger than 0.011. Given the congestion elasticity of city i is $\eta_n^g = \lambda_0 + 2\lambda_n \cdot \ln n_i + \sum_j \lambda_j a_{ji}$, this positive value ($\lambda_n > 0$) implies that congestion accelerates as population increases. Note that since the values for the λ_0s and some λ_js are negative, the value of η_n^g can be negative if $2\lambda_n \cdot \ln n_i < -\lambda_0 - \sum_j \lambda_j a_{ji}$, allowing for the possibility of the camaraderie effect. We calculate the congestion elasticities for the 572 cities that constitute the observations in our sample. Table 12.3 reports the summary statistics of the 572 η_n^gs which, with a minimum of 1.269, show no evidence of the camaraderie effect.

Several, but not all, of the coefficients on the environment factors (λ_js and δ_js) are statistically significant, implying that the degree of congestion depends on factors other than population. It is also interesting to note that the environmental factors that have significant effects through δ_j also do so through λ_j. These factors are the city's total area, share of forested area, daytime population ratio, and elder population ratio. Note that the effect of a marginal change in one of the environmental factors is not easily read from the coefficient estimates, since it also depends on the values of β_w and $\ln n_i$:

$$\frac{\partial \ln c_i}{\partial a_{ji}} = \beta_g \cdot \left(\lambda_j \cdot \ln n + \delta_j\right) \tag{12.17}$$

We calculate $\partial \ln c_i / \partial a_{ji}$ for the 572 cities for each environmental factor. Table 12.4 provides the summary statistics, which clearly show that cities with different populations implies different and sometimes opposite effects of the environmental factors.

[6]Note that we assume the output variable z_i is independent of total cost c_i and therefore of error term u_i, which is a common assumption in the literature. The basis for this assumption is a public

Table 12.2 Estimation results

Coef.	Estimates	P value	Coef.	Estimates	P value
B_0	21.942***	0.000			
β_w	0.179**	0.035			
β_g	0.328***	0.002			
b_1	0.272***	0.000			
b_2	0.57	0.371			
b_3	0.281***	0.001			
λ_n	0.375**	0.011			
λ_0	$-8.254**$	0.012			
λ_1	$-0.242**$	0.031	δ_1	3.013**	0.023
λ_2	−0.167	0.480	δ_2	2.012	0.414
λ_3	0.014	0.841	δ_3	0.013	0.987
λ_4	0.702*	0.078	δ_4	$-8.250*$	0.075
λ_5	0.761	0.207	δ_5	−8.998	0.194
λ_6	−0.247	0.297	δ_6	3.019	0.362
λ_7	−0.174	0.533	δ_7	2.548	0.430
λ_8	0.254***	0.006	δ_8	$-2.551***$	0.007
λ_9	5.310*	0.090	δ_9	−56.605	0.106
λ_{10}	4.516**	0.033	δ_{10}	$-45.176**$	0.047
λ_{11}	0.203	0.333	δ_{11}	−2.451	0.314
λ_{12}	1.707	0.163	δ_{12}	−18.105	0.166
λ_{13}	1.290	0.270	δ_{13}	−13.224	0.288
Adjusted R^2		0.974			
Number of observations		572			
Degree of freedom		538			

$***p < 0.01$; $**0.01 \leq p < 0.05$; $*0.05 < 0.10$

Table 12.3 Summary statistics of congestion elasticities for 572 cities

Congestion elasticity	Mean	Std. Error.	Minimum	Maximum
η_n^g	2.607	0.561	1.269	8.584

Lastly, the coefficient estimates for the dummies for the designated cities (b_1) and the Kobe Earthquake (b_3) have the expected signs, as they both tend to increase local public expenditures. The positive sign of the coefficient for the special wards in Tokyo (b_2) are unexpected, since they have less administrative responsibilities than the ordinary cities. However, the estimate is not significant at the standard levels of significance.

---

choice process where some public decision makers (e.g., voters) decide the desired level of output, which is exogenous to public officials who minimize costs subject to their technical constraint.

Table 12.4 Marginal effects of the environmental factors

Effects	Mean	Std. Error	Minimum	Maximum
$\partial \ln c_i/\partial a_{1i}$	0.080	0.074	−0.206	0.237
$\partial \ln c_i/\partial a_{2i}$	0.036	0.051	−0.160	0.144
$\partial \ln c_i/\partial a_{3i}$	0.056	0.004	0.047	0.072
$\partial \ln c_i/\partial a_{4i}$	−0.078	0.213	−0.532	0.749
$\partial \ln c_i/\partial a_{5i}$	−0.100	0.231	−0.592	0.798
$\partial \ln c_i/\partial a_{6i}$	0.067	0.075	−0.223	0.226
$\partial \ln c_i/\partial a_{7i}$	0.182	0.053	−0.023	0.295
$\partial \ln c_i/\partial a_{8i}$	0.114	0.077	−0.050	0.413
$\partial \ln c_i/\partial a_{9i}$	1.322	1.610	−2.111	7.581
$\partial \ln c_i/\partial a_{10i}$	2.096	1.369	−0.823	7.418
$\partial \ln c_i/\partial a_{11i}$	−0.046	0.061	−0.176	0.193
$\partial \ln c_i/\partial a_{12i}$	0.453	0.517	−0.650	2.465
$\partial \ln c_i/\partial a_{13i}$	0.493	0.391	−0.341	2.013

Table 12.5 Summary statistics of scale index values for 572 cities

	Mean	Std. Error.	Minimum	Maximum
$S(n)$	0.855	0.184	0.416	2.816
η_n	−0.145	0.184	−0.584	1.816

Table 12.5 shows the summary statistics of the scale index $S(n) \equiv \eta_n^g/\eta_c^g$ calculated for the 572 cities. Since the value for η_c^g is fixed at $1/\beta_w = 3.049$, variations in $S(n)$ across the cities are due to those in η_n^g, which in turn depend on local population as well as other environmental factors. The *rank* correlation coefficient between $S(n)$ and n is 0.921, which appears to be rather high despite the inclusion of the environmental factors in the formula for η_n^g. This may be because some of the environmental factors like the DID ratio and per capita income are highly correlated with population.

Figure 12.1 then shows the cumulative distribution of the scale index. Since population n^*, such that $S(n) = 1$ is MES, about 81% of the 572 cities is regarded as underpopulated whereas only 19% is regarded as overpopulated. To give a preview of the results, Table 12.6 lists the top 20 cities ranked according to the scale index along with their population ranks. In addition, Table 12.7 shows the cities in the vicinity of the optimal size with $S(n)$ such that $0.990 < S(n) > 1.001$.

12.5 Concluding Remarks

In this paper, we analytically delineated the factors that account for the "U-shaped" per capita local public expenditures. The important concepts we examined are (i)

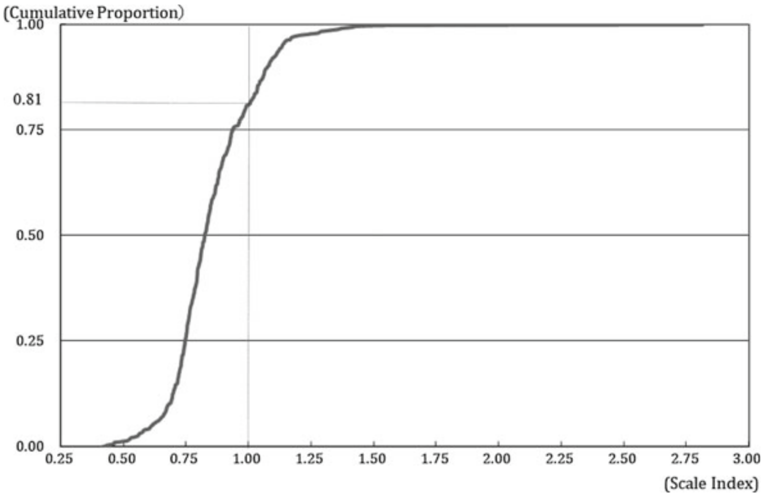


Fig. 12.1 Distribution of cities according to the scale index

Table 12.6 Top 20 over-congested cities

Rank	*Name*	*S(n)*	*Pop. Rank*	*Population*
1	Chiyoda (Tokyo)	2.816	504	34.8
2	Chuo (Tokyo)	1.570	329	63.9
3	Yokohama	1.427	1	3,307.1
4	Osaka	1.418	2	2,602.4
5	Kyoto	1.384	5	1,463.8
6	Nagoya	1.372	3	2,152.2
7	Kobe	1.360	6	1,423.8
8	Sapporo	1.334	4	1,757.0
9	Hiroshima	1.308	9	1,108.9
10	Fukuoka	1.291	7	1,284.8
11	Minato (Tokyo)	1.284	154	144.9
12	Kitakyusyu	1.279	10	1,019.6
13	Sendai	1.245	11	971.3
14	Kawasaki	1.242	8	1,202.8
15	Kumamoto	1.207	15	650.3
16	Okayama	1.192	19	615.8
17	Hamamatsu	1.178	22	561.6
18	Kagoshima	1.173	23	546.3
19	Nagasaki	1.170	39	438.6
20	Chiba	1.168	12	856.9

Table 12.7 Minimally efficient cities

Rank	*Name*	*S(n)*	*Pop. Rank*	*Population*
104	Nagaoka	1.009	118	190.5
105	Suita	1.009	59	342.8
106	Takarazuka	1.009	110	202.5
107	Hitachi	1.008	115	199.2
108	Odawara	1.004	113	200.1
109	Kashiwa	1.001	70	317.8
110	Nakano (Tokyo)	0.994	72	306.6
111	Meguro (Tokyo)	0.991	96	243.1
112	Kishiwada	0.991	117	194.8
113	Ishihara	0.990	85	277.1

technical elasticity of scale, which characterizes technology for the direct outputs produced by a government, and (ii) congestion elasticity, which characterizes consumption technology for the public service level consumed by citizens. By articulating that population-induced changes in per capita cost are related to the relative magnitude between the two elasticities, we constructed an indicator that quantifies the distance of a local population from its MES for local public expenditures. We then estimated the urban public production structure in Japan and applied the analysis to the Japanese case. With the estimates obtained, we ranked the Japanese cities according to the calculated values of the indicator.

Let us conclude our argument with a cautionary note. In popular writings, the MES are often used as the target population for municipal mergers to attain. Recall however that we calculate the MES of local population, holding all other variables in the cost function constant. As briefly mentioned in Introduction, such all "other" variables also change in typical cases of municipal mergers. Obviously, a municipal merger alters not only the population of a given fiscal unit, but also its surface area and demographics. In addition, if the wage systems of municipal employees were different before merging, they are typically adjusted in a new fiscal unit. Furthermore, mergers also change political landscape which would result in a different level and spatial distribution of public service. These are the cases of merging local fiscal units with a fixed level of population in each unit. In contrast, the focus of this study was the case of population changes in a fixed jurisdiction of local fiscal unit. The MES only indicates the cost-minimizing population for a fixed territorial boundary of a municipalities.

References

Bergstrom, T.C., and R.P. Goodman. 1973. Private demands for public goods. *American Economic Review* 63: 280–296.

Blom-Hansen, J., K. Houlberg, S. Serritzlew, and D. Treisman. 2016. Jurisdiction size and local government policy expenditure: Assessing the effect of municipal amalgamation. *American Political Science Review* 110 (4): 812–831.

Borcherding, T.E., and R.T. Deacon. 1972. The demand for the services of non-federal government. *American Economic Review* 62: 891–901.

Bradford, D.F., R.A. Malt, and W.E. Oates. 1969. The rising cost of local public services: Some evidence and reflections. *National Tax Journal* 22: 185–202.

Brueckner, J.K. 1981. Congested public goods: The case of fire protection. *Journal of Public Economics* 15: 45–58.

Davidson, R., and J.G. MacKinnon. 1993. *Estimation and inference in econometrics*. New York: Oxford University Press.

Duncombe, W., and J. Yinger. 1993. An analysis of returns to scale in public production, with an application to fire protection. *Journal of Public Economics* 52: 49–72.

Edwards, J.H.Y. 1990. Congestion function specification and the "publicness" of local public goods. *Journal of Urban Economics* 27: 80–96.

Hayashi, M. 2002a. Congestion, technical returns, and the minimum efficient scales of local public expenditures: The case of Japanese municipalities. Discussion paper series 01-01, institute for research in business and economics, Meiji Gakuin university.

Hayashi, M. 2002b. Chiho jichitai no saisho koritsu kibo (The minimum efficient scales of municipalities). *Financial Review* 61: 59–89. [in Japanese].

Hayes, K. 1986. Local public good demand and demographic effects. *Applied Economics* 18: 1039–1045.

Hayes, K.J., and D.J. Slottje. 1987. Measure of publicness based on demographic scaling. *Review of Economics and Statistics* 69: 713–718.

Hirsch, W.Z. 1959. Expenditure implications of metropolitan growth and consolidation. *Review of Economics and Statistics* 41: 232–241.

Hirsch, W.Z. 1965. Cost functions of an urban government service: Refuse collection. *Review of Economics and Statistics* 47: 87–93.

Kitchen, H. 1976. A statistical estimation of an operating cost function for municipal refuse collection. *Public Finance Quarterly* 4: 56–76.

Ministry of Internal Affairs and Communication, 1997. Shichoson Kessan Jokyo Shirabe 1995 (Local Public Expenditures of Municipalities 1997).

Nihon Keizai Shinbunsya, 1998. 610 shi no gyosei sahbisu suijun ichi-ran (Table of public service levels of 610 cities). Nikkei Chi-iki Joho, 0(303).

Reiter, M., and A.J. Weichenrieder. 1999. Public goods, club goods, and the measurement of crowding. *Journal of Urban Economics* 46: 69–79.

Statistics Bureau, Management and Coordination Agency, 1997. Kokusei Chosa 1995 (National Census 1995).

Stevens, B.J. 1978. Scale, market structure, and the cost of refuse collection. *Review of Economics and Statistics* 60: 438–448.

Chapter 13
Fiscal Stabilization Under Government Spending Reversal

Kazuki Hiraga

13.1 Introduction

Many advanced countries face on the fiscal consolidation to set some fiscal rule. For example, countries which join the Euro currency followed the Maastricht Treaty.[1] And Japan set some commitments of fiscal discipline; e.g. "Basic Policy 2006" mentioned about the primary balance would be positive until 2011.

This paper investigates whether the fiscal stabilization rules contribute to its object in literature of Woodford (1998)'s "Locally" Ricardian, which means the condition of converging to some finite steady state value of debt.[2] Corsetti et al. (2012) analyzes the effect of fiscal policy under government spending reversals, which mean that the government decreases their spending when increases the government debt. Intuitively, government spending reversals consider about one of the instruments of fiscal stabilization rule. In this paper, however, we show that improvement of fiscal stabilization rule such as government spending reversals may worsen its object; i.e. attaining debt stability. This result occurs when the persistency of the government spending and interest rate are sufficient high. Concretely, higher degree of the spending reversal may occur cyclical behavior of the economy and more fluctuated. In addition, this paper expands the model of the Fiscal Theory of Price Level which is discussed by Leeper (1991), Sims (1994) and Woodford (1994, 1998, 2001) and obtains the explicit condition of stability.

This paper is organized as follows. Section 13.2 states the basic structure of the model and analyzes the stability of the model. Section 13.3 enriches the model and discusses the interpretation of the result, and we state the conclusion in Sect. 13.4.

[1] GIIPS crisis actually ruined the Maastricht Treaty.
[2] Bohn (1998, 2008) mention about the same concept of the sustainability of the government debt. Hiraga (2013) discusses about the condition of satisfying the "Globally" Ricardian.

K. Hiraga (✉)
Tokai University, Tokyo, Japan

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_13

13.2 The Model

13.2.1 Simple Model of the Government Budget Constraint

To grasp simply the structure of the government debt stability, we set the model of the government budget constraint and two fiscal stabilization rules using in Corsetti et al. (2012): the one is the tax reaction function with respect to the government debt; the other is the government spending reaction function as follows:

$$\hat{b}_{t+1} = (1 + r)\,\hat{b}_t + \hat{g}_t - \hat{t}_t \tag{13.1}$$

$$\hat{g}_{t+1} = -\psi \hat{b}_{t+1} + \rho \hat{g}_t \tag{13.2}$$

$$\hat{t}_t = \varphi \hat{b}_t \tag{13.3}$$

Where $\hat{b}_t$ is the (linearized) government debt at the beginning at period t, $\hat{g}_t$ is the government spending at the beginning at the period t and $\hat{t}$ is the tax revenue. Equation (13.1) is the (linearized) government, Eq. (13.2) is the spending reversal rule a la Corsetti et al. (2012) and Eq. (13.3) is tax revenue rule.[3]

Substituting Eq. (13.3) to (13.1), we rewrite the two-dimensional system:

$$\begin{bmatrix} 1 & 0 \\ \psi & 1 \end{bmatrix} \begin{bmatrix} \hat{b}_{t+1} \\ \hat{g}_{t+1} \end{bmatrix} = \begin{bmatrix} 1 + r - \varphi & 1 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \hat{b}_t \\ \hat{g}_t \end{bmatrix}$$
$$\Leftrightarrow A \begin{bmatrix} \hat{b}_{t+1} \\ \hat{g}_{t+1} \end{bmatrix} = B \begin{bmatrix} \hat{b}_t \\ \hat{g}_t \end{bmatrix}$$

where $A = \begin{bmatrix} 1 & 0 \\ \psi & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 + r - \varphi & 1 \\ 0 & \rho \end{bmatrix}$. Using inverse matrix of A, we obtain the following equation:

$$\begin{bmatrix} \hat{b}_{t+1} \\ \hat{g}_{t+1} \end{bmatrix} = A^{-1} B \begin{bmatrix} \hat{b}_t \\ \hat{g}_t \end{bmatrix} = \begin{bmatrix} 1 + r - \varphi & 1 \\ -\psi\,(1 + r - \varphi) & \rho - \psi \end{bmatrix} \begin{bmatrix} \hat{b}_t \\ \hat{g}_t \end{bmatrix}. \tag{13.4}$$

13.2.2 Stability Analysis of the Model

To inspect the stability of the model, we induce the eigenvalues as follows:

[3] We obtain the similar fiscal stabilization rule to Bohn (1998, 2008) combining with Eqs. (13.2) and (13.3).

$$\begin{vmatrix} 1+r-\varphi-\lambda & 1 \\ -\psi\,(1+r-\varphi) & \rho-\psi-\lambda \end{vmatrix} = (1+r-\varphi-\lambda)\,(\rho-\psi-\lambda)+\psi\,(1+r-\varphi)=0$$
$$\Rightarrow \lambda^2-(1+r+\rho-\varphi-\psi)\,\lambda+(1+r-\varphi)\,\rho=0$$

$$\lambda=\frac{1+r+\rho-\varphi-\psi\pm\sqrt{(1+r+\rho-\varphi-\psi)^2-4\,(1+r-\varphi)\,\rho}}{2} \quad (13.5)$$

Since $\hat{b}_t$ and $\hat{g}_t$ are state variables, there are two eigenvalues and we can obtain the following proposition which shows the sufficient condition of the stability.[4]

Proposition 13.1 *The system of government debt and expenditure must be stable if* $1+r+\rho-\varphi-\psi+\sqrt{(1+r+\rho-\varphi-\psi)^2-4\,(1+r-\varphi)\,\rho}<2$.

As for the necessary condition of the stability of the system, we can show it using trace and determinant of the system.

$$det\;(\mathrm{C})=\lambda_1\lambda_2=(1+r-\varphi)\,(\rho-\psi)+\psi\,(1+r-\varphi)=(1+r-\varphi)\,\rho \quad (13.6)$$

$$tr\;(\mathrm{C})=\lambda_1+\lambda_2=1+r+\rho-\varphi-\psi \quad (13.7)$$

where $C=\begin{bmatrix} 1+r-\varphi & 1 \\ -\psi(1+r-\varphi) & \rho-\psi \end{bmatrix}$ We can obtain the necessary condition the following proposition.

Proposition 13.2 *The system of government debt and expenditure must be unstable if* det (C) >1 *and* tr (C) >2. *In other word, the necessary condition of the stability is* det (C) <1 *and* tr (C) <2.

To focus on the role of the spending reversal parameter ψ, we set r $=0.05$, $\rho=0.5$, $\varphi=0.01$ and show the eigenvalues in each ψ in Fig. 13.1. Figure 13.1 shows that stranger the degree of spending reversal ψ helps the stability of the system but much stronger ψ may cause the cyclical behavior of the system.[5]

13.2.3 Phase Diagram Explanation

To grasp visually the mechanism of the system, we also analyze the model using phase diagram.

[4]As you can see, the eigenvalues are real solutions iff $(1+r+\rho-\varphi-\psi)^2>4\,(1+r-\varphi)\,\rho\Rightarrow\psi<1+r+\rho-\varphi+2\sqrt{(1+r-\varphi)\,\rho}$. Otherwise, they are imaginary solutions.

[5]Because the eigenvalues are imaginary solutions.

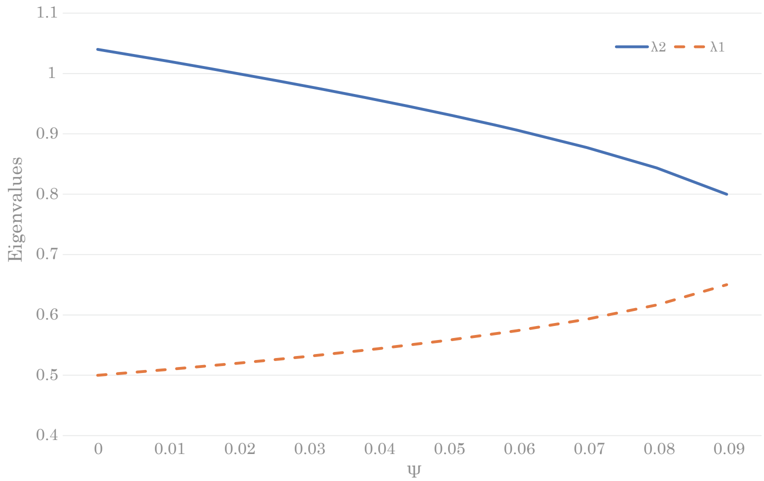


Fig. 13.1 Eigenvalues ($\lambda 2$ is larger than $\lambda 1$) in each spending reversal parameter ψ (We define that $\lambda 2 = \frac{1+r+\rho-\varphi-\psi+\sqrt{(1+r+\rho-\varphi-\psi)^2-4(1+r-\varphi)\rho}}{2}$, and $\lambda 1 = \frac{1+r+\rho-\varphi-\psi-\sqrt{(1+r+\rho-\varphi-\psi)^2-4(1+r-\varphi)\rho}}{2}$). (Note) The eigenvalues are imaginary solutions when $\psi > 0.1$

Substituting Eq. (13.3) to (13.1), we obtain the condition of $\Delta\hat{b}_{t+1} = \hat{b}_{t+1} - \hat{b}_t \geq 0$ as follow:

$$\Delta\hat{b}_{t+1} \geq 0 \;\Rightarrow (r-\varphi)\,\hat{b}_t + \hat{g}_t \geq 0 \Rightarrow \begin{cases} \hat{b}_t \geq -\frac{\hat{g}_t}{r-\varphi}, & if\ r \geqslant \varphi \\ \hat{b}_t \leq -\frac{\hat{g}_t}{r-\varphi}, & if\ r < \varphi \end{cases} \tag{13.7}$$

As for the condition of $\Delta\hat{g}_{t+1} = \hat{g}_{t+1} - \hat{g}_t \geq 0$, we can rewrite as follow:

$$\Delta\hat{g}_{t+1} = -\psi\,(1+r-\varphi)\,\hat{b}_t - (1-\rho+\psi)\,\hat{g}_t \geq 0 \Rightarrow \hat{b}_t \leq -\frac{1-\rho+\psi}{\psi\,(1+r-\varphi)}\hat{g}_t \tag{13.8}$$

Using Eqs. (13.7) and (13.8), we obtain the following proposition and show in Figs. 13.2, 13.3 and 13.4.

Proposition 13.3 *1. The system is sink if $r < \varphi$.*
2. The system is saddle if $r \geq \varphi$ and $\psi < (1-\rho)\,(r-\varphi)$
3. The system is source if $r \geq \varphi$ and[6] *$\psi > (1-\rho)\,(r-\varphi)$.*

Figure 13.2 shows the system is stable in any points. Figure 13.3 shows the system is only stable at the saddle-path (i.e. Knife-Edge stable), Fig. 13.4 shows the system is source; i.e. unstable. Interesting result is that larger ψ may disturb stability of the

[6]It may cause the limit cycle.

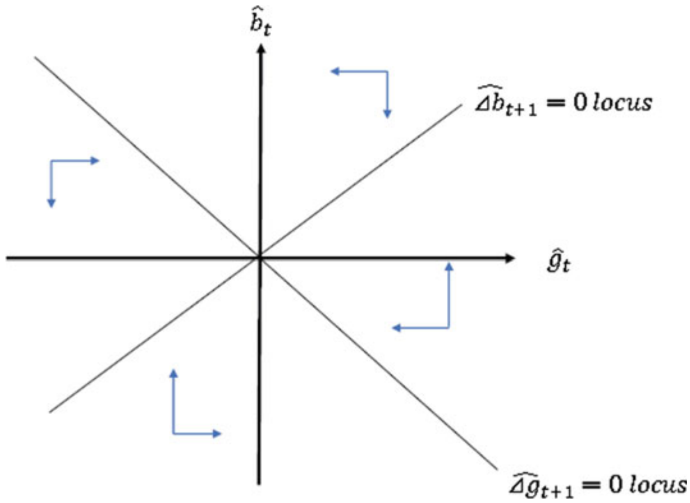


Fig. 13.2 The case in the system is sink ($r < \varphi$)

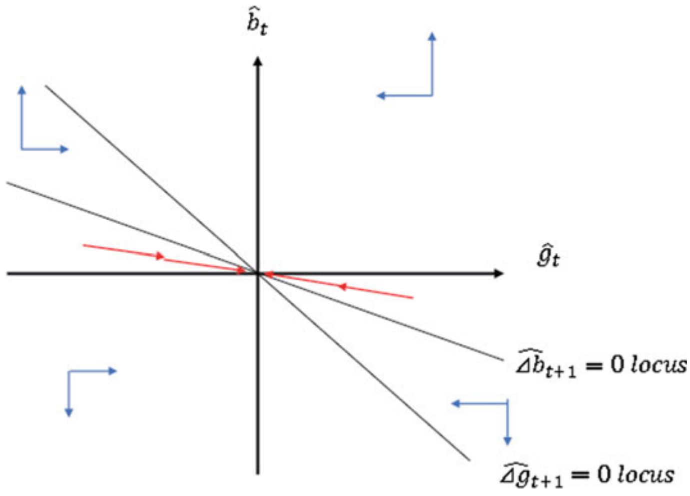


Fig. 13.3 The case in the system is saddle ($r \geq \varphi$ and $\psi \leq (1 - \rho)(\rho - \varphi)$)

system shown in Proposition 13.3(3). Especially, larger persistency of the government spending is easier to cause the case in source.

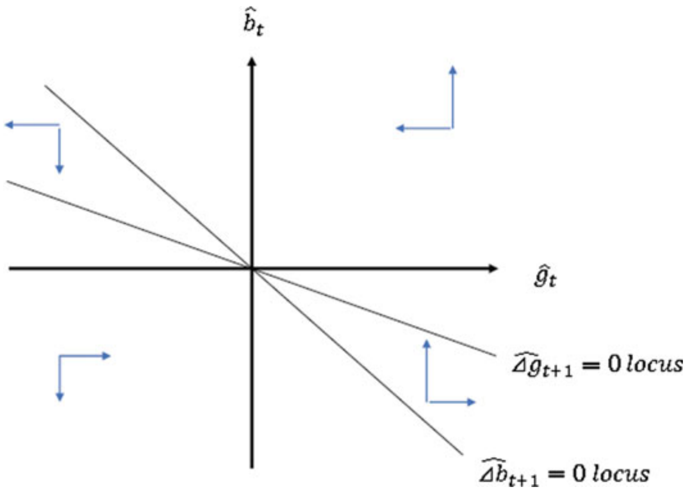


Fig. 13.4 The case in the system is saddle ($r \geq \varphi$ and $\psi \leq (1-\rho)(\rho-\varphi)$)

13.3 Extension to the Fiscal Theory of Price Level

In this section, we extend the model which connects with the Fiscal Theory of Price Level (FTPL) discussed by Leeper (1991), Sims (1994) and Woodford (1994, 1998, 2001).

13.3.1 The Model of FTPL with Spending Reversal

We construct the model of Leeper (1991) with the fiscal spending rule a la Corsetti et al. (2012). For simplicity, we assume that the economy is perfect foresight and representative infinite horizon model without capital.

13.3.2 Representative Agent

Representative agent chooses sequences real consumption c_t, nominal money stock M_t and nominal bond B_t to solve the following lifetime utility,

$$\max \sum_{t=0}^{\infty} \beta^t \left[log c_t + \delta log \frac{M_t}{P_t} \right]$$

Subject to the following budget constraint;

$$c_t + \frac{M_{t+1} + B_{t+1}}{P_t} + \tau_t = y + \frac{M_t + R_t B_t}{P_t}$$

Where $\beta(<1)$ is the discount factor, δ is the preference for real money stock, y is the exogenous fixed income, P_t is the price level at period t, τ_t is the lump-sum tax (or transfer if negative) and R_{t-1} is the nominal gross interest rate at the end of period $t-1$ (equals to the beginning at period t).

Aggregate resource constraint for this economy is as follow:

$$c_t + g_t = y.$$

Solving the first order conditions, we can obtain the following equations;

$$\frac{c_{t+1}}{c_t} = \beta \frac{R_{t+1}}{\pi_{t+1}}$$

$$m_t = \frac{\delta c_t}{R_t - 1}$$

where $\pi_{t+1} = \frac{P_{t+1}}{P_t} m_t = \frac{M_t}{P_t}$.

13.3.3 Unionized Government Budget and Policies

The government budget constraint which combines fiscal and monetary authorities as follow:

$$\pi_{t+1}\left(b_{t+1} + m_{t+1}\right) + \tau_t = g_t + R_t b_t + m_t$$

Fiscal policies obey Eqs. (13.2) and (13.3). In addition, monetary authority obeys the following equation:

$$\hat{R}_t = \alpha \hat{\pi}_t. \tag{13.9}$$

13.3.4 Linearized Equilibrium Conditions

To combine the equilibrium conditions and linearized equilibrium conditions and government budget constraint, we obtain the following equations:

$$\hat{g}_t - \hat{g}_{t+1} = (\alpha - 1)\,\hat{\pi}_{t+1} \tag{13.10}$$

$$\hat{m}_t = -\hat{g}_t - \alpha \hat{\pi}_t \tag{13.11}$$

$$\pi\left(\hat{b}_{t+1}+\hat{m}_{t+1}\right)+(b+m)\,\hat{\pi}_{t+1}+\varphi\hat{b}_t=\hat{g}_t+R\hat{b}_t+b\hat{R}_{t+1}+b\hat{\pi}_t+\hat{m}_t$$
$$\Rightarrow (b+m+\alpha(b-1))\,\hat{\pi}_{t+1}-\alpha\pi\hat{g}_{t+1}+\pi\hat{b}_{t+1}=(b-\alpha)\,\hat{\pi}_t+(R-\varphi)\hat{b}_t \tag{13.12}$$

13.3.5 Stability Analysis

In this system of FTPL with spending reversal, there are three equations in two state variables (government debt $\hat{b}$ and government spending $\hat{g}$) and one control variable (inflation rate $\hat{\pi}$).

$$\begin{bmatrix}\omega & -\alpha\pi & \pi\\ 0 & 1 & \psi\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\hat{\pi}_{t+1}\\ \hat{g}_{t+1}\\ \hat{b}_{t+1}\end{bmatrix}=\begin{bmatrix}b-a & 0 & R-\varphi\\ 0 & \rho & 0\\ 0 & 1 & R-\varphi\end{bmatrix}\begin{bmatrix}\hat{\pi}_t\\ \hat{g}_t\\ \hat{b}_t\end{bmatrix}$$
$$\Rightarrow F\begin{bmatrix}\hat{\pi}_{t+1}\\ \hat{g}_{t+1}\\ \hat{b}_{t+1}\end{bmatrix}=G\begin{bmatrix}\hat{\pi}_t\\ \hat{g}_t\\ \hat{b}_t\end{bmatrix}, \tag{13.13}$$

where $\omega=b+m+\alpha\,(b-1)$ $\quad F=\begin{bmatrix}\omega & -\alpha\pi & \pi\\ 0 & 1 & \psi\\ 0 & 0 & 1\end{bmatrix}$ $\quad \mathrm{G}=\begin{bmatrix}b-a & 0 & R-\varphi\\ 0 & \rho & 0\\ 0 & 1 & R-\varphi\end{bmatrix}$
and b, m and π are the steady state values which are exogenous by assumption.

Using inverse matrix of F, we obtain the following equation:

$$\begin{bmatrix}\hat{\pi}_{t+1}\\ \hat{g}_{t+1}\\ \hat{b}_{t+1}\end{bmatrix}=F^{-1}G\begin{bmatrix}\hat{\pi}_t\\ \hat{g}_t\\ \hat{b}_t\end{bmatrix}$$
$$=\begin{bmatrix}\frac{b-a}{\omega} & \frac{(1+a\psi)\pi}{\omega} & \frac{(R-\varphi)[1-(1+a\psi)\pi]}{\omega}\\ 0 & \rho & -\psi\,(R-\varphi)\\ 0 & 0 & R-\varphi\end{bmatrix}\begin{bmatrix}\hat{\pi}_t\\ \hat{g}_t\\ \hat{b}_t\end{bmatrix}. \tag{13.14}$$

To inspect the stability of the model, we induce the eigenvalues as follows:

$$\begin{vmatrix}\frac{b-a}{\omega}-\lambda & \frac{(1+a\psi)\pi}{\omega} & \frac{(R-\varphi)[1-(1+a\psi)\pi]}{\omega}\\ 0 & \rho-\lambda & -\psi\,(R-\varphi)\\ 0 & 1 & R-\varphi-\lambda\end{vmatrix}=0$$
$$\Rightarrow\left(\frac{b-a}{\omega}-\lambda\right)[(\rho-\lambda)\,(R-\varphi-\lambda)+\psi\,(R-\varphi)]=0 \tag{13.15}$$

Therefore,

$$\lambda = \frac{b-a}{\omega} \frac{R+\rho-\varphi \pm \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2}$$

Using the explicit solution of eigenvalues, we can obtain the following proposition.

Proposition 13.4 *(1) The system faces on indeterminacy, when*

$$\frac{b-a}{\omega} < 1, \ \frac{R+\rho-\varphi - \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2} < 1,$$
$$\frac{R+\rho-\varphi + \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2} < 1.$$

(2) The system is no bounded equilibrium, when

$$\frac{b-a}{\omega} > 1, \ \frac{R+\rho-\varphi - \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2} > 1$$

$$\frac{R+\rho-\varphi + \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2} > 1.$$

(3) The system has unique equilibrium when

(i)

$$\frac{b-a}{\omega} > 1, \ \frac{R+\rho-\varphi - \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2}$$
$$< 1, \frac{R+\rho-\varphi + \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2} < 1$$

or

(ii)

$$\frac{b-a}{\omega} < 1, \ fracR+\rho-\varphi - \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}2 < 1$$

$$\frac{R+\rho-\varphi + \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2} > 1$$

The case in (i) is regime in Ricardian and case in (ii) is regime in FTPL (Non-Ricardian).

Similar to the previous section, To focus on the role of the spending reversal parameter ψ, we set r = 0.05, æ = 0.5, ' = 0.01 and show the two eigenvalues; i.e. case in $\lambda = \frac{R+\rho-\varphi \pm \sqrt{(R+\rho-\varphi)^2 - 4(\rho+\psi)(R-\varphi)}}{2}$ in each ψ in Fig. 13.5. Similar to Figs. 13.1, 13.5 shows that stranger the degree of spending reversal ψ helps the

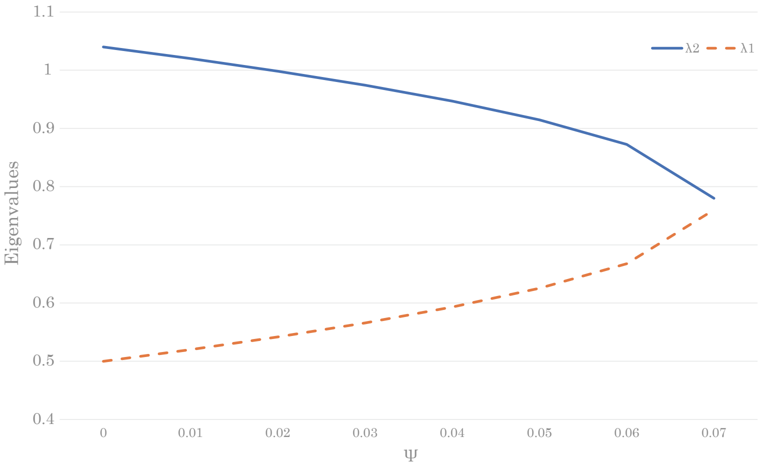


Fig. 13.5 Eigenvalues(λ_2 is larger than λ_1) in each spending reversal parameter ψ (We define that $\lambda_2 = \frac{R+\rho-\varphi+\sqrt{(R+\rho-\varphi)^2-4(\rho+\psi)(R-\varphi)}}{2}$, and $\lambda_1 = \frac{R+\rho-\varphi-\sqrt{(R+\rho-\varphi)^2-4(\rho+\psi)(R-\varphi)}}{2}$). (Note) The eigenvalues are imaginary solutions when $\psi > 0.08$

stability of the system, but much stronger ψ may cause the cyclical behavior of the system.[7]

13.4 Conclusion

This paper investigates whether the fiscal stabilization rules contribute to its object in literature of Woodford (1998)'s "Locally" Ricardian, which means that the condition of converging to some finite steady state value of debt.

This paper shows that the stronger fiscal stance may worsen its possibility of target level of the debt. This result implies that some countries and local government fail to fiscal reconstruction because their higher interest rate of debt and persistency of the government spending.

There are some possible improvement points. First, we inspect the model with sovereign default. Second, this paper compares with the welfare criteria. We will deal with them as future researches.

[7]Because the eigenvalues are imaginary solutions.

References

Bohn, H. 1998. The behavior of U.S. public debt and deficits. *Quarterly Journal of Economics* 113: 949–963.

Bohn, H. 2008. The sustainability of fiscal policy in the United States. In *Sustainability of public debt*, vol. 2008, ed. Reinhard, Neck and Jan-Egbert Sturm, 15–49. Cambridge: MIT Press.

Corsetti, G., A. Meier, and G.J. Müller. 2012. Fiscal stimulus with spending reversals. *Review of Economics and Statistics* 94: 878–895.

Hiraga, K. 2013. A note on the weak condition of "Globally" Ricardian, *Economics bulletin* 33: 361–371.

Leeper, E. 1991. Equilibria under 'Active' and 'Passive' monetary policies. *Journal of Monetary Economics* 27: 129–147.

Sims, C. 1994. A simple model for the study of the determination of the price level and the interaction of monetary and fiscal policy. *Economic Theory* 4: 381–399.

Woodford, M. 1994. Monetary policy and price level determinacy in a cash in advance economy. *Economic Theory* 4: 345–380.

Woodford, M. 1998. *Public debt and the price level*. New Jersey: Princeton University.

Woodford, M. 2001. Fiscal requirements for price stability. *Journal of Money, Credit and Banking* 33: 669–728.

Chapter 14
Welfare Effects of Public Education and National Security

Tatsuya Omori

14.1 Introduction

In the 21st century, nations such as Japan must confront not only an aging society, but also national disasters (e.g., earthquake, typhoon, and flood) and international disputes (e.g., war and terrorism). In an aging society, the population of older generation increases but the number of children and the population of working generation decreases. Productivity efficiency decreases and economic growth decline even if productivity is improved. To ameliorate those effects, economic agents promote the accumulation of human capital. Additionally, the enrichment of education makes agents secure with life risk because education enhances the agents' abilities and their income. Nevertheless, the nation must also confront national risks and controversial issues related to national security policy. For the agents, national risk affects consumption and saving behaviour because the environment around agents changes.

On the other hand, Musgrave (1958) specifies that government plays roles of stabilising the economy with economic growth and of providing public goods and/or public services to agents when private agents have difficulty securing national risk. Therefore, government is expected to resolve difficulties of how to spend public funds to pursue public policies.

A pioneering work by Lucas (1988) describes development of a neoclassical growth model with emphasis on human capital accumulation through formal education and learning-by-doing. Glomm and Ravikumar (1992) present an overlapping-generations model with human capital investment by public education and private education. de la Croix and Doepke (2004) show that, although private schooling engenders higher growth, public education can be predominant when inequality is sufficiently high. Introducing a fertility decision and child care costs into an overlapping-generations model with public education and social security, Omori

T. Omori (✉)
Chukyo University, Nagoya, Japan
e-mail: tomori@mecl.chukyo-u.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_14

(2009) examines the effects of these public policies on fertility and demonstrates that, with a constant tax rate, a change in the allocation from social security benefits to public investment in education decreases fertility. Dioikitopoulos (2014) develops an overlapping-generations model with public education and public health and demonstrates how government can use the allocation of both public expenditures to maximise growth. Using an overlapping-generations model with public education and public health expenditure, Omori (2018) discusses how public health expenditure and public education affect human capital formation and social welfare in a small open economy and shows that, with a constant wage tax rate, changing the allocation rate from public health improvements to public investment in education increases the growth rate of human capital and private expenditure for education. However, it is difficult to find a model of human capital accumulation that includes discussion of national security.

For discussion of national security, we applied the economic theory of public goods and/or public services to these issues. In the context of a discussion of national security, Smith (1995) provides an analytical survey of military expenditures using an empirical econometric model. Murphy and Topel (2013) show that national security has three salient characteristics. First, national security investment provides societal insurance against widespread harm. Second, national security is useful in various circumstances such as military preparedness. Third, potential national security threats are uncertain in terms of their occurrence and magnitude. Naito and Omori (2016) use a simple overlapping-generations model economy to discuss effects of natural disaster prevention behaviour of households on population distribution in certain areas, and on fertility. Some room remains to elucidate the relation between public education and national security.

As discussed earlier, an important issue is the allocation of public funds between public education and national security as fundamental functions of government. Regarding the allocation of public funds between social security and public education, Kaganovich and Zilcha (1999) discuss government allocation between public investment in education and social security benefits to the older generation. However, in the literature that addresses public education and national security as government functions, finding models that include both public education and national security is difficult. As described herein, we discuss public policy that is inclusive of public education and national security in the economy. The question addressed in this paper is the following: How do public education and national security affect social welfare?

The remainder of the paper is organised as follows: Sect. 14.2 presents the model. Section 14.3 presents balanced growth path characteristics. Based on Sects. 14.3, 14.4 clarifies the effects of public education and national security on welfare. The last section presents concluding remarks.

14.2 Model

Based on Diamond (1965), for the simplification of analysis, we consider an overlapping-generations model of a small open economy.[1] In this economy, the world interest rate is assumed to remain constant over time.[2] The economy comprises identical three-period-lived consumers, perfectly competitive firms, and a government.

14.2.1 Consumers

Consumers in the first period of their lives, as the young generation, are raised by their parents. They combine inputs provided by their parents and government to develop their human capital.

Consumers in the second period of their lives, as the middle-aged generation, supply their labour to firms inelastically. They have children and divide their one unit of time between raising their children and working. They divide their after-tax income among current consumption, saving for consumption when older, and private investment in education.

Consumers in the final period of life, as the older generation, enjoy consumption with their accumulated savings.

As described in this paper, because we discuss national security for national risk, we introduce longevity as life risk into the model for the comparison with national risk. With probability p_s, a consumer who worked during period t will live throughout the final period. With probability $1 - p_s$, the consumer will die before the onset of the final period. In this model, similar to those presented by Yakita (2001), Chakraborty (2004), Pestieau et al. (2008), Omori (2018), and others, when introducing longevity into the overlapping-generations model, we assume that the probability of survival, p_s, is the equal and given exogenously for all individuals. However, because accidental bequests arise if a consumer dies at the onset of old age, by introducing an annuity market into the model, we do not presume accidental bequests. The return in the annuity market at period t is the efficiency interest rate, $1 + r$, divided by $\left(\text{i.e., } \frac{1+r}{p_s}\right)$.

Following Naito and Omori (2016) who develops a simple overlapping-generations model with introducing the probability of national risk into the overlapping-generations model, we assume that the utility function includes damage caused by national risk, which is denoted by D_i in period i. Consumers deal with that damage as given and cannot control this variable because we assume that the

[1] Similar to Omori (2018), we develop the overlapping-generations model in a small open economy. We can develop a similar model in a closed economy. The path in a closed economy might not fundamentally differ from that in a small open economy.

[2] The capital–labour ratio and wage rate are also constant. Production technology is assumed to be governed by a standard neoclassical constant-returns-to-scale production function.

national risk is a "public bad" and that it generates negative external effects on the economy.[3]

Because the middle-aged generation at period t is called generation t, the preference of a representative consumer of generation t is

$$U = \ln c_t^t + p_s \ln c_{t+1}^t + \beta \ln n_{t+1} h_{t+1} - p_D \ln D_t - p_s p_D \ln D_{t+1}, \quad (14.1)$$

where c_t^t and c_{t+1}^t respectively denote the consumption of generation t during the second period and the older period, n_{t+1} is the number of children, h_{t+1} represents the human capital at $t+1$, and β is the positive parameter.[4] In addition, p_D is the probability of national risk.[5] We assume $0 < p_D < 1$. Let N_t be the total middle-aged population at period t. Therefore, we have $N_{t+1} = n_{t+1} N_t$.

Similarly to Glomm and Ravikumar (1992) and Omori (2018), we assume public investment in education for which all children are provided the same quality and which governments determine as compulsory education. We also assume that parents determine private investment in education based on the individual optimal behaviour, such that they consume their disposable income to buy picture books and other educational materials for children.

The human capital of each consumer in generation $t+1$ is a function of the following: private investment in education made by parents, e_t, as the material sacrifice; public investment in education made by the government, E_t; and the parents' child care time per child, Λ, as the opportunity cost.[6]

The human capital of each consumer in generation $t+1$ is given as

$$h_{t+1} = e_t^{\eta} E_t^{1-\eta} \Lambda^{\gamma}, \quad (14.2)$$

where $\gamma > 0$ and $0 < \eta < 1$.

The budget constraints of a representative consumer of generation t in middle-age and the old period are given, respectively, as

$$(1-\tau)\left[1 - n_{t+1}\Lambda\right] w h_t = c_t^t + e_t + s_t, \quad (14.3)$$

[3]Fennema and Van Assen (1998) and Slovic et al. (1977) show that the marginal disutility is decreasing with respect to damage by empirical analysis. Based on these results, Naito and Omori (2016) develop the model and we also assume that the marginal disutility of damage is decreasing in our model.

[4]As shown in Omori (2018), in this utility function, parents obtain utility from consumption and from education of their children. The value of this education is summarised by the child's human capital. This utility is derived from parents' love of or duty to their children.

[5]Assuming that the marginal disutility of damage is decreasing in our model, p_D in the utility function denotes the subjective provability of national risk rather than the occurrence probability of national risk.

[6]As explained in Kaganovich and Zilcha (1999) and Omori (2009, 2018), in raising children, communication between parents and their children is helpful for human capital accumulation of their children.

and

$$\frac{(1+r)\,s_t}{p_s} = c_{t+1}^t, \tag{14.4}$$

where τ denotes the wage income tax rate, w stands for the wage rate, s_t stands for savings, and r signifies the interest rate.

Given the wage rate, the interest rate, the human capital, the wage income tax rate, and the child care time per child, a representative consumer chooses c_t^t, c_{t+1}^t, n_{t+1}, and e_t to maximise his utility, (14.1), subject to the budget constraints, (14.3) and (14.4). We can present the first-order conditions as

$$\frac{1}{c_t^t} = \lambda, \tag{14.5}$$

$$\frac{1}{c_{t+1}^t} = \frac{\lambda}{1+r}, \tag{14.6}$$

$$\frac{\beta}{n_{t+1}} = \lambda\,(1-\tau)\,\Lambda w h_t, \tag{14.7}$$

$$\frac{\beta\eta}{e_t} = \lambda, \tag{14.8}$$

and

$$(1-\tau)\left[1-n_{t+1}\Lambda\right] w h_t - c_t^t - e_t - \frac{p_s c_{t+1}^t}{1+r} = 0, \tag{14.9}$$

where λ is a Lagrangian multiplier.

From these first-order conditions, we can show the optimal plans as

$$n_{t+1} = \frac{\beta}{(1+p_s+\beta+\beta\eta)\,\Lambda} \tag{14.10}$$

and

$$e_t = \frac{(1-\tau)\,w h_t \beta\eta}{(1+p_s+\beta+\beta\eta)}. \tag{14.11}$$

From (14.10), the number of children is shown by constant parameters. Hereinafter, we designate n_{t+1} as n^* $(n_{t+1} = n^*)$. That is,

$$n^* = \frac{\beta}{(1+p_s+\beta+\beta\eta)\,\Lambda} \tag{14.12}$$

In fact, not p_D but p_s affects the decision of fertility.

14.2.2 National Risk and National Security

We discuss how governments secure for national risk. For natural disasters, governments adopt policies to mitigate damage generated by disasters. For national defence, governments execute policies to deter enemies and mitigate damage to national interests caused by the enemy.[7] As described herein, we presume that national security mitigates damage generated by national risks which are posed by natural disasters, international disputes, and other eventualities.

In this subsection, we formulate the damage that might be caused by national risk in the model and consider national security as a means of mitigating national risk. Consumers face damage by national risk during second and final periods. Because D_t and D_{t+1} show national risk and because consumers deal with them as given, consumers cannot ascertain the level of this damage. However, as discussed in the Introduction, we assume that the government can mitigate the impact of the damage by public expenditure for national security. Therefore, we assume the damage function because of national security as

$$D_t = \frac{a}{G_t}, \tag{14.13}$$

and

$$D_{t+1} = \frac{a}{G_{t+1}}, \tag{14.14}$$

where a stands for the primary damage, G_t signifies public expenditure for national security at period t, and G_{t+1} represents public expenditure for national security in period $t+1$. Based on (14.13) and (14.14), the damage to national security is a decreasing function of public expenditure for national security.[8]

14.2.3 Government

During period t, the government imposes a wage income tax on consumers of the middle-aged generation and allocates tax revenue to public investment in education and national security benefits. The government's budget constraint in period t is

$$\tau\left(1 - n^* \Lambda\right) w h_t N_t = E_t N_t + G_t N_t + G_t p_s N_{t-1}. \tag{14.15}$$

The left-hand-side of (14.15) represents income tax revenues of period t. However, the first term on the right-hand-side in (14.15) represents public investment in educa-

[7]In many economically developed countries, military power presents a deterrent to other countries and lowers the probability of international disputes.

[8]We assume that the government expends resources for national risk as a government function.

tion at period t. The second and third terms on the right-hand-side in (14.15) denote public expenditure for national security at period t.

We assume that each consumer has n children in the working generation. Therefore, n is equal to N_t/N_{t-1}.

Let Δ represent the allocation rate of tax revenue for public investment in education, where $0 < \Delta < 1$. Therefore, one can show public investment in education and national security as

$$E_t = \Delta\tau\left(1 - n^*\Lambda\right)wh_t, \tag{14.16}$$

and

$$G_t = \frac{n}{n + p_s}(1 - \Delta)\tau\left(1 - n^*\Lambda\right)wh_t. \tag{14.17}$$

14.3 Balanced Growth

Based on (14.2), (14.10), (14.11), and (14.16), the growth rate of human capital, Γ, is derived as

$$\Gamma \equiv \frac{h_{t+1}}{h_t} = \left[\frac{\beta\eta}{(1 + p_s + \beta + \beta\eta)}\right]^{\eta}\left[\Delta\tau\left(\frac{1 + p_s + \beta\eta}{1 + p_s + \beta + \beta\eta}\right)\right]^{1-\eta} w\Lambda^{\gamma}. \tag{14.18}$$

From (14.18), the growth rate of human capital depends on not the probability of national risk but on the probability to survive. Although public expenditure for national security does not affect the economic growth rate, this results shows that government must expend that as government's functions in the economy.

On a balanced growth path, the growth rate of human capital, Γ is a constant and time-invariant variable.[9] Human capital grows at a rate of Γ, given positive human capital at initial period, h_0. Therefore, human capital at period t is shown as $h_t = \Gamma^t \times h_0$. In the balanced growth path, c_t^t, c_{t+1}^t, and e_t can be rewritten as

$$\hat{c}_t^t = \frac{(1 - \tau)w\Gamma^t h_0}{(1 + p_s + \beta + \beta\eta)}, \tag{14.19}$$

$$\hat{c}_{t+1}^t = \frac{(1 + r)(1 - \tau)w\Gamma^t h_0}{(1 + p_s + \beta + \beta\eta)}, \tag{14.20}$$

and

$$\hat{e}_t = \frac{\beta\eta(1 - \tau)w\Gamma^t h_0}{(1 + p_s + \beta + \beta\eta)}, \tag{14.21}$$

[9] This economy is regarded as a small open economy in which the capital–labour ratio, interest rate, and wage rate are constant.

where ˆ denotes the variables on the balanced growth path. We also rewrite E_t on (14.16) and G_t on (14.17) in the balanced growth path as

$$\hat{E}_t = \Delta\tau\left[1 - n^*\Lambda\right] w\Gamma^t h_0, \tag{14.22}$$

and

$$\hat{G}_t = \frac{n^*}{n^* + p_s}(1 - \Delta)\tau\left[1 - n^*\Lambda\right] w\Gamma^t h_0. \tag{14.23}$$

The national security at period $t + 1$ is also shown as

$$\hat{G}_{t+1} = \frac{n^*}{n^* + p_s}(1 - \Delta)\tau\left[1 - n^*\Lambda\right] w\Gamma^{t+1} h_0. \tag{14.24}$$

14.3.1 Policy Effects on Economic Growth

Based on (14.18), the effects of wage income tax on economic growth rate are shown as

$$\frac{d\Gamma}{d\tau} = (1 - \eta)\left[\frac{\beta\eta}{(1 + p_s + \beta + \beta\eta)}\right]^{-\eta}\left[\Delta\left(\frac{1 + p_s + \beta\eta}{1 + p_s + \beta + \beta\eta}\right)\right]^{\eta} w\Lambda^{\gamma} > 0. \tag{14.25}$$

A higher wage tax rate enhances the economic growth rate because increasing the tax rate raises public investment in education.

Similarly, the effects of allocation from national security to public investment in education on growth rate are shown as

$$\frac{d\Gamma}{d\Delta} = (1 - \eta)\left[\frac{\beta\eta}{(1 + p_s + \beta + \beta\eta)}\right]^{-\eta}\left[\tau\left(\frac{1 + p_s + \beta\eta}{1 + p_s + \beta + \beta\eta}\right)\right]^{\eta} w\Lambda^{\gamma} > 0. \tag{14.26}$$

Changing the allocation of public funds from national security to public investment in education enhances the economic growth rate because of the accumulation of human capital.

14.3.2 Cross-Policy Effects on Economic Growth

This model includes the assumption that government has two policy instruments: tax to finance public investment in education and national security as well as allocation of public funds to them. However, as government should adopt them, one policy change influences another policy. The relation between tax policy and allocation policy is interdependent, if the tax policy is changed, then the change affects the allocation policy and vice versa. There are cross-policy effects on economic growth that can be shown as

$$\frac{d\Gamma}{d\tau d\Delta} = \frac{d\Gamma}{d\Delta d\tau} = \eta(1-\eta)\left[\frac{\beta\eta}{(1+p_s+\beta+\beta\eta)}\right]^{-\eta}\left[\frac{1+p_s+\beta\eta}{1+p_s+\beta+\beta\eta}\right]^{\eta-1} w\Lambda^{\gamma} > 0. \tag{14.27}$$

Based on (14.25) and (14.26), because both tax and allocation effects on economic growth are positive, the cross effects are also positive.

14.4 Welfare Effects

Before discussing welfare effects, we assume a social welfare function in this small open economy. Substituting (14.19)–(14.24) into (14.1), we derive the indirect utility function in equilibrium as

$$W_t = \ln\hat{c}_t^t + p_s \ln\hat{c}_{t+1}^t + \beta\ln n^*\hat{h}_{t+1} - p_D\ln\hat{D}_t - p_s p_D\ln\hat{D}_{t+1}, \tag{14.28}$$

where $\hat{h}_{t+1} = \hat{e}_t^{\eta}\hat{E}_t^{1-\eta}\Lambda^{\gamma}$, $\hat{D}_t = \frac{a}{\hat{G}_t}$ and $\hat{D}_{t+1} = \frac{a}{\hat{G}_{t+1}}$.

14.4.1 Tax Effects

Differentiating (14.28) with respect to τ, the effect of increasing wage income tax rate on welfare is shown as

$$\frac{dW}{d\tau} = \frac{-\tau(1+p_s+\beta)+(1-\eta)\beta}{(1-\tau)} + \{p_s[t+p_D(t+1)]+t(1+\beta+p_D)\}\frac{1}{\Gamma}\frac{d\Gamma}{d\tau}. \tag{14.29}$$

Specifically, addressing the first term of right-hand-side in (14.29), if $\frac{(1-\eta)\beta}{(1+p_s+\beta)} > \tau$, then $\frac{dW}{d\tau}$ is positive. A higher tax rate has two effects. One is a positive benefit effects on welfare because increasing taxes increases tax revenues; also, government expends funds for public investment in education. The other is the negative tax effects on welfare because a higher tax decreases the disposable income for consumers and because consumers expend private investment in education less. When the former effects are greater than the latter effects, a higher tax enhances social welfare.

From (14.29), cross effects on welfare are shown as

$$\frac{dW}{d\tau d\Delta} = -\left(p_s[t+p_D(t+1)]+t(1+\beta+p_D)\right)\frac{1}{\Gamma}\left(\frac{1}{\Gamma}-1\right)\frac{d^2\Gamma}{d\tau d\Delta}. \tag{14.30}$$

If $\frac{1}{\Gamma}-1$ is negative, i.e., $\Gamma > 1$, then $\frac{dW}{d\tau d\Delta}$ is positive. When the growth rate of human capital is greater than one, the cross policy effects is positive and vice versa.

Finally, we note that we can derive the optimal tax rate to maximise social welfare in this model. However, above discussed, as such optimal tax rate would be affected by allocation policy, the discussion on optimal tax is too complex to pursue the purpose of this model, that is to examine the welfare effects of both tax and allocation policies. Therefore, we discuss the welfare effects in this paper.

14.4.2 Allocation Effects

Similarly, the effects of changing allocation from national security to public investment in education on welfare are shown as

$$\frac{dW}{d\Delta} = \{p_s [t + p_D (t+1)] + t (1 + \beta + p_D)\} \frac{1}{\Gamma} \frac{d\Gamma}{d\Delta} + \frac{\beta (1-\eta)}{\Delta} + \frac{p_D (1 + p_s)}{(1-\Delta)} > 0. \tag{14.31}$$

As changing allocation of public funds from national security to public investment in education accumulates human capital and consumers can receive the wage income more, they enjoy consumption more. Therefore, such reallocation enhances social welfare.[10]

We confirm the cross-policy effects on welfare in the case of this reallocation at the first stage and the increasing tax rate at second stage. These effects are shown as

$$\frac{dW}{d\Delta d\tau} = -\{p_s [t + p_D (t+1)] + t (1 + \beta + p_D)\} \frac{1}{\Gamma} \left(\frac{1}{\Gamma} - 1\right) \frac{d^2\Gamma}{d\Delta d\tau}. \tag{14.32}$$

As discussed earlier, if $\frac{1}{\Gamma} - 1$ is negative, (i.e., $\Gamma > 1$), $\frac{dW}{d\Delta d\tau}$ is positive. Similarly to the discussion presented above, if $\Gamma > 1$, then the cross policy effects are positive and vice versa.

14.5 Concluding Remarks

As described above, by introducing national security into the overlapping-generations model with public education, we studied how public education and national security affect social welfare. We presented the result that a higher tax rate enhances social welfare when positive benefit effects on welfare are greater than the negative tax effects on welfare. We also presented that changing the allocation of public funds from national security to public investment in education accumulates human capital and enhances social welfare. Furthermore, we assume that the government has two policy instruments and that the government should adopt them: one policy change

[10] From (14.31), we find that the allocation rate, Δ to maximise social welfare is one. However, because of government's functions in the economy, governments must allocate public expenditure for national security. Therefore, it is not realistic that such optimal rate Δ is one.

influences another policy. We discussed the cross effects of two policies on welfare and demonstrated that the cross policy effects are positive if the growth rate of human capital is greater than one.

In the economy in which the population decreases but the probability of national risk increases, government should examine public education specifically, rather than national security. Then, the private agents can better secure their life risk.

Finally, for simplification of discussion, we suppose the constant interest rate and we also assume the probability to survive and that for national risk are constant. In fact, this rate and these probabilities change as time goes on. For future research, it is left how we introduce these insights into the model.

Acknowledgements This work was supported by USPS KAKENHI Grant Numbers JP16H03635 and JP16K03719. A previous version of this paper was presented at the 2018 Fall conference of Japan Association for Applied Economics at Daito Bunka University. The author thanks Junko DOI, Daisuke IKAZAKI, Masafumi TSUBUKU, and other participants for useful comments to the previous version of this paper.

References

Chakraborty, S. 2004. Endogenous lifetime and economic growth. *Journal of Economic Theory* 116: 119–137.

de la Croix, D., and M. Doepke. 2004. Public versus private education when differential fertility matters. *Journal of Development Economics* 73: 607–629.

Diamond, P. 1965. National debt in a neoclassical growth model. *American Economic Review* 55: 1126–1150.

Dioikitopoulos, E.V. 2014. Aging, growth and the allocation of public expenditures on health and education. *Canadian Journal of Economics* 47: 1173–1194.

Fennema, H., and M. Van Assen. 1998. Measuring the utility of losses by means of the tradeoff method. *Journal of Risk and Uncertainty* 17: 277–296.

Glomm, G., and B. Ravikumar. 1992. Public versus private investment in human capital: Endogenous growth and income inequality. *Journal of Political Economy* 100: 818–834.

Kaganovich, M., and I. Zilcha. 1999. Education, social security, and growth. *Journal of Public Economics* 71: 289–309.

Lucas, R.E. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22: 3–42.

Murphy, K.M., and R.H. Topel. 2013. Some basic economics of national security. *American Economic Review* 103: 508–511.

Musgrave, R.A. 1958. *The theory of public finance: A study in public economy*. New York: McGraw-Hill.

Naito, T., and T. Omori. 2016. Household's disaster prevention activities, agglomeration, and economic growth. *Regional Science Policy & Practice* 8: 177–195.

Omori, T. 2009. Effects of public education and social security on fertility. *Journal of Population Economics* 22: 585–601.

Omori, T. 2018. Parent life and children's life: Public health expenditure and human capital formation. *Studies in Applied Economics* 11: 72–78.

Pestieau, P., G. Ponthiere, and M. Sato. 2008. Longevity, health spending, and pay-as-you-go pensions. *FinanzArchiv* 64: 1–18.

Smith, R. 1995. The demand for military expenditure. In *Handbook of defence economics*, vol. 1, 69–87. North Holland, Netherlands.

Slovic, P., B. Fischhoff, S. Lichtenstein, B. Corrigan, and B. Combs. 1977. Preference for insuring against probable small losses: Insurance implications. *Journal of Risk and Insurance* 44: 237–258.

Yakita, A. 2001. Uncertain lifetime, fertility and social security. *Journal of Population Economics* 14: 635–640.

Chapter 15
Optimal Cooperation of Medical Care and Nursing Care in a Two-Region Spatial Model

Tohru Naito

Abstract This paper presents an analysis of the optimal cooperation of medical care service and nursing care service under two regions with an asymmetric density of patients. We apply Aiura and Sanjo (2010) by introducing cooperation between hospitals and nursing care facilities and analyze the effects of cooperation on the equilibrium medical service level, nursing care service level, optimal medical and nursing care fee, and social welfare. Results of the analysis show that the introduction of cooperation between hospitals and nursing care facilities increases social welfare, although it decreases the medical care service and nursing care service level in equilibrium.

15.1 Introduction

Aging is an important issue to be addressed in economically developed countries. For example, the share of elderly persons 65 years old or older among the total population in Japan exceeded 20% in 2005 and rose to 26.3% in 2015. Moreover, Japan has had a declining population since 2010, presenting difficulties of sustainability of the social security system. Although the medical expenditures of elderly people in 1985 were 389.88 billion yen, they increased to 144,927 billion yen in 2014 because the increase in medical expenditures for elderly people is constraining government finances. Meaningful fiscal improvement is not expected. Recently, the operation of an efficient medical system has, therefore, become necessary. When we acutely operate efficient medical care and nursing care systems, it is important to address the current circumstances of elderly people and patients using hospitals and nursing care facilities. For elderly people, the convenience of hospitalization is an important issue. Hospitals and nursing care facilities are not necessarily located in convenient

T. Naito (✉)
Doshisha University, Kyoto, Japan
e-mail: tnaito@mail.doshisha.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_15

places for all elderly people and patients. Situations of low fertility and aging are advancing more rapidly in peripheral areas than in urban areas. However, hospitals and nursing care facilities tend to agglomerate in urban areas. Although hospitals providing advanced medical services agglomerate in urban areas, elderly people and patients in non-medical regions have less access even to area hospitals. Increasingly, they must travel to other areas for consultation. An increase in the number of patients using extraterritorial hospitals can strongly affect the operation of hospitals in different regions.

Doctors engaged at medical facilities per 100,000 in Tokushima prefecture were 315.9 in 2016, which is the highest figure in Japan. Although Tokushima Prefecture has the largest number of doctors aggregated at the prefecture level, its distribution is not necessarily uniform in the prefecture. Figures 15.1 and 15.2 respectively portray the density of medical care staff and nursing care staff in Tokushima prefecture in 2016. The density of medical staff in Western area of Tokushima prefecture is but the density of medical staff is high in eastern and southern areas from Fig. 15.1. Medical institutions that supply advanced medical services such as Tokushima university hospital or Japanese Red Cross hospitals are located in eastern areas in Tokushima Prefecture. Therefore, patients demanding advanced medical services provided by

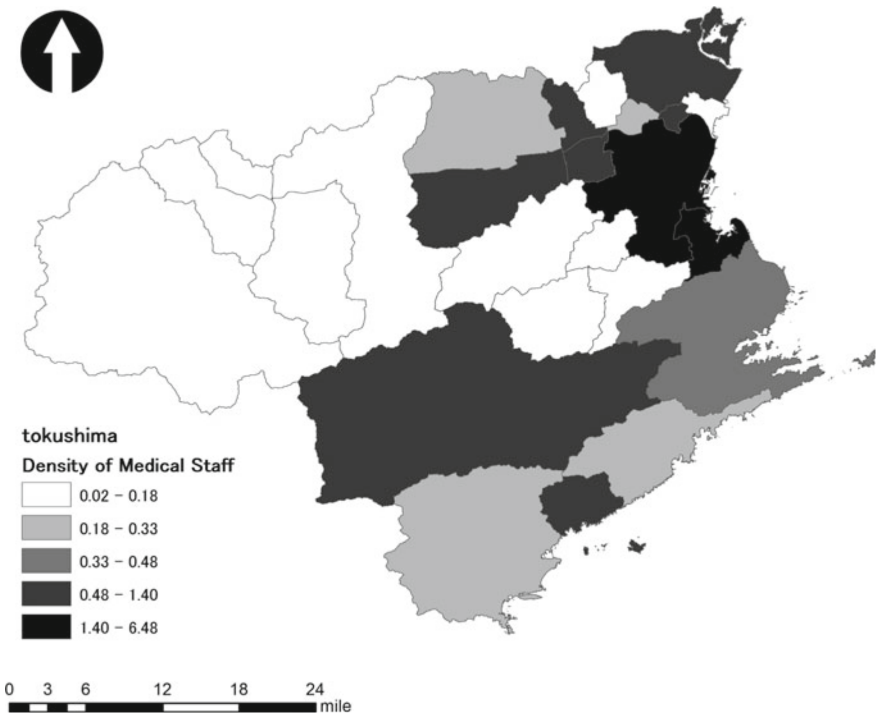


Fig. 15.1 Density of medical staff in Tokushima Pref. Japan

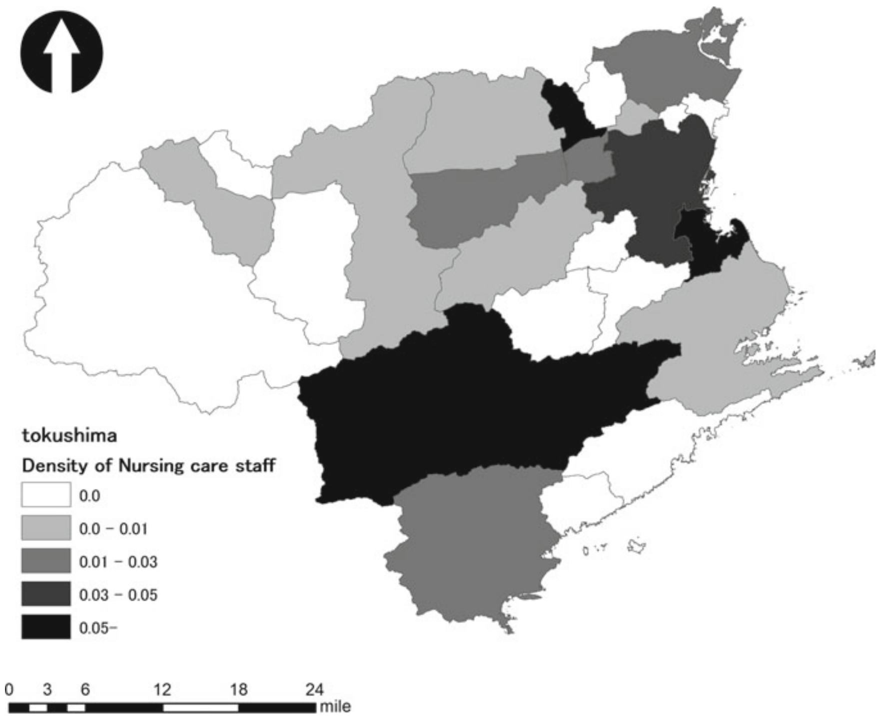


Fig. 15.2 Density of nursing care staff in Tokushima Pref. Japan

these hospitals must bear extra transportation expenditures. Regarding the nursing care staff density, from Fig. 15.2, the nursing care staff is density is known not to be uniform, just as medical staff are not uniform. Furthermore, the population density, the density of doctors, and the density of nursing care staff are not uniform among region. We know that the distribution of population over 65 years of age including the majority of patients who need nursing care and medical care also are not uniform from Fig. 15.3. Such geographical situations should also be examined to construct an optimal medical care and nursing care system.

Banks et al. (2001) describe consideration of the vertical integration of hospitals and the nursing care facilities, analyze the interaction between the medical institution and the nursing care facilities, and derive an optimum payment system. They derive that, if nursing care facilities are reimbursed for services performed, bundling prospective payments for hospitals and nursing care facilities does not remove incentives for vertical integration or induce production efficiency without vertical integration. Mougeot and Naegelen (2005) or Miura and Tajika (2015) extend a bundling payment system model by taking account of Hotelling (1929) and analyzed service provided by hospitals and nursing care facilities under optimum medical service fee and nursing care service fee, respectively. Moreover, they introduce procurement costs into Banks et al. (2001) and confirm the feasibility of cooperation between hos-

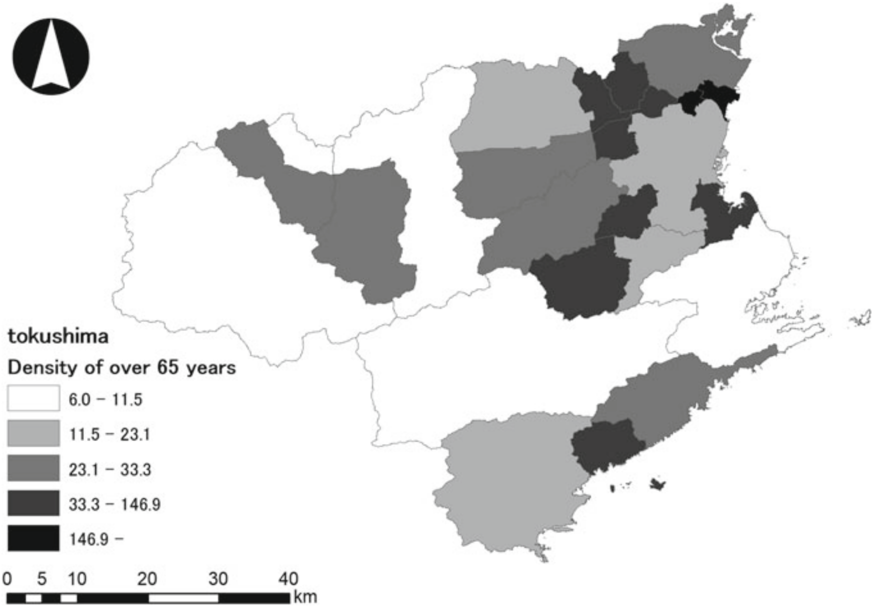


Fig. 15.3 Density of people over 65 years old of Tokushima in 2015

pitals and nursing care facilities. Ma (1994) investigates cost reduction and quality enhancement incentive of cost raiment can implement the first best under adequate condition. As for empirical studies Ettner (1993) adopts date of National Long-term Care Survey and the Area Resources File to analyze the issue of diminished access to nursing home care for elderly medical patients.

Sanjo (2009) takes account of a public hospital, which maximize the social welfare, and a private hospital, which maximize his profit and analyze their medical service level and privatization of public hospital under a traditional Hotelling model. Aiura and Sanjo (2010) consider a duopolistic health care market in which a rural pubic hospital competes against an urban hospital on medical quality under Hotelling spatial competition model which has two regions with asymmetric patients distribution and show that some patients can use a cross border medical service, they do not take into account that hospitals cooperate with nursing care facilities or optimal bundled payment. On the other hand, Mougeot and Naegelen (2005) and Miura and Tajika (2015) do not include two regions with asymmetric patients distribution though they analyze the cooperation between hospitals and nursing care facilities.

The purpose of this paper is to analyze the cooperation between hospitals and nursing care facilities and derive the optimal fee bundling medical service and nursing care service fee. Moreover, we compare medical service level, nursing care service level, and social welfare under cooperation between hospitals and nursing care facilities with these under non-cooperation and analyze the possibility of cooperation. Therefore, this paper complements the blank of the knowledge of previous studies.

The organization of this paper is as follows. The next section presents the basic model, which describes the behavior of patients, hospitals, and nursing care facilities. Section 15.3 analyze the case where hospitals cooperate with nursing care facilities. In Sect. 15.4, we derive the optimal medical fee, nursing care fee, and social welfare under non-cooperation and cooperation and compare them. Finally, we conclude with a discussion of the results derived in this model and present the remaining unsolved issues.

15.2 The Model

15.2.1 Patients

We consider an economy that consists of two regions with a linear city. As described by Aiura and Sanjo (2010), region 1 exists from 0 to $-\frac{1}{p}$, although region 2 exists from 0 to 1. Now patients in region $i (= 1, 2)$ are spread with density p and 1, respectively, because we assume that the distribution of patients in each is a uniform distribution.[1] Therefore, both regions have the homogeneous patients, although the density on each point in region 1 differs from that in region 2. Similar to Aiura and Sanjo (2010) and Miura and Tajika (2015), each hospital is located at the city edge, i.e., the respective locations of hospital in region 1 and region 2 are $-\frac{1}{p}$ and 1. All patients consume one unit of medical service inelastically and gain benefits from their consulted hospital. Let q_i represent medical services from hospital $i (= 1, 2)$. Because patients must move from their residential points to the hospital, the patient residing at $x_i \in [-\frac{1}{p}, 1]$ must bear the transportation cost of $t\left(x_i + \frac{1}{p}\right)$ to go to hospital 1. On the other-hand, the patient must bear the transportation cost of $t(1 - x_i)$ to go to hospital 2. Therefore the utility function of patient who resides at x_i and goes to hospital i is the following.[2]

$$u_{i1} = \alpha q_1 - sM - t\left(x_i + \frac{1}{p}\right) \tag{15.1}$$

and

$$u_{i2} = \alpha q_2 - sM - t(1 - x_i), \tag{15.2}$$

where α, s, M, and t respectively denote the marginal benefit of medical service, the burden rate of health-care costs, medical fee, and unit transportation cost. Although the price is determined by the market mechanism in the general market, medical fees are often regulated by the national health system. Consequently, each patient deals

[1] Here we assume that the range of p satisfies $p \in (0, 1)$.

[2] Although Aiura and Sanjo (2010) assume a quadratic transportation cost, we follow Miura and Tajika (2015) and assume a linear transportation cost with respect to distance.

with s and M as given. Here, we define the residential location of patient for whom the benefits are indifferent, irrespective of the hospital as $\bar{x}$.

$$\bar{x} = \frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha\,(q_1 - q_2)}{2t} \tag{15.3}$$

When a patient resides at the left of $\bar{x}$, they go to hospital 1 to receive medical services. However, when a patient resides at the right of it, they make use of hospital 2 to receive it. Consequently, the demand for medical services of hospital 1 depends on the medical service quality of hospital 1 and hospital 2. Let D_1 and D_2 respectively represent the demand for medical services supplied by hospital 1 and hospital 2. D_1 and D_2 are given as

$$D_1 = \begin{cases} 1 + p\bar{x} & (\bar{x} \leq 0) \\ 1 + \bar{x} & (\bar{x} > 0) \end{cases} \tag{15.4}$$

$$D_2 = 2 - D_1 \tag{15.5}$$

15.2.2 Hospital

Next, we refer to hospitals providing medical services in each region. Medical fees denoted by M are determined by the medical institutions of central governments administering both regions. Consequently, because both hospitals deal with this fee as given, they determine the quality of medical service to maximize profit.[3]

$$\pi_i = (M - q_1)\,D_i - \frac{\beta}{2}q_i^2 \quad (i = 1, 2), \tag{15.6}$$

First, we consider the case in which $\bar{x}$ is not positive, i.e., $\bar{x} \leq 0$. Taking account of (15.1), (15.4), and (15.6), the profits of hospitality 1 are given as follows.

When $\bar{x}$ is not positive,

$$\pi_1(q_1, q_2) = (M - q_1)\left[1 + p\left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha\,(q_1 - q_2)}{2t}\right)\right] - \frac{\beta}{2}q_1^2 \tag{15.7}$$

When $\bar{x}$ is positive,

$$\pi_1(q_1, q_2) = (M - q_1)\left[1 + \left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha\,(q_1 - q_2)}{2t}\right)\right] - \frac{\beta}{2}q_1^2 \tag{15.8}$$

[3] Although Aiura and Sanjo (2010) assume that the fixed cost is constant, we consider that the demand independent cost depends on the medical service quality proposed by the hospital, as do Miura and Tajika (2015).

Because the demand of medical service of hospital 2 is given as (15.5), the profits of hospitality 2 are the following.

When $\bar{x}$ is not positive, then

$$\pi_2(q_1, q_2) = (M - q_2)\left[1 - p\left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha\,(q_1 - q_2)}{2t}\right)\right] - \frac{\beta}{2}q_2^2. \quad (15.9)$$

When $\bar{x}$ is positive, then

$$\pi_2(q_1, q_2) = (M - q_2)\left[1 - \left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha\,(q_1 - q_2)}{2t}\right)\right] - \frac{\beta}{2}q_2^2. \quad (15.10)$$

After differentiating (15.7) and (15.9) with respect to q_1 and q_2, respectively, we solve first-order conditions to maximize (15.7) and (15.9). Consequently, when $\bar{x}$ is not positive, equilibrium medical services of both hospitals are derived as follows.

$$q_1^* = \frac{(3p\alpha + 2t\beta)\,p\alpha M - (5 + p)\,tp\alpha - 2\,(1 + p)\,t^2\beta}{(p\alpha + 2t\beta)\,(3p\alpha + 2t\beta)} \quad (15.11)$$

$$q_2^* = \frac{(3p\alpha + 2t\beta)\,p\alpha M - (7 - p)\,pt\alpha - 2\,(3 - p)\,t^2\beta}{(p\alpha + 2t\beta)\,(3p\alpha + 2t\beta)} \quad (15.12)$$

Here we compare q_1^* with q_2^*, we derive the following relations and consider that q_1^* is smaller than q_2^*.

$$q_1^* - q_2^* = \frac{2t\,(1 - p)}{(3p\alpha + 2t\beta)} > 0. \quad (15.13)$$

Substituting (15.13) for (15.3), $\bar{x}$ under equilibrium medical service is given as

$$\bar{x}^* = \frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha\,(1 - p)}{3p\alpha + 2t\beta} < 0, \quad (15.14)$$

where $\bar{x}^*$ denotes the residential location of a patient whose benefit is indifferent, irrespective of the hospital under equilibrium. Because we assume that $\bar{x}$ is not positive, $\bar{x}^*$ supports this assumption. Consequently, some patients in region 1 use not hospital 1 but hospital 2 although hospital 1 provides higher medical quality than hospital 1. We can present the following economic interpretation for this result. The patients residing in the area near the border between region 1 and region 2 can save transportation costs using hospital 2 rather than hospital 1 because the density of patients in region 1 is lower than that in region 2. Hospital 1 is located at the edge of region 1 $(1/p)$.

Next we consider a case in which $\bar{x}$ is positive, i.e., $\bar{x} > 0$. When $\bar{x}$ is positive, the profit functions of each hospital are given as (15.8) and (15.10). Similar to the process presented above, we respectively differentiate (15.8) and (15.10) with respect to $q_i (i = 1, 2)$. Moreover, solving these first-order conditions to maximize each hospital's profit, equilibrium medical services are derived as presented below.

$$q_1^{**} = \frac{(3\alpha - 2t\beta)\, p\alpha M - t\,(5p-3)\,\alpha + 2t^2\,(3p-1)\,\beta}{p\,(3\alpha - 2t\beta)\,(\alpha - 2t\beta)} \tag{15.15}$$

and

$$q_2^{**} = \frac{(3\alpha - 2t\beta)\, p\alpha M - t\,(3-p)\,\alpha - 2t^2\,(p+1)\,\beta}{p\,(3\alpha - 2t\beta)\,(\alpha - 2t\beta)} \tag{15.16}$$

From (15.15) and (15.16), we can note differences of equilibrium medical services of hospital 1 and hospital 2.

$$q_1^{**} - q_2^{**} = \frac{2t\,(3\alpha\,(1-p) + 4pt\beta)}{p\,(3\alpha - 2t\beta)\,(\alpha - 2t\beta)} > 0 \tag{15.17}$$

Here we assume that marginal benefit of medical service is sufficiently large that all patients in the economy consume one unit of medical service. Particularly, we assume that α is larger than $2t\beta$. We exclude cases, in which $\bar{x}^{**}$ are positive, because (15.15) and (15.16) do not hold that $\bar{x}^{**}$ is positive when α is sufficiently large.[4] Consequently, some patients in region 1 with low density do not go to the hospital in the same region but go to a hospital in region 2. From (15.13), the medical service quality of hospital in region 1, which has low population density, is greater than that of hospital in region 2.

15.2.3 Nursing Care Facility

Next we refer to the behavior of nursing care service in the same spatial configuration. Although one purpose of this paper is to analyze the effects of integration of medical services and nursing care services on social welfare, we consider only the behavior of nursing care service as a benchmark. Similar to a hospital's location, we assume that each nursing care facility is also located at the edge of each region. Nursing care facility 1 and nursing care facility 2 are located respectively in $\frac{1}{p}$ and 1. Following Miura and Tajika (2015), let w_i represent the services level of nursing care facility $i(=1,2)$. Moreover, we assume that patients get benefit γw_i from nursing care facility i. Let v_i represent the net benefits of patient with residential point x_i to make use of nursing care facility i are

$$v_1 = \gamma w_1 - sN - t\left(x_i + \frac{1}{p}\right) \tag{15.18}$$

and

[4]Please see Appendix A for this reason.

$$v_2 = \gamma w_2 - sN - t(1 - x_i), \tag{15.19}$$

where s and N respectively denote the burden rate of nursing care costs and nursing care fees.[5] Let $\hat{x}$ represent the residential location of a patient who is indifferent to the choice between nursing care facility 1 and nursing care facility 2. Taking account of (15.18) and (15.19), $\hat{x}$ is given as shown below.

$$\hat{x} = \frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\gamma}{2t}(w_1 - w_2) \tag{15.20}$$

Because all patients consume one unit of nursing care service inelastically, patients who reside on the left of $\hat{x}$ consume it from nursing care service facility 1. However, the other consumes it from nursing care facility 2. Now we describe d_1 and d_2 as the total numbers of patients who consume nursing care services from facility 1 and facility 2, respectively, i.e., d_1 and d_2 are denoted as

$$d_1 = \begin{cases} 1 + p\hat{x} & (\hat{x} \leq 0) \\ 1 + \hat{x} & (\hat{x} > 0) \end{cases} \tag{15.21}$$

$$d_2 = 2 - \hat{x} \tag{15.22}$$

Next we consider the behaviors of respective nursing care service facilities based on (15.21) and (15.22). Similar to hospitals described in the previous section, their objective function is to maximize his profit denoted as π_i^n $(i = 1, 2)$. The profit functions of the respective nursing care service facilities are the following.[6]

When $\hat{x}$ is not positive,

$$\pi_1^n(w_1, w_2) = (N - c)\left[1 + p\left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\gamma(w_1 - w_2)}{2t}\right)\right] - \frac{\delta}{2}\left(w_q\right)^2. \tag{15.23}$$

When $\hat{x}$ is positive,

$$\pi_1^n(w_1, w_2) = (N - c)\left[1 + \left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\gamma(w_1 - w_2)}{2t}\right)\right] - \frac{\delta}{2}(w_1)^2. \tag{15.24}$$

When $\hat{x}$ is not positive,

[5]Similarly to medical service fees, we assume that s and N are also determined by the central government. For simplification, burden rates of medical care cost and nursing care costs has a common rate.

[6]It is noteworthy that the cost function of nursing care service facility differs from that of a hospital. Although the marginal cost of medical service is an increasing function with respect to service quality, that of nursing care service is constant. Hospitals must install advanced medical equipment to increase medical service quality generally. However, the job of nursing care service requires no advanced medical equipment than medical service. Therefore, we consider that the cost for a unit of nursing care service is constant in our model.

$$\pi_2^n(w_1, w_2) = (N-c)\left[1 - p\left(\frac{1}{2}\left(1-\frac{1}{p}\right) + \frac{\gamma(w_1-w_2)}{2t}\right)\right] - \frac{\delta}{2}(w_2)^2 . \tag{15.25}$$

When $\hat{x}$ is positive,

$$\pi_2^n(w_1, w_2) = (N-c)\left[1 - \left(\frac{1}{2}\left(1-\frac{1}{p}\right) + \frac{\gamma(w_1-w_2)}{2t}\right)\right] - \frac{\delta}{2}(w_2)^2 . \tag{15.26}$$

To begin with, we consider the equilibrium nursing care service under $\hat{x} \leq 0$. Here we differentiate (15.23) and (15.25) to derive the respective first-order conditions to maximize profit with respect to w_1 and w_2. Moreover, solving these first-order conditions, the equilibrium nursing care service of the respective facilities is the following.

When $\hat{x}$ is not positive,

$$w_1^* = w_2^* = \frac{(N-c)\,\gamma}{2t\delta}. \tag{15.27}$$

When $\hat{x}$ is positive, the equilibrium nursing care service of each facility is given as

$$w_1^{**} = w_2^{**} = \frac{(N-c)\,\gamma}{2t\delta}. \tag{15.28}$$

From (15.27) and (15.28) equilibrium nursing care service level is known to be independent of the density of p; both nursing care service facilities supply the same level. Substituting (15.27) and (15.28) for $\hat{x}$, respectively, the residential point of a patient who is indifferent to the choice between the nursing care facilities, in equilibrium is the following.

$$\hat{x}^* = \hat{x}^{**} = \frac{1}{2}\left(1-\frac{1}{p}\right) < 0 \tag{15.29}$$

Similar to discussions of a point at which the net benefits of patients from each hospital are not different, cases in which $\hat{x}$ is positive is excluded from analysis. Here we compare (15.14) with (15.29). $\bar{x}^*$ is smaller than $\hat{x}^*$. Without cooperation between a hospital and a nursing care service facility, the types of patients are classified as represented below (Fig. 15.4).

$$Paitient\ type = \begin{cases} hospital\ 1\ and\ nursing\ care\ facility\ 1\ (x < \bar{x}^*) \\ hospital\ 2\ and\ nursing\ care\ facility\ 1\ (\bar{x}^* \leq x < \hat{x}^*) \\ hospital\ 2\ and\ nursing\ care\ facility\ 2\ (otherwise) \end{cases} \tag{15.30}$$

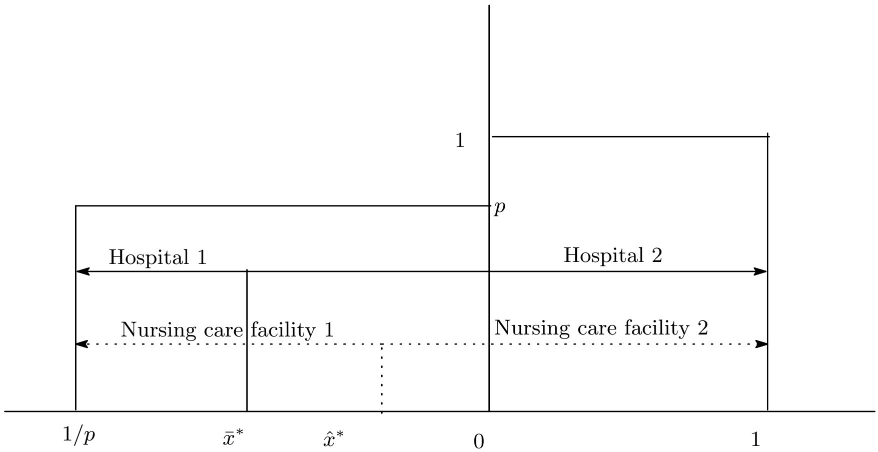


Fig. 15.4 Patients access type of hospital and nursing care facility

15.2.4 Optimal Medical Service Fee

We derived the equilibrium medical service level as described in the section above. Because we assume that medical fees are determined by the medical care system of central government, patients and hospitals cannot determine it. They deal with it as given. In this section, we derive the optimal medical fee denoted by M to maximize social welfare. First, we must define a social welfare function. The social welfare function in each region consists of total net benefit of patients and the profit of hospital located in each region. Let W_i^* represent the social welfare function in region $i (= 1, 2)$ as

$$W_1^* = p\left\{\int_{-\frac{1}{p}}^{\bar{x}^*}\left(aq_1^* - sM - t(y + \frac{1}{p})\right)dy + \int_{\bar{x}}^{0}\left(aq_2^* - sM - t(1-y)\right)dy\right\} + \pi_1^* \tag{15.31}$$

and

$$W_2^* = \int_0^1 \left(aq_2^* - sM - t(1-y)\right)dy + \pi_2^* \tag{15.32}$$

We assume that the objective function of central government is to maximize the sum of W_1^* and W_2^*, which is denoted by W^*. Differentiating W^* with respect to M, the first-order condition for social welfare maximization with respect to M is given as

$$\frac{\partial W^*}{\partial M} = \left(-\frac{2\beta p^2\alpha^2}{(p\alpha + 2t\beta)^2}\right)M + \left(\frac{2\beta\left(4\beta t^2 + 4p\alpha t\right)}{(p\alpha + 2t\beta)^2} + \frac{2p\alpha^2}{p\alpha + 2t\beta} - 2s\right) = 0. \tag{15.33}$$

Solving (15.33) with respect to M, the optimal medical service fee M^* is the following.

$$M^* = \frac{\left(p^2\alpha^3 + 4t^2\beta^2 + 4pt\alpha\beta + 2pt\alpha^2\beta\right)}{\beta p^2\alpha^2} - \frac{s\,(p\alpha + 2t\beta)^2}{\beta p^2\alpha^2} \tag{15.34}$$

Next we analyze the effect of each parameter on optimal medical service fees with comparative static analysis. Differentiating M^* with respect to α, β, t, s and p, we obtain the following results.

$$\frac{\partial M^*}{\partial \alpha} \gtreqless 0,\ \frac{\partial M^*}{\partial \beta} \gtreqless 0,\ \frac{\partial M^*}{\partial t} > 0,\ \frac{\partial M^*}{\partial s} < 0,\ \frac{\partial M^*}{\partial p} < 0$$

Because the signs of $\frac{\partial M^*}{\partial \alpha}$ and $\frac{\partial M^*}{\partial \beta}$ are ambiguous, the effect of α or β on M^* is not determined uniquely. However, the effect of t, s, and p on M^* is determined uniquely. Although the increase in transportation cost increases the optimal medical service fee, the increases in s and p decrease it.

15.2.5 Optimum Nursing Care Service Fee

Next we consider the optimal nursing care service fee under the case in which nursing care service facilities and hospitals do not cooperate. The point at which the patient is indifferent to net benefits from nursing care services irrespective of the facility is given as (15.29). Let W_i^{n*} represent the social welfare function in region $i\,(= 1, 2)$.[7] Substituting (15.29) for (15.20), we obtain the point at which a patient is indifferent to net benefits from nursing care services, irrespective of the facility. That point is given as shown below.

$$\hat{x}^* = \frac{1}{2}\left(1 - \frac{1}{p}\right) \tag{15.35}$$

Considering (15.35), W_1^{n*} and W_2^{n*} in equilibrium are the following.

$$W_1^{n*} = p\left[\int_{-\frac{1}{p}}^{\hat{x}^*}\left(\gamma w_1 - sN - t\left(y + \frac{1}{p}\right)\right)dy + \int_{\hat{x}^*}^{0}(\gamma w_2 - sN - t(1-y))dy\right] + \pi_1^{n*} \tag{15.36}$$

and

$$W_2^{n*} = \int_0^1 (\gamma w_2^* - sN - t(1-y))dy + \pi_2^{n*} \tag{15.37}$$

[7]Strictly speaking, we should consider both medical service and nursing care service when we consider social welfare.

Similarly to the medical service fee, we assume that the central government determines the nursing care fee to maximize W^{n*}, which is equal to $W_1^n + W_2^n$, in equilibrium. Differentiating W^{n*} with respect to N and solving the first-order condition with respect to N, the optimal nursing care service fee is derived as shown below.

$$N^* = \frac{4t^2\delta(1-s) + p\gamma^2(2t+cp)}{p^2\gamma^2} \tag{15.38}$$

Here we analyze the effect of parameters on optimal nursing care service fees. Differentiating N^* with respect to γ, c, δ, t, s, and p, we obtain the following results.

$$\frac{\partial N^*}{\partial \gamma} < 0, \frac{\partial N^*}{\partial c} > 0, \frac{\partial N^*}{\partial \delta} > 0, \frac{\partial N^*}{\partial t} > 0, \frac{\partial N^*}{\partial s} < 0, \frac{\partial N^*}{\partial p} < 0$$

As the results of comparative static analysis showed, we know that increases in γ, s, and p decrease the optimal nursing care service fee, although increased δ, c, and t increase it.

15.3 Cooperation Between Medical and Nursing Care

In the previous section, we respectively analyzed the markets of medical services and nursing care services. We do not consider cooperation between a hospital and a nursing care service facility. Recently, cooperation between nursing care and medical care for patients, especially elderly people, has been very important in aging society. In reality, medical care and nursing care often have a close relation. Therefore, we reconstruct the earlier models of medical service market and nursing care service to describe cases in which hospitals cooperate with nursing care service facilities in each region. Similar to previous studies, medical and nursing care facility supplies both medical service and nursing care service. Because the behavior of hospitals and nursing care facilities were considered independently, as for patients, the locations of nursing care facilities are not necessarily the same locations as hospitals in equilibrium, as

$$V_1 = \alpha\tilde{q}_1 + \gamma\tilde{w}_1 - sR - t\left(x_i + \frac{1}{p}\right) \tag{15.39}$$

and

$$V_2 = \alpha\tilde{q}_2 + \gamma\tilde{w}_2 - sR - t(1 - x_i), \tag{15.40}$$

where $\tilde{q}_i$, $\tilde{w}_i$, and R respectively denote medical service level and nursing care service level provided by medical and nursing care hospital $i (= 1, 2)$ and medical and nursing care fee. Let $\tilde{x}$ represent the point of a patient who gets the same surplus from either of medical and nursing care hospital. Consequently, $\tilde{x}$ satisfies the following relation.

$$\tilde{x} = \frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha(\tilde{q}_1 - \tilde{q}_2)}{2t} + \frac{\gamma(\tilde{w}_1 - \tilde{w}_2)}{2t} \tag{15.41}$$

When we define $\tilde{D}_1$ as the demand of service provided by the medical and nursing care hospitals in region 1, it is given as

$$\tilde{D}_1 = \begin{cases} 1 + p\tilde{x} & (\tilde{x} \leq 0) \\ 1 + \tilde{x} & (\tilde{x} > 0) \end{cases} \tag{15.42}$$

$$\tilde{D}_2 = 2 - \tilde{D}_1 \tag{15.43}$$

Now we assume that the marginal benefit of medical service is efficiently large, i.e., $2t\alpha\delta - p\theta\gamma^2 > 0$. Because the medical and nursing care hospitals in region i provide both medical services and nursing care services, the objective functions of medical and nursing care hospitals in region i are presented below.

When $\tilde{x}$ is not positive,

$$\begin{aligned}\tilde{\pi}_1(\tilde{q}_1, \tilde{q}_2, \tilde{w}_1, \tilde{w}_2) = (R - \theta\tilde{q}_1 - \theta c)\left[1 + p\left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha(\tilde{q}_1 - \tilde{q}_2)}{2t} + \frac{\gamma(\tilde{w}_1 - \tilde{w}_2)}{2t}\right)\right] \\ - \frac{\beta}{2}(\tilde{q}_1)^2 - \frac{\delta}{2}(\tilde{w}_1)^2\end{aligned} \tag{15.44}$$

and

$$\begin{aligned}\tilde{\pi}_2(\tilde{q}_1, \tilde{q}_2, \tilde{w}_1, \tilde{w}_2) = (R - \theta\tilde{q}_2 - \theta c)\left[1 - p\left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha(\tilde{q}_1 - \tilde{q}_2)}{2t} + \frac{\gamma(\tilde{w}_1 - \tilde{w}_2)}{2t}\right)\right] \\ - \frac{\beta}{2}(\tilde{q}_2)^2 - \frac{\delta}{2}(\tilde{w}_2)^2,\end{aligned} \tag{15.45}$$

where θ denotes the efficient cooperation effect parameter, smaller than one.

When $\tilde{x}$ is positive,

$$\begin{aligned}\tilde{\pi}_1(\tilde{q}_1, \tilde{q}_2, \tilde{w}_1, \tilde{w}_2) = (R - \theta\tilde{q}_1 - \theta c)\left[1 + \left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha(\tilde{q}_1 - \tilde{q}_2)}{2t} + \frac{\gamma(\tilde{w}_1 - \tilde{w}_2)}{2t}\right)\right] \\ - \frac{\beta}{2}(\tilde{q}_1)^2 - \frac{\delta}{2}(\tilde{w}_1)^2\end{aligned} \tag{15.46}$$

and

$$\begin{aligned}\tilde{\pi}_2(\tilde{q}_1, \tilde{q}_2, \tilde{w}_1, \tilde{w}_2) = (R - \theta\tilde{q}_2 - \theta c)\left[1 - \left(\frac{1}{2}\left(1 - \frac{1}{p}\right) + \frac{\alpha(\tilde{q}_1 - \tilde{q}_2)}{2t} + \frac{\gamma(\tilde{w}_1 - \tilde{w}_2)}{2t}\right)\right] \\ - \frac{\beta}{2}(\tilde{q}_2)^2 - \frac{\delta}{2}(\tilde{w}_2)^2\end{aligned} \tag{15.47}$$

Differentiating (15.44) and (15.45) with respect to $\tilde{q}_i$ and $\widetilde{w}_i$, respectively, and solving $\frac{\partial \tilde{\pi}_i}{\partial \tilde{q}_i} = 0$ and $\frac{\partial \tilde{\pi}_i}{\partial \tilde{q}_i} = 0$, $(i = 1, 2)$, we obtain the following $\tilde{q}_i^*$ and $\widetilde{w}_i^*$, $(i = 1, 2)$ under

$\tilde{x} < 0$

$$\tilde{q}_1^* = \frac{2pt\alpha\delta\,(R - c\theta)\,(A - B) - 2t^2\theta\delta[(1+p)\,A - (3-p)\,B]}{A^2 - B^2}, \tag{15.48}$$

$$\tilde{q}_2^* = \frac{2pt\alpha\delta\,(R - c\theta))(A - B) - 2t^2\theta\delta\,[(3-p)\,A - (1+p)\,B]}{A^2 - B^2}, \tag{15.49}$$

$$\tilde{w}_1^* = \frac{p\gamma}{2t\delta}\left(R - \theta(c + \tilde{q}_1^*)\right), \tag{15.50}$$

and

$$\tilde{w}_2^* = \frac{p\gamma}{2t\delta}\left(R - \theta(c + \tilde{q}_2^*)\right), \tag{15.51}$$

where A and B are defined as shown below.

$$A \equiv 4t\delta(t\beta + p\theta\alpha) - p^2\gamma^2\theta^2$$

$$B \equiv p^2\gamma^2\theta^2 - 2t\delta p\theta\alpha$$

When a hospital in each region cooperates with a nursing care facility in the same region, the nursing care service level depends on the medical service level and the efficient cooperation effect denoted by θ. When $\tilde{g}_i{}^*$ decreases $\tilde{q}_i{}^*$, the equilibrium nursing care service level decreases in region i. However, the effect of θ on nursing care service level in equilibrium is ambiguous because the effect of θ on $\tilde{q}_i{}^*$ is ambiguous though the direct cooperation efficient effect is decrease $\tilde{w}_i^*$.

Substituting (15.48), (15.49), (15.50), and (15.51) for (15.41), equilibrium point, in which the net benefit of patient is the same level regardless of medical and nursing care hospital, is given as shown below.

$$\tilde{x}^* = \frac{1}{2}\left(1 - \frac{1}{p}\right) - \frac{p\theta\gamma^2 - 2t\alpha\delta}{A - B}\,(1-p)\,\theta < 0 \tag{15.52}$$

Because $\tilde{x}^*$ is not positive under $\tilde{q}_1^*$, $\tilde{q}_2^*$, $\tilde{w}_1^*$, and $\tilde{w}_2^*$,we know $\tilde{x}^*$ meets the assumption. Here we define the social welfare foundation to consider the optimal medical and nursing care fee denoted by R. Let $\tilde{W}^*$ represent the equilibrium social welfare under cooperation of medical services and nursing care services. Let $\tilde{W}_i^*$ represent the equilibrium social welfare in region i under cooperation of medical services and nursing care services. $\tilde{W}_i^*$ are as described below.

$$\tilde{W}_1^* = p\left[\int_{\frac{1}{p}}^{\tilde{x}^*}\left(\alpha\tilde{q}_1 + \gamma\tilde{w}_1 - sR - t\left(y + \frac{1}{p}\right)\right)dy + \int_{\tilde{x}^*}^{0}(\alpha\tilde{q}_2 + \gamma\tilde{w}_2 - sR - t\,(1-y))dy\right] + \tilde{\pi}_1^* \tag{15.53}$$

and

$$\tilde{W}_2^* = \int_0^1 (\alpha \tilde{q}_2^* + \gamma \tilde{w}_2^* - sR - t(1-y))dy + \tilde{\pi}_2^* \tag{15.54}$$

Next we define the social welfare of the economy as the sum of social welfare in region 1 and region 2. Consequently, $\tilde{W}^*$ is given as $\tilde{W}_1^* + \tilde{W}_2^*$ because R^* is determined to maximize $\tilde{W}^*$ with respect to R. Differentiating $\tilde{W}^*$ with respect to R, we derive the following optimal medical and nursing care fee under cooperation between medical services and nursing care services.

$$\begin{aligned}\tilde{R}^* = & \frac{1}{p\beta\left(Z(p\theta\gamma^2 - 2t\alpha\delta) - p\gamma^2\right)} \\ & \times \left\{p\,(Z\theta - 1)\,(2t\beta + (\alpha + c\beta - \theta)\,p\theta)\,\gamma^2 - 2Zpt\theta\delta\alpha^2 \right. \\ & \left. - \; 2t\delta\left((1-s)\,p\theta + \left((2t + cp\theta)\beta - p\theta^2\right)Z\right)\alpha - 4t^2\beta\delta\,(1-s)\right\},\end{aligned} \tag{15.55}$$

where Z is defined as shown below.

$$Z \equiv \frac{4\,(1-p)\,t^2\delta\theta}{-2p^2\theta^2\gamma^2 + 6\alpha\delta pt\theta + 4\beta\delta t^2}$$

15.4 Comparison

15.4.1 Medical Service Level

We derive the medical service, nursing care service, and optimal medical or nursing care fee in equilibrium at each institution. Particularly, we compare the case without cooperation in hospital and nursing care facility with that with mutual cooperation. In our model medical service fee, nursing care service fee, and inclusive fee are determined by the central governments administering the respective regions. Furthermore, they are given, respectively, as (15.34), (15.38), and (15.55). When the hospitals do not cooperate with nursing care facilities, the medical service fee and nursing care service fee are determined independently. Hospitals and nursing care facilities choose medical service level and nursing care service levels respectively under (15.15), (15.16), and (15.27).

Although we derive the optimal medical fee and nursing care fee under cooperation and non-cooperation in (15.34), (15.38), and (15.55), it is difficult to compare them because (15.55) is complex in particular. Therefore, we investigate their mutual difference with an appropriate numerical example.[8] The red line and blue line in

[8]The values of the parameters used in Fig. 15.5 are $\alpha = 4,\ \gamma = 2,\ \beta = \delta = c = t = 1,\ s = 0.3$. Because the self-pay ratio is 30% of the total medical fee in Japan, we set the value of s to 0.3.

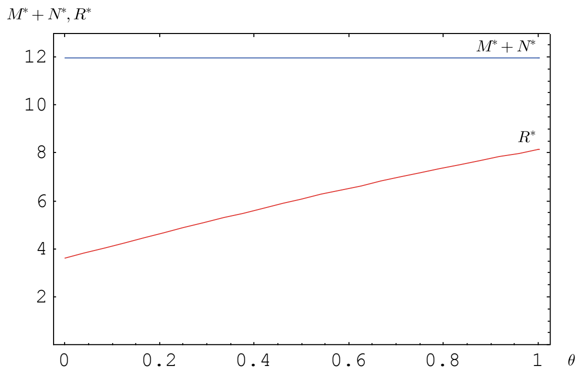


Fig. 15.5 Medical fee under nursing care fee under cooperation and non-cooperation

Fig. 15.5 respectively denote the optimal cooperation fee of medical service and nursing care service and the sum of optimal medical service fee and nursing care service fee under non-cooperation. As shown by Fig. 15.5, the optimal cooperation fee of medical service and nursing care service is less than the sum of optimal medical service fee and nursing care service fee under non-cooperation. Here we refer to the effect of efficient cooperation effect parameter on R^*. Because θ denotes the efficient cooperation effect parameter and because it is smaller than one, the decrease in θ represents improvement of the efficiency of cooperation between medical care and nursing care. Consequently, we derive the following proposition.

Proposition 1

As the efficient cooperation effect between medical service and nursing care service becomes more efficient, the optimal cooperation fee of medical service and nursing care service decreases. Moreover, it is smaller than the sum of optimal medical service fee and nursing care service fee under non-cooperation.

We compare the level of medical service provided by each hospital under cooperation and non-cooperation. The optimum medical service fee under non-cooperation and the medical and nursing care service fee under cooperation are given respectively as (15.34) and (15.55). Because it is difficult to compare $\tilde{q}_1^*$ with $\tilde{q}_2^*$ analytically, we confirm the difference between $\tilde{q}_1^*$ and $\tilde{q}_2^*$ using a numerical example. From Fig. 15.6, we can infer that $\tilde{q}_1^*$ and $\tilde{q}_2^*$ under cooperation are smaller, respectively, than q_1^* and q_2^* under non-cooperation. In comparison with $\tilde{q}_1^*$, we know that the value of $\tilde{q}_2^*$ is large irrespective of θ. Therefore, cooperation between medical care services and nursing care services engenders decreases of medical service levels before it. Although each hospital increases the medical service level along with cooperation efficiency, the medical service level is lower than the medical service level under non-cooperation.

Moreover, we assume that the marginal benefits of medical service are greater than those of nursing care services.

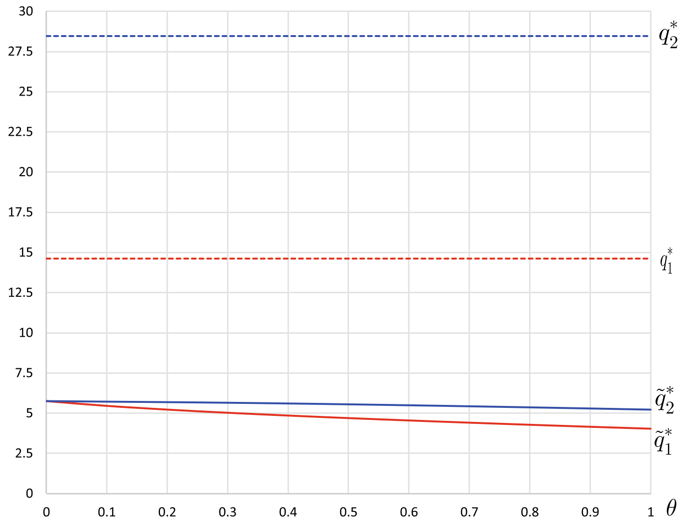


Fig. 15.6 Medical service level under optimal medical fee

15.4.2 Nursing Care Service Level

$\tilde{w}_1^*$ and $\tilde{w}_2^*$ are affected not only by a cooperation fee of medical service and nursing care service but also the medical service level. Differentiating $\tilde{w}_i^*$ with respect to θ, one obtains the following relation.

$$\frac{\partial \tilde{w}_i^*}{\partial \theta} = \frac{p\gamma}{2t\delta}\left\{\frac{\partial R^*}{\partial \theta} - (c + \tilde{q}_i^*) - \theta \frac{\partial \tilde{q}_i^*}{\partial \theta}\right\} \gtreqless 0, \quad (i = 1, 2) \tag{15.56}$$

In general, the sign of (15.56) is ambiguous because the effect of θ on optimal medical and nursing care fee is not determined uniquely. The increase in θ increases the nursing care service level in equilibrium if the effect of θ on optimal medical and nursing care fee under their cooperation is sufficiently large. Next, we confirm the effect of θ on $\tilde{w}_i^*$ under the same parameters as Fig. 15.2. We describe the effect of θ on optimal nursing care service under the same parameters in Fig. 15.6 and 15.7.

15.4.3 Social Welfare

Finally, we compare social welfare under non-cooperation with that under cooperation. Here we define the sum of the patient's benefit and hospitals, and nursing care facilities in both regions because the social welfare in each region under non-cooperation is given as (15.53) and (15.54). Substituting (15.48), (15.49), (15.50), (15.51), (15.52), and (15.55) for (15.53) and (15.54), the equilibrium social welfare

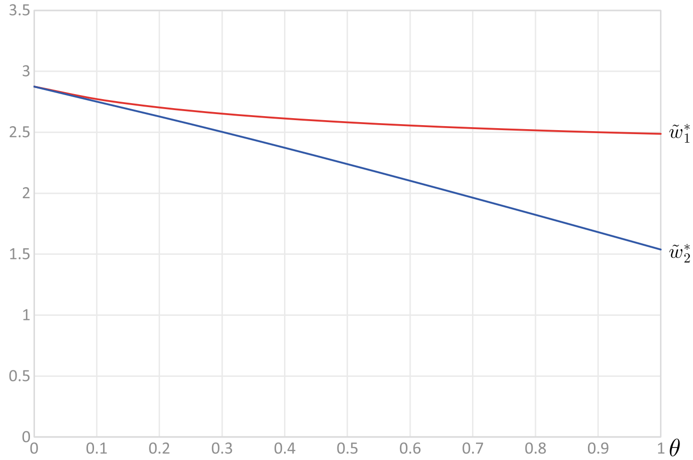


Fig. 15.7 Nursing care service level under cooperation and non-cooperation

under cooperation is the following.

$$\tilde{W}^* = \tilde{W}_1^* + \tilde{W}_2^* \tag{15.57}$$

We define $W^* + W^{n*}$ as the equilibrium social welfare function under non-cooperation. We compare equilibrium social welfare under cooperation with that under non-cooperation. Similarly to previous analyses, we adopt the same parameters and assess their mutual differences with a numerical example. $\tilde{W}^*$ and $W^* + W^{n*}$ are described, respectively, red line and blue line in Fig. 15.8. As the efficient cooperation effect between medical service and nursing care service becomes more efficient, the social welfare increases in equilibrium. That is, W^* is a decreasing function of θ. We know that the social welfare under cooperation always exceeds social welfare under non-cooperation in $\theta \in [0, 1]$. It is possible to improve social welfare by cooperation between medical institutions (hospital) and nursing care facilities. Therefore, the following proposition is derived.

Proposition 2

As the efficient cooperation effect between medical service and nursing care service becomes more efficient, social welfare increases in equilibrium. Moreover, it is possible to improve social welfare through cooperation between medical institutions (hospitals) and nursing care facilities.

15.5 Concluding Remarks

In our model, we consider the regions with asymmetric patients' distribution, in which the patient density in region 1 is lower than that in region 2. First, we derive the behavior of hospitals, nursing care facility, and patients under the case in which no cooperation exists among hospitals and nursing care facility. Moreover, we derive the optimal medical service fee and nursing care service fee to maximize social welfare under no cooperation between them. We show that a part of patients in region 1 with low density go not to hospital 1 but hospital 2, i.e., transboundary usage is caused. Next, we consider the case in which cooperation exists between hospitals and nursing care facility. When we consider their mutual cooperation, we consider the efficiency of cooperation. After obtaining the optimal comprehensive service fee for medical and nursing care denoted by R^*, we compare the difference of optimal service fee, the medical service, nursing care service, and social welfare between cooperation and non-cooperation. Because some endogenous variables have complicated values in equilibrium, in general, the effects of parameters on these endogenous variables are ambiguous. Therefore, we adopted reasonable parameters and confirmed and compared these effects by numerical examples. Regarding medical service fees with nursing care fees under cooperation and non-cooperation, the sum of optimal medical service fee and nursing care service fee under non-cooperation is greater than that under cooperation irrespective of the cooperation efficiency. Moreover, we confirm that the optimal comprehensive service fee of medical service and nursing care service decreases as the efficient cooperation effect between medical service and nursing care service becomes more efficient. Next we confirmed differences of medical service levels provided by hospitals under cooperation and non-cooperation. This result is presented in Fig. 15.6. Figure 15.6 shows that medical service levels under non-cooperation are lower than those under cooperation. Although cooperation in the supply of medical care and the nursing care increases the standard of a

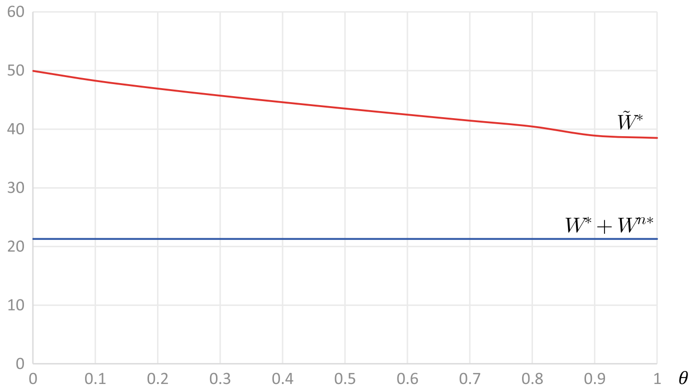


Fig. 15.8 Comparison of social welfare under cooperation and non-cooperation

health care service and the care service intuitively, the results of simulation do not necessarily support that intuition, possibly because of cooperation mechanisms similar to mechanisms of spill over effects as public goods. Moreover, the improvement of cooperation efficiency is known to increase the respective medical service levels. Also, differences of medical service levels between hospitals under cooperation are reduced by improvement of the cooperation efficiency. The marginal cost for the medical service level becomes zero when the cooperation is most efficient. Because both hospitals have equal fixed costs for medical service levels in this case, no differences exist between medical service provision costs. In addition, nursing care service levels depend not only on the cooperation efficiency effect but also on the medical service level provided by hospitals in the same region. Although the improvement of cooperation efficiency increases the nursing care service level provided by both nursing care facilities, the effect of cooperation efficiency on nursing care service in region 2 is greater than that in region 1.

Finally we demonstrate that the social welfare obtained under cooperation is greater than that obtained under non-cooperation. Improving the cooperation efficiency, the social welfare increases under cooperation. Although $R*$ and $\tilde{q}_i (i = 1, 2)$ under cooperation is smaller than these under non-cooperation, the social welfare under cooperation is greater than that obtained under non-cooperation. Recently the promotion of cooperation between medical care service and nursing care service has been studied in Japan. Consequently, the results derived in our analysis of this paper support the promotion policy of cooperation between hospitals and nursing care facilities.

This paper presented analysis of the cooperation between medical care (hospital) and nursing care facility in regions with an asymmetric patient distribution. However, we do not compare equilibriums among areas of different patient densities. For complexity of some endogenous variables in equilibrium, we conduct a comparison between cooperation and non-cooperation with numerical example and simulations. Qualitative analyses must be applied to elucidate comparisons between them. These points will be addressed in our future work.

Acknowledgements This study is supported by Grant-in-Aid for Scientific Research B (17H02533) and Grant-in-Aid for Scientific Research C (18K01632) from Japan Society of the Promotion of Science. The previous version of this chapter was presented in 2018 annual meeting of Korean Association of Applied Economics. The author thanks Seik Kim (Korean University) for very useful comments to the previous version of this chapter.

Appendix

Proof of $\bar{x}^{**} \not> 0$

Substituting (15.17) for (15.3), $\bar{x}^{**}$ in equilibrium is given as shown below.

$$\bar{x}^{**} = \frac{1}{2}\left(1 - \frac{1}{p}\right) + \left(\frac{\alpha\left(3\alpha\left(1-p\right) + 4pt\beta\right)}{p\left(3\alpha - 2t\beta\right)\left(\alpha - 2t\beta\right)}\right) > 0 \tag{15.58}$$

Now we assume that α is sufficiently large. When α is sufficiently large, p is larger than 1. However, (15.58) is not satisfied because the range of p is from 0 to 1. Consequently, $\bar{x}^{**}$ is not positive.

Proof of $\tilde{x}^* \leq 0$

Substituting (15.48), (15.49), (15.50), and (15.51) for (15.52), we derive the $\tilde{x}^*$ in equilibrium as follows.

$$\tilde{x}^* = \frac{1}{2}\left(1 - \frac{1}{p}\right) - \frac{\left(p\theta\gamma^2 - 2t\alpha\delta\right)}{\Gamma}\left[\theta\left(1-p\right)2t\delta\left(2t\beta + p\theta\alpha\right)\right] \tag{15.59}$$

Because $\tilde{x}^*$ is not positive, the following inequality must hold.

$$\frac{1}{2}\left(1 - \frac{1}{p}\right) - \frac{\left(p\theta\gamma^2 - 2t\alpha\delta\right)}{\Gamma}\left[\theta\left(1-p\right)2t\delta\left(2t\beta + p\theta\alpha\right)\right] < 0 \tag{15.60}$$

Simplifying (15.60), we obtain the simplified inequality as follows.

$$\frac{1}{2}\left(1 - \frac{1}{p}\right) - \frac{\left(p\theta\gamma^2 - 2t\alpha\delta\right)}{\Gamma}\left[\theta\left(1-p\right)2t\delta\left(2t\beta + p\theta\alpha\right)\right] < 0 \Leftrightarrow -p\theta\alpha < 2\beta t \tag{15.61}$$

Now $-p\theta\alpha < 2\beta t$ always holds because we assume that p, θ, α, β, and t are positive. Consequently, we can prove that $\tilde{x}^*$ is not positive under $\tilde{q}_1^*$, $\tilde{q}_2^*$, $\tilde{w}_1^*$, and $\tilde{w}_2^*$.

References

Aiura, H., and Y. Sanjo. 2010. Privatization of local public hospitals: effect on budget, medical service quality, and social welfare. *International Journal of Health Care Finance and Economics* 10: 275–299.

Banks, D., E. Parker, and J. Wendel. 2001. Strategic interaction among hospitals and nursing facilities: The efficiency effects of payment systems and vertical integration. *Health Economics* 10 (2): 119–134.

Ettner, S.L. 1993. Do elderly Medicaid patients experience reduced access to nursing home care? *Journal of Health Economics* 12 (3): 259–280.

Hotelling, H. 1929. Stability in competition. *Economic Journal* 39: 41–57.

Ma, C.A. 1994. Health care payment systems: Cost and quality incentives. *Journal of Economic Manage Strategy* 3: 93–112.

Miura, I., and H. Tajika. 2015. Cooperation of medical and nursing care service and optimal bundled fee. *Kyushu University Discussion Paper Series* No. 2015-3. (in Japanese).

Mougeot, M., and F. Naegelen. 2005. Hospital price regulation and expenditure cap policy. *Journal of Health Economics* 24 (1): 55–72.

Sanjo, Y. 2009. Quality choice in a health care market: A mixed duopoly approach. *The European Journal of Health Economics* 10 (2): 207–215.

Chapter 16
Towards Deeper Integration Among China, Japan and Korea

Bo-Young Choi and Seo Young Lee

16.1 Introduction

While the negotiations of the World Trade Organization (WTO) have not shown much progress, preferential trade agreements (PTAs) have become the main trade policy ro pursue liberalization in international trade for decades. Modern PTAs have evolved, including not only provisions to liberalize trade in goods and services, but also provisions to harmonize the regulatory environment among partner countries. Given the low tariff barriers and the prevalence of multinationals, PTAs' role of regulatory coherence and harmonization have become more important than before.

Because trade costs are accumulated as parts and components cross borders multiple times, more attention has been given to the role of deep agreements in economies where production networks are prevalent (e.g. Orefice and Rocha 2014; Kim 2012, 2015a; Osnago et al. 2016).[1] In this context, harmonization in the behind-the-border measures have become especially relevant for economies in East Asia, as these economies are interconnected with regional supply chains.

Modern PTAs include deep integration provisions to increase compatibility across different member countries.[2] These deep integration PTAs go beyond the existing WTO agreements with broader coverage of policy areas. For instance, the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) includes

---

[1] See Balassa (1961) for discussions on varying degrees of integration.

[2] Kim (2015a) characterizes provisions that promote deep integration as follows: protection of foreign firms' economic interests; liberalization of trade, especially 'beyond-the-border' barriers that go beyond traditional border barriers; and harmonization of trade rules across partner countries.

B.-Y. Choi (✉)
Kyungpook National University, Daegu, South Korea
e-mail: bychoi2@knu.ac.kr

S. Y. Lee
Korea Institute of Origin Information,Seoul, South Korea

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_16

labor, environment, electronic commerce provisions. Also, the United States-Mexico Canada (USMCA) trade agreement includes labor provisions that requires wage convergence between member countries to some extent. Such provisions deal with areas that were not included in traditional PTAs.

In terms of PTAs in Northeast Asia, there have been 14 rounds of negotiations of the trilateral China-Japan-Korea (CJK) FTA since 2010 but little progress has been made. Japan-Korea FTA talks have been suspended since November 2004. Although the free trade agreement (FTA) signed by Korea and China entered into effect in 2015, some argue that the Korea-China FTA has not been satisfactory enough in terms of the depth of liberalization and the scope of obligations.[3] Nonetheless, China, Japan and Korea are important trade partners to each other, and thus it is worthwhile to analyze and compare the PTAs of China, Japan and Korea to move further toward achieving deeper integration in the region.

This paper provides a comprehensive analysis for various issues other than tariff liberalization dealt in PTAs signed by China, Japan and Korea. We examine how the PTAs of China, Japan and Korea evolved through time, and how each country's approach is different in various trade-related issues. We analyze the PTAs of China, Japan and Korea based on Horn et al. (2010). Horn et al. (2010) classifes the PTA provisions into two groups: 14 WTO+ areas (provisions reconfirming the existing commitments in the WTO agreement and providing for additional obligations) and 38 WTO-X areas (provisions in areas outside the mandate of WTO). All three countries' PTAs show an extensive coverage of WTO+ areas; Korea, Japan and China each include 12.9, 11.4 and 10 areas per PTA in average. On the other hand, all three countries show less coverage in WTO-X areas, which are new policy areas that are not covered in the WTO agreement. Out of 38 areas, only 11.1, 9.4 and 7.9 areas were included for each Korea, Japan and China in average. While we examine the overall coverage of areas, we also analyze whether the included provisions in each area are legally enforceable. In order to measure the extent whether the PTAs are deep agreements with legal enforceability, we also take into account legal inflation, defined as the share of the total number of non-legally enforceable provisions out of the total number of covered provisions. Like most of the countries, Korea, Japan and China's PTAs exhibit a higher legal inflation rate in WTO-X areas than in WTO+ areas. Among the three countries, Korea and Japan's PTAs exhibit strong law enforcements for provisions in their PTAs while the legal inflation rate is relatively higher in China's PTAs. However, the legal inflation rate drops notably since its first PTA with a developed country, New Zealand, reaching to a similar level to Korea and Japan's PTAs.

This paper is organized as follows. Section 16.2 reviews the literature related to this study. Section 16.3 presents the methodology of the study including the assessment of coverage and legal enforceability of provisions. Section 16.4 analyzes the PTAs of the three countries with third countries. Section 16.5 concludes.

[3] Schott et al. (2015), p. 1.

16.2 Literature Review

With the prevalence of the multinational firms and production networks, much attention has been given to the relation between PTAs and trade. These works especially scrutinized the contents of the PTAs, measuring the depth of PTAs. Here 'deep integration' refers to a process that allows the convergnce of regulatory heterogeneity across different economies to facilitate trade. Thus, deep integration commitments not only includes market access for goods, but also includes other commitments related to issues such as competition, intellectual property rights and labor and environment standards. Kim (2015b) calculates the depth of integration score of agreements in Asia, which is essentially a count of all explicit provisions under items of each of these categories included in the PTA. She finds that the long-term pattern of depth of integration in PTAs shows an upward trend, although the pattern varies widely in the short term. As for country average, the scores were highest for Australia, Singapore and Japan, while Korea and China ranked 14th and 17th out of the 23 countries examined.

Horn et al. (2010) further classifies the provisions into two areas, WTO+ and WTO-X, and examines whether the provisions are legally enforceable. They specifically compare PTAs of the two representative rule setters, the US and the EU. They find that while the EU PTAs include four times as many WTO-X provisions as the US PTAs, many of the provisions appear not to be legally enforceable. Based on the classification of areas that PTAs cover, and the judgment of the provisions' legal enforceability proposed by Horn et al. (2010), the WTO constructed a database that contains an analysis of the world's PTAs. The database is also used in many quantitative studies to assess the effect of the depth of PTAs in not only trade in goods but also in migration. Orefice and Rocha (2014) find that deeper agreements tend to increase network trade amongst member countries. They also find that higher levels of network trade relative to total trade increase the likelihood of signing deeper agreements. Orefice (2015) estimates a modified gravity model to assess the role of PTAs and their content on bilateral migration flows. He not only finds a strong positive relationship between PTAs and migration flows, but also finds that the content of PTAs matters; visa, asylum and labor market-related provisions further stimulate bilateral migration flows.

16.3 Terminology and Methodology

16.3.1 Definition of PTAs

PTAs commonly refer to the regional trade agreements (RTAs) in the WTO, using that term because a large number of agreements are no longer in geographical proximity, and are not limited to countries within a single region[4] (e.g. Korea-Chile Free Trade Agreement). Therefore, this study uses the more generic term PTAs, instead of the term RTAs.[5]

With a series of exceptions for PTAs in the WTO agreements, WTO members are permitted to derogate from WTO rules in order to establish PTAs under certain conditions. The PTA exceptions are allowed by either Article XXIV of the GATT 1994 or paragraph 2 (c) of the Enabling Clause, which apply to PTA provisions relating to goods. The exception for PTAs is also contained in Article V (7) of the GATS, which applies to PTA provisions relating to services.[6]

In this study, PTAs notified under the Enabling Clause are included along with others notified under GATT Article XXIV and GATS Article V. Among all of the agreements covered in this study, only the Korea-ASEAN FTA, Korea-India CEPA are notified to the WTO under GATT Article XXIV and GATS Article V, as well as the Enabling Clause.[7] Also, China-ASEAN FTA is notified to the WTO under the Enabling Clause and GATS.

Horn et al. (2010) only cover PTAs concluded by WTO members under GATT Article XXIV or GATS Article V, and do not take into account agreements notified under the Enabling Clause. Since the Enabling Clause can only be used when all parties of the PTA are developing countries, these agreements are not related to the EU and US PTAs. Therefore, all the PTAs considered in Horn et al. (2010) are FTAs, except for the EU-Turkey agreement, which is a customs union. In contrast, World Trade Organization (2011) includes PTAs notified under the Enabling Clause, together with other PTAs notified under GATT Article XXIV and GATS Article V.

16.3.2 Measure of Deep Trade Agreements

Horn et al. (2010) suggest a method for qualitative analysis for examining the depth of PTAs. Horn et al. (2010) classify all the provisions of 14 EU PTAs and 14 US PTAs with third countries, notified to the WTO as of 2008, into 52 policy areas based on the chapter and heading of each PTA. World Trade Organization (2011) extends further from Horn et al. (2010), covering 96 PTAs entered into force during the period from 1958 through 2010, including most regions of the world. The World Trade Organization (2011) analysis is slightly different from the Horn et al. (2010) approach, in that it also covers several provisions that are not mentioned as an article or chapter heading but expressed in the context of other provisions.

Horn et al. (2010) classify 52 policy areas into two groups. The first policy area, referred to as WTO Plus (+) provisions, comes under the current mandate of the

[4]World Trade Organization (2011), p. 44; Mitchell and Lockheart (2016), p. 81; Matsushita et al. (2015), p. 507.

[5]World Trade Organization (2011), p. 44; Matsushita et al. (2015), p. 507.

[6]Mitchell and Lockheart (2016), p. 82.

[7]WTO's Regional Trade Agreements Information System (RTA-IS).

Table 16.1 List of 52 policy areas

WTO Plus (+) areas	WTO-Extra (X) areas	
FTA industrial	Anti-corruption	Cultural cooperation political dialogue
FTA agriculture	Competition policy	Economic policy dialogue public administration
Customs	Environmental laws	Education and training regional cooperation
Export taxes	IPR (Intellectual Property Rights)	Energy research and technology
SPS (Sanitary and Phytosanitary Measures)	Investment	Financial assistance SME (Small and Medium-sized Enterprises)
TBT (Technical Barriers to Trade)	Labour market regulation	Health social matters
STE (State Trading Enterprise)	Movement of capital	Human rights statistics
AD (Anti-Dumping)	Consumer protection	Illegal immigration taxation
CVM (Countervailing Measures)	Data protection	Illicit drugs terrorism
State aid	Agriculture	Industrial cooperation visa and asylum
Public procurement	Approximation of legislation	Information society
TRIMs (Trade Related Investment Measures)	Audio visual	Mining
GATS (General Agreement on Trade in Service)	Civil protection	Money laundering
TRIPs (Trade-related Aspects of Intellectual Property Rights)	Innovation policies	Nuclear safety

Note See Appendix 1 for detailed description of provisions
Source World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements)

WTO, where the parties undertake commitments at the bilateral level going beyond those they have accepted at the multilateral level (e.g. tariff reduction). The WTO+ provisions of PTAs can reconfirm existing commitments in the WTO agreement and provide for additional obligations. The second area, referred to as WTO-Extra (X) provisions, covers obligations lying outside the current WTO mandate (e.g. labor standards).[8] Table 16.1 lists both WTO+ areas and WTO-X areas, and Appendix provides detailed descriptions of each provision in both areas. Based on this classification, Horn et al. (2010) examine the coverage of WTO+ and WTO-X areas in each agreement.

[8]Several policy areas are divided into two groups, WTO+ and WTO-X. For example, TRIPs in WTO+ includes anything reconfirming commitments in the WTO, but IPR in WTO-X includes anything beyond commitments in TRIPs. See Appendix for the description of the 52 policy areas.

The standard judgment of legal enforceability in Horn et al. (2010) focuses on the explicit statement containing actual terminology of a provision that leads to the enforcement of dispute settlement proceedings under the PTA.[9] In other words, a provision that contains clearly defined legal obligations, likely to be invoked by a complainant in dispute settlement proceedings, is classified as legally enforceable.[10] Whether or not the provisions include legally enforceable obligations depends on the treaty interpretation. Since PTAs are recognized as treaties,[11] PTAs have to be interpreted in accordance with traditional treaty interpretation, stemming from the Vienna Convention on the Law of Treaties of 1969. The most commonly exemplified terms implying legally enforceable obligations are "shall," "agree," "undertake" and the like.[12] The WTO Appellate Body has also found the terms "shall" and"should" to be in accordance with the principle of effectiveness in treaty interpretation.[13] The following statements are examples of implications of legally enforceable obligations: "Each Party shall eliminate its customs duties on originating goods of the other Party… (Korea-EU FTA)," "A Party is hereby committed to allow such movement of capital… (Japan-Switzerland FTA)."

Meanwhile, if the provision is explicitly excluded by dispute settlement under PTAs, it is considered as a non-legally enforceable provision. In addition, vague expressions, which are difficult to be invoked by a complainant in a dispute settlement proceeding, are also considered as non-legally enforceable. For example, the following statements contain vague expressions for binding parties: "The Parties shall cooperate and exchange information with each other… (Japan-India CEPA)," "The Parties shall enhance their communication and cooperation on labor and environment matters… (China-New Zealand FTA)."

In addition to Horn et al. (2010), World Trade Organization (2011) explained that other factors, such as political factors and having recourse to WTO dispute settlement procedures, should be taken into consideration when classifying the legally enforceable obligations.[14] Therefore, the World Trade Organization (2011) considered an area as legally enforceable if the language is sufficiently precise to provide a legally enforceable obligation and the dispute settlement under the PTA has been excluded. Based on this approach measuring the depth of PTAs, the WTO and the World Bank

[9] See p. 1572 of Horn et al. (2010) for a detailed explanation on Legal Enforceability.

[10] World Trade Organization (2011), pp. 128–130.

[11] In Article 2.1.(a) of the Vienna Convention on the Law of Treaties (1969), a "Treaty" is defined as an international agreement concluded between States in written form and governed by international law, whether embodied in a single instrument or in two or more related instruments and whatever its particular designation.

[12] McCaffrey (2006), p. 82.

[13] World Trade Organization (2011), p. 129.

[14] According to World Trade Organization (2011) pp. 129–130, legally enforceable provisions with dispute settlement proceedings should take into consideration other factors that might be important for the actual enforcement of a provision in practice. For example, obligations under the WTO agreements may have an effect on the legal enforceability of obligations under PTAs. Also, obligations limiting the use of trade counter-measures under the WTO may have an effect on the enforceability of PTA provisions, and there are non-legal considerations such as external political factors.

publish data for PTAs signed by countries. While Horn et al. (2010) only analyzed the EU and US PTAs, the WTO expands the analysis to the world, covering PTAs signed in the period of 1958–2010. Further expanding the existing dataset, the World Bank recently opened a new dataset on [15] [16] 279 PTAs signed by 189 countries that entered into force between January 1958 and August 2015. We examine 11 of Korea's, 11 of China's and 14 of Japan's PTAs which entered into force between November 2002 and January 2015, based on the World Bank database (Table 16.2).

16.4 WTO+ and WTO-X Areas in PTAs of China, Japan and Korea

In this section, we analyze and compare the coverage and legal enforceability of the WTO+ and WTO-X provisions in PTAs. In Sect. 16.4.1, we review the overall trend of the world's PTAs in terms of its coverage and legal enforceability of WTO+ and WTO-X provisions. In Sect. 16.4.2–16.4.5, we analyze the PTAs signed by China, Japan, and Korea with third countries.

16.4.1 WTO+ and WTO-X Areas in PTAs Worldwide

Figure 16.1 shows the average number of WTO+ and WTO-X provisions covered by 279 PTAs notified to the WTO and in force between 1958 and 2015. The heading AC represents "Area Covered," while the heading LE represents "Legally Enforceable."[17] For both WTO+ provisions and WTO-X provisions, the average number decreased over time until the 1990s. The average number of WTO+ provisions, however, rose sharply from 6.6 in the 1990s to 10.6 in the 2000s, and the number further jumped to 11.5 for PTAs signed during 2010–2015. Similarly, the average number of WTO-X provisions increased from 4.6 in the 1990s to 7.9 in the 2000s, and to 10.8 in the early 2010s. Meanwhile, most of the provisions are legally enforceable in the WTO+ area, while only half of the provisions in covered areas are legally enforceable for WTO-X.

Figure 16.2 presents the number of PTAs for each specific WTO+ area. All of the 279 agreements contain provisions relating to tariffs on industrial goods and agriculture goods, with an exception for the ASEAN FTA (AFTA) in agriculture tariff elimination. Provisions regarding customs and export taxes are frequently included in PTAs. Non-tariff measures such as anti-dumping, TBT and SPS are included in many PTAs but a smaller share of provisions appear to have legal enforceability.

[15] World Bank dataset, content of Deep Trade Agreements, http://data.worldbank.org/data-catalog/deep-trade-agreements (accessed on October 30, 2016).

[16] Hofmann et al. (2017).

[17] Horn et al. (2010), p. 1573, p. 1575.

Table 16.2 List of China's, Japan's, and Korea's PTAs

Korea Agreement	Date of Signature by Parties	Date of Entry into force of agreement	GATT Art. XXIV	GATS Art V	Enabling Clause
Chile	01/02/2003	01/04/2004	08/04/2004	08/04/2004	Not applicable
Singapore	04/08/2005	02/03/2006	21/02/2006	21/02/2006	Not applicable
EFTA	15/12/2005	01/09/2006	23/08/2006	23/08/2006	Not applicable
ASEAN	24/08/2006(G) 21/11/2008(S)	01/01/2010(G) 14/10/2010(S)	applicable	applicable	applicable
India	07/08/2009	01/01/2010	23/06/2010	23/06/2010	applicable
EU	06/10/2010	01/07/2011	07/07/2011	07/07/2011	Not applicable
Peru	21/03/2011	01/08/2011	09/08/2011	09/08/2011	Not applicable
US	30/06/2007	15/03/2012	15/03/2012	15/03/2012	Not applicable
Turkey	01/08/2012	01/05/2013	30/04/2013	30/04/2013	Not applicable
Australia	08/04/2014	12/12/2014	22/12/2014	22/12/2014	Not applicable
Canada	22/09/2014	01/01/2015	20/01/2015	20/01/2015	Not applicable

China Agreement	Date of Signature by Parties	Date of Entry into force of agreement	GATT Art. XXIV	GATS Art V	Enabling Clause
Hong Kong	29/06/2003	29/06/2003	27/12/2003	27/12/2003	Not applicable
Macao	17/10/2003	17/10/2003	27/12/2003	27/12/2003	Not applicable
ASEAN	29/11/2004(G) 14/01/2007(S)	01/01/2005(G) 01/07/2007(S)	Not applicable	26/06/2008	21/09/2005
Chile	18/11/2005(G) 13/04/2008(S)	01/10/2006(G) 01/08/2010(S)	20/01/2007	18/11/2010	Not applicable
Pakistan	24/11/2006(G) 21/02/2009(S)	01/07/2007(G) 10/10/2009(S)	18/01/2008	20/05/2010	Not applicable
New Zealand	07/04/2008	01/10/2008	21/04/2009	21/04/2009	Not applicable
Singapore	23/10/2008	01/01/2009	02/03/2009	02/03/2009	Not applicable
Peru	28/04/2009	01/03/2010	03/03/2010	03/03/2010	Not applicable
Costa Rica	08/04/2010	01/08/2011	27/02/2012	27/02/2012	Not applicable
Iceland	15/04/2013	01/07/2014	10/10/2014	10/10/2014	Not applicable
Switzerland	06/07/2013	01/07/2014	30/06/2014	30/06/2014	Not applicable

Japan Agreement	Date of Signature by Parties	Date of Entry into force of agreement	GATT Art. XXIV	GATS Art V	Enabling Clause
Singapore	13/01/2002	30/11/2002	08/11/2002	08/11/2002	Not applicable
Mexico	17/09/2004	01/04/2005	31/03/2005	31/03/2005	Not applicable
Malaysia	13/12/2005	13/07/2006	12/07/2006	12/07/2006	Not applicable
Chile	27/03/2007	03/09/2007	24/08/2007	24/08/2007	Not applicable
Thailand	03/04/2007	01/11/2007	25/10/2007	25/10/2007	Not applicable
Indonesia	20/08/2007	01/07/2008	27/06/2008	27/06/2008	Not applicable
Brunei	18/06/2007	31/07/2008	31/07/2008	31/07/2008	Not applicable
ASEAN	26/03/2008	01/12/2008	23/11/2009	Not applicable	Not applicable
Philippines	09/09/2006	11/12/2008	11/12/2008	11/12/2008	Not applicable
Switzerland	19/02/2009	01/09/2009	01/09/2009	01/09/2009	Not applicable
Viet Nam	25/12/2008	01/10/2009	01/10/2009	01/10/2009	Not applicable
India	16/02/2011	01/08/2011	14/09/2011	14/09/2011	Not applicable
Peru	31/05/2011	01/03/2012	24/02/2012	24/02/2012	Not applicable
Australia	08/07/2014	15/01/2015	12/01/2015	12/01/2015	Not applicable

Note (1) Korea's, China's and Japan's PTAs with other WTO Members, entered into force as of January 2015
(2) (G) applies to PTA provisions relating to goods, while (S) applies to PTA provisions relating to services
(3) Both the Korea-ASEAN FTA and Korea-India CEPA are notified under GATT XXIV, the Enabling Clause and GATS V. The dates of notification are not indicated in the WTO RTA-IS, but in the ADB Asia Regional Integration Center (ARIC)

Source World Trade Organization database, "Regional Trade Agreements Information System (RTA-IS)," http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx (accessed on October 30, 2016)

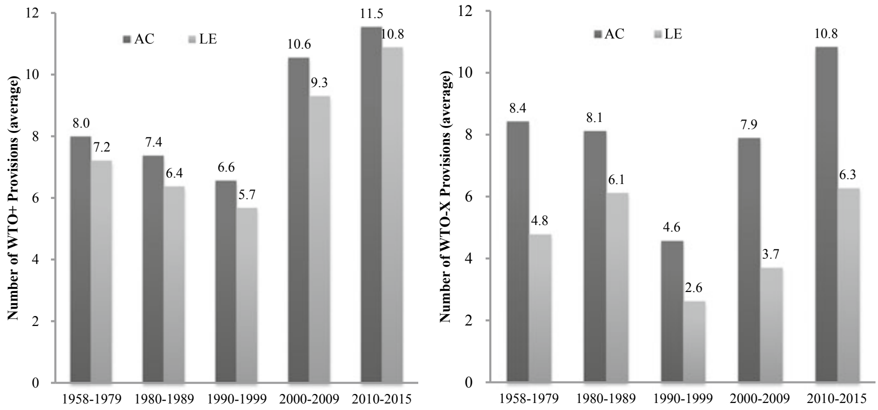


Fig. 16.1 WTO+ and WTO-X areas in PTAs worldwide, over time
Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Figure 16.3 shows the number of PTAs including specific WTO-X areas. Although there are a greater variety of WTO-X provisions than WTO+ provisions, substantially less WTO-X provisions are included in PTAs compared to WTO+ provisions. Competition policy is the only exception; 208 agreements out of 279 include competition policy-related provisions. Legal enforceability also varies by area. While about 88% of provisions in competition policy, 93% of movement of capital and 83% of IPR provisions are legally enforceable, only half of environmental law provisions appear to be so.

16.4.2 WTO+ Areas for PTAs of China, Japan, and Korea

Tables 16.3, 16.4 and 16.4 each present the coverage and legal enforceability of WTO+ provisions in PTAs of China, Japan and Korea. As in the previous analysis, the heading AC indicates "Area Covered," while the heading LE represents "Legally Enforceable." In the AC column, a dark box indicates policy areas covered in the agreement, while the white box indicates policy areas are not covered.[18] In the LE column, a dark box indicates that the language is sufficiently precise or commits to providing a legally enforceable obligation,[19] while a white box indicates areas that are legally non-enforceable due to imprecise language. The cross-hatched box

[18]Horn et al. (2010), p. 1573.
[19]Horn et al. (2010), p. 1575.

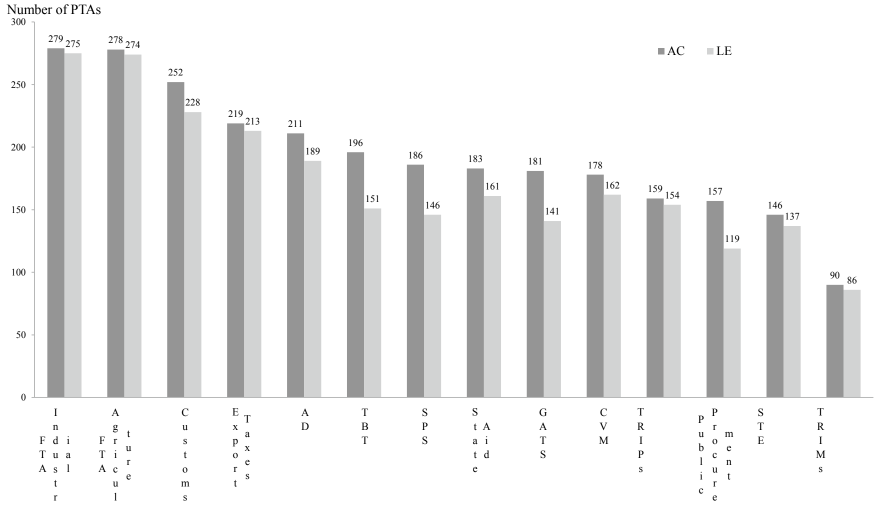


Fig. 16.2 Number of PTAs worldwide covering WTO+ areas
Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

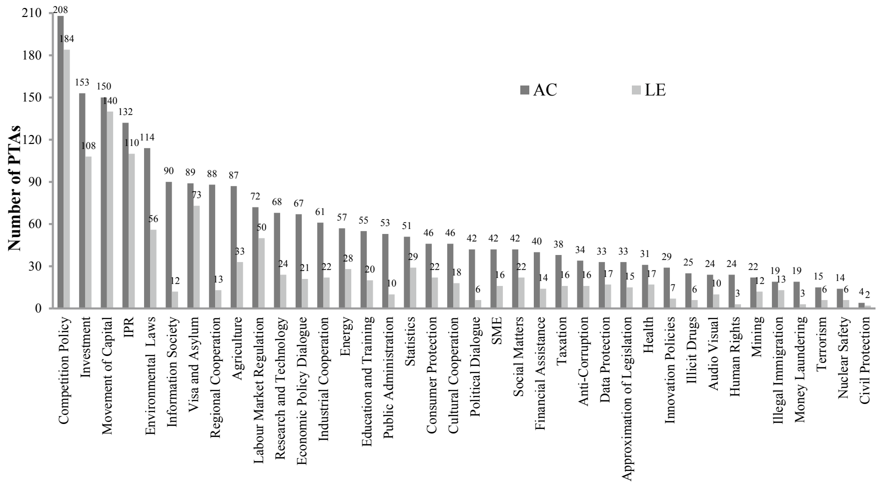


Fig. 16.3 Number of PTAs worldwide covering WTO-X areas
Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

indicates that the language is sufficiently precise or commits to providing a legally enforceable obligation, but dispute settlement is not available.[20]

We find extensive coverage for WTO+ provisions in all agreements of China, Japan and Korea. Among the three countries, Korea's PTAs have the most extensive coverage of WTO+ areas. Korea's PTAs include 12.9 areas per PTA, while Japan's PTAs and China's PTAs include 11.4 and 10 areas each per PTA.

There are two WTO+ areas for which all three countries' agreements contain provisions: tariff liberalization on industrial goods and agriculture goods. This is not surprising, in that tariff liberalization is generally the main purpose of PTAs.

Provisions that are covered in all PTAs of Korea are Customs, Sanitary and Phytosanitary Measures (SPS), Technical Barriers to Trade (TBT), Anti- Dumping (AD), Countervailing Measures (CVM), General Agreement on Trade in Service (GATS) and Trade-related aspects of Intellectual Property Rights (TRIPs). Note that a majority of non-tariff measures provisions, SPS, TBT, AD, and CVM are not only included but also legally enforceable. In contrast, state aid is included relatively less than other WTO+ provisions. Examining the PTAs across partner countries, we find that FTAs with Chile and India cover all 14 WTO+ provisions, and 13 of those are legally enforceable.[21] Korea's PTAs with ASEAN, Turkey and Singapore include less WTO+ provisions. However, even in the case of these PTAs, 12 out of 14 WTO+ provisions are included and most of them are legally enforceable.

Japan's PTAs include less WTO+ provisions than Korea. In addition to FTA industrial and FTA agriculture provisions, GATS and customs are covered and legally enforceable in all PTAs. TRIPs is also covered and legally enforceable in all PTAs except for the Japan-ASEAN FTA. However, when looking at Japan's PTAs with individual ASEAN member states, we observe that the TRIPs provision is included and legally enforceable in all cases.[22] The least covered and legally enforceable provisions are state aid and state trading enterprises (STEs). Examining the PTAs across partner countries, we find that there are eight PTAs with ASEAN and members of ASEAN, three PTAs with Latin America and two PTAs with developed countries. Japan tends to include less WTO+ provisions for PTAs with ASEAN member states, while relatively more WTO+ provisions are included for PTAs with Chile and Australia.

For China, anti-dumping (AD) and countervailing measures (CVM) are commonly included in each country's PTA. These two provisions are covered and legally enforceable in all of China's 11 PTAs, with one exception; the AD provision is not legally enforceable in the China-Macao PTA. Considering that China has become

[20]Horn et al. (2010) consider the cross-hatched box as non-enforceable, due to the exemption from dispute settlement proceedings.

[21]Although the Korea-India CEPA (Comprehensive Economic Partnership Agreement) is widely known for its low tariff concession levels compared to other FTAs signed by Korea, the agreement covers diverse areas with legal enforcement as shown in Tables 16.3 and 16.6.

[22]The least developed countries (LDCs), such as Myanmar, Cambodia, and Laos, are given an extended transition period to protect intellectual property under the WTO's Agreement on TRIPs. "Responding to least developed countries' special needs in intellectual property," https://www.wto.org/english/tratop_e/trips_e/ldc_e.htm (accessed on November 11, 2016).

Table 16.3 Classification of WTO+ areas in Japan's agreements

	Korea-Chile		Korea-Singapore		Korea-EFTA		Korea-ASEAN		Korea-India		Korea-EU		Korea-Peru		Korea-US		Korea-Turkey		Korea-Australia		Korea-Canada		AC	LE
	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE		
FTA Ind.																							11	11
FTA Agr.																							11	11
Customs																							11	10
Export Taxes																							10	10
SPS																							11	11
TBT																							11	11
STE																							9	9
AD																							11	11
CVM																							11	11
State Aid																							7	6
Public Procurement																							9	8
TRIMs																							8	8
GATS																							11	10
TRIPs																							11	11
WTO+ Total	14	13	12	12	13	13	12	11	14	13	13	13	13	13	13	13	12	11	13	13	13	13	142	138

Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Table 16.4 Classification of WTO+ areas in Japan's agreements

	Japan-Singapore		Japan-Mexico		Japan-Malaysia		Japan-Chile		Japan-Thai Land		Japan-Indonesia		Japan-Brunei		Japan-ASEAN		Japan-Philippines		Japan-Switzerland		Japan-Viet Nam		Japan-India		Japan - Peru		Japan - Australia		AC	LE
	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE		
FTA Ind.	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	14	14
FTA Agr.	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	14	14
Customs	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	14	14
Export Taxes	■	■	■	■			■	■					■	■			■		■	■			■	■	■	■	■	■	9	8
SPS			■	▨	■	■	■	▨							■	■			■	▨	■	▨	■	▨	■	▨	■	▨	9	9
TBT	■	■	■	▨	■	■	■	▨					■		■	■			■	▨	■	▨	■	▨	■	▨	■	▨	11	10
STE	■	■					■	■					■				■	■	■	■	■	■					■	▨	7	6
AD	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■			■	■	■	■			■	■	12	11
CVM	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■			■	■	■	■			■	■	12	11
State Aid							■	■					■	■					■	■	■	■	■		■	■	■	■	7	6
Public Procurement	■		■	■			■	■	■	■	■	■	■	▨			■		■	▨	■	▨	■	■	■	■	■	■	12	10
TRIMs	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■			■	■	12	12
GATS	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	▨	14	14
TRIPs	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■	■	■	■	■	■	■	■	■	■	13	13
WTO+ Total	12	11	12	12	10	10	14	14	9	9	9	9	13	9	9	9	11	9	12	12	12	12	13	12	10	10	14	14	160	152

Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

a major target of trade remedy measures by trading partners, this is not surprising. China's anti-dumping problem is related to the country's Market Economy Status (MES) under the WTO accession protocol.[23] In order to mitigate antidumping problems, China has made market economy status a prerequisite for launching FTA talks.

Currently, China is not a member of the WTO Government Procurement Agreement (GPA), and it tends not to be actively involved in FTA negotiations regarding government procurement.[24] China's most recent FTAs with Switzerland and Iceland are an exception. The China-Switzerland and China- Iceland FTAs have reached consensus on government procurement, which China seldom met in previous FTA negotiations. In fact, China's FTAs with Switzerland and Iceland touch on many new rules and subjects including government procurement, demonstrating China's confidence toward building high-standard FTAs. While China mostly concluded PTAs with developing countries in the earlier stages, more recently, China's FTA strategy has changed, as it started to sign PTAs with developed countries. Since 2008, starting with the China-New Zealand FTA, China has included more WTO+ provisions in its PTAs with developed countries.

In addition to coverage of provisions, Horn et al. (2010) define the legal inflation rate as the share of the number of total non-legally enforceable provisions out of the number of total covered provisions. Calculating the legal inflation rates for WTO+ provisions, we find that the degree of legal inflation is relatively low in the three countries' agreements: 2.8% for Korea (138 out of 142), followed by 5.0% for Japan (152 out of 160), and 8.2% (101 out of 110) for China. Korea's and Japan's PTAs exhibit strong law enforcement for their provisions. China, on the other hand, had a relatively lower rate of legally enforceable obligations until the mid-2000s, but after signing agreements with New Zealand, it reached a level similar to that of Japan and Korea (Table 16.5).

16.4.3 WTO-X Areas for PTAs of China, Japan, and Korea

While there are a wider variety of WTO-X areas (38 areas) than WTO+ (14 areas), less WTO-X provisions are included in the PTAs of Korea, Japan and China. Tables 16.6, 16.7 and 16.8 show the coverage and legal enforceability of the WTO-X provisions for each PTA of Korea, Japan and China. With respect to the coverage of WTO-X provisions, Korea includes 11.1 areas per PTA, while Japan includes 9.4 and China includes 7.9.

The two areas with the largest number of legally enforceable provisions among the 38 WTO-X areas are investment and intellectual property rights (IPRs) in all three countries' agreements. All PTAs include investment provisions except for the

[23] Salidjanova, N. (May 28, 2015). "China agreed to be treated as a non-market economy for 15 years after China's WTO accession."

[24] Public procurement is generally referred to as government procurement.

Table 16.5 Classification of WTO+ areas in China's agreements

	China-Hong Kong		China - Macao		China-ASEAN		China-Chile		China-Pakistan		China-New Zealand		China-Singapore		China-Peru		China - Costa Rica		China-Iceland		China-Switzer land		AC	LE
	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE		
FTA Ind.	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	11	11
FTA Agr.	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	11	11
Customs			■	■			■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	9	9
Export Taxes							■	■	■	■	■	■			■	■	■	■	■	■	■		7	6
SPS			■	■					■	■	■	■	■	■	■	■	■	■	■	■	■	■	8	8
TBT							■		■	■	■	■	■	■	■	■	■	■	■	■	■	■	8	7
STE							■				■	■	■	■	■	■					■	■	5	4
AD	■	■	■		■	■	■	■	■	▨	■	■	■	■	■	■	■	▨	■	▨	■	■	11	10
CVM	■	■	■	■	■	■	■	■	■	▨	■	■	■	■	■	■	■	▨	■	▨	■	■	11	11
State Aid			■				■	■	■	▨	■	■					■	■	■	■	■		7	5
Public Procurement																			■	■	■	▨	2	2
TRIMs					■		■				■	■	■	■	■	■							5	3
GATS	■	■	■	■	■						■	■	■	■	■	■	■	■	■	■	■	■	9	8
TRIPs							■	■			■	■			■	■	■	■	■	■	■	■	6	6
WTO+ Total	5	5	8	6	6	4	11	8	9	9	13	13	10	10	12	12	11	11	12	12	13	11	110	101

Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Table 16.6 Classification of WTO-X areas in Korea's agreements

	Korea-Chile		Korea-Singapore		Korea-EFTA		Korea-ASEAN		Korea-India		Korea-EU		Korea-Peru		Korea-US		Korea-Turkey		Korea-Australia		Korea-Canada			
	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE
Anti Corruption															■						■		2	0
Competition Policy	■	▨	■	■	■	▨			■	▨	■	▨	■	▨	■	▨	■	▨	■	▨	■	▨	10	10
Environmental Laws							■				■	▨	■	▨	■	▨	■	▨	■	▨	■	▨	7	6
IPR	■	■	■	■	■	▨	■	■	■	▨	■	■	■	■	■	■	■	■	■	■	■	■	11	11
Investment	■	■	■		■	▨	■		■		■	■	■	■	■	■			■	■	■	■	10	7
Labour Market Regulation											■	▨	■	▨	■	▨	■	▨	■	▨	■	▨	6	6
Movement of Capital	■	■	■	■	■	▨			■	■	■	■	■	■	■	■			■	■	■	■	9	9
Consumer Protection	■												■	▨	■	▨			■	▨	■		5	3
Data Protection													■						■		■		3	0
Agriculture							■	■					■	▨	■	■			■				4	3
Apporx. Of Legislation																							0	0
Audio Visual			■						■		■	▨	■	▨					■	▨	■	▨	6	4
Civil Protection																							0	0
Innovation Policies															■				■				2	0
Cultural Cooperation							■	■	■		■	▨	■	▨							■	▨	5	4
Economic Policy Dialogue													■	▨									1	1
Education and Training			■				■												■				3	0
Energy			■				■	■	■				■	▨					■				5	2
Financial Assistance																					■		1	0
Health									■						■	■							2	1
Human Rights																							0	0
Illegal Immigration																							0	0
Illicit Drugs																							0	0
Industrial Cooperation							■																1	0
Information Society	■	■									■		■	▨	■		■		■		■		7	2
Mining							■	■					■	▨									2	2
Money Laundering																							0	0
Nuclear Safety																							0	0
Political Dialogue													■	▨					■		■		3	1
Public Administration															■				■		■		3	0
Regional Cooperation																			■				1	0
Research and Technology			■				■	■	■				■	▨					■				5	2
SME							■	■	■				■	▨									3	2
Social Matters																							0	0
Statistics																							0	0
Taxation																							0	0
Terrorism																							0	0
Visa and Asylum	■	■	■	■					■	■									■	▨	■	▨	5	5
WTO-X Total	7	6	9	4	4	4	11	7	11	4	9	8	18	17	13	9	5	4	19	9	16	9	122	81

Note "Area Covered"; LE: "Legally Enforceable"
Source AC: Source: Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Table 16.7 Classification of WTO-X areas in Japan's agreements

	Japan-Singapore		Japan-Mexico		Japan-Malaysia		Japan-Chile		Japan-Thailand		Japan-Indonesia		Japan-Brunei		Japan-ASEAN		Japan-Philippines		Japan-Switzerland		Japan-Viet Nam		Japan-India		Japan-Peru		Japan–Australia		AC	LE
	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE		
Anti Corruption									■								■						■	■	■		■		5	1
Competition Policy	■	▨	■	▨	■	▨	■	▨	■		■	■			■	▨	■	▨	■	▨	■	▨	■	▨	■	▨	■	▨	13	12
Environmental Laws			■	▨							■		■		■	▨	■						■	■	■		■		8	3
IPR	■		■	▨	■	■			■	■			■	▨	■	▨	■	■	■	■	■	■			■	■	■		11	9
Investment	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■		■	■	13	12
Labour Market Regulation							■										■												2	0
Movement of Capital	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■	■	■	■	■	■	■	■	■	■	13	13
Consumer Protection																			■								■	▨	2	1
Data Protection							■																		■		■		3	0
Agriculture			■	▨									■	▨	■	▨									■				4	3
Apporx. of Legislation							■																						1	0
Audio Visual																													0	0
Civil Protection																													0	0
Innovation Policies																									■				1	0
Cultural Cooperation																													0	0
Economic Policy Dialogue																			■	▨							■	▨	2	2
Education and Training	■		■	▨									■	▨	■	▨									■				5	3
Energy											■		■	■	■	▨							■	▨			■	■	5	4
Financial Assistance													■	■													■	■	2	2
Health																							■	▨					1	1
Human Rights																													0	0
Illegal Immigration									■																				1	0
Illicit Drugs																									■				1	0
Industrial Cooperation													■	▨													■		2	1
Information Society													■		■	▨					■		■	▨	■		■	▨	6	3
Mining											■																		1	0
Money Laundering																													0	0
Nuclear Safety																													0	0
Political Dialogue																													0	0
Public Administration																											■		1	0
Regional Cooperation					■				■		■				■	▨							■	▨			■		6	2
Research and Technology	■		■	▨									■	▨											■				4	2
SME	■		■	▨					■				■	▨	■	▨							■	▨			■		7	4
Social Matters																							■						1	0
Statistics																													0	0
Taxation																			■	▨									1	1
Terrorism																													0	0
Visa and Asylum					■	■			■	■	■	■					■	■	■	■	■	■	■	▨	■	▨	■	▨	9	9
WTO-X Total	7	3	9	9	6	5	6	3	9	4	8	4	12	10	10	10	8	5	8	7	5	4	12	11	14	4	17	9	131	88

Note AC: "Area Covered"; LE:"Legally Enforceable"

Source Source: Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Table 16.8 Classification of WTO-X areas in China's agreements

	China-Hong Kong		China-Macao		China-ASEAN		China-Chile		China-Pakistan		China-New Zealand		China-Singapore		China-Peru		China - Costa Rica		China-Iceland		China-Switzer land		AC	LE
	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE	AC	LE		
Anti Corruption																							0	0
Competition Policy															■		■		■	▨	■	▨	4	2
Environmental Laws							■	▨			■	▨			■				■	▨	■	▨	5	4
IPR	■		■	■			■	▨	■	■	■	■					■	■	■	■	■	■	8	7
Investment	■		■	■	■		■	▨	■	■	■	■	■	■			■	■	■		■	■	10	7
Labour Market Regulation							■	▨			■	▨			■				■	▨			4	3
Movement of Capital															■	■	■	■	■	■			3	3
Consumer Protection											■	■											1	1
Data Protection																							0	0
Agriculture															■		■				■	▨	3	1
Apporx. of Legislation																							0	0
Audio Visual																	■						1	0
Civil Protection																							0	0
Innovation Policies																	■				■	▨	2	1
Cultural Cooperation	■						■	▨					■		■		■						5	1
Economic Policy Dialogue							■	▨											■	▨			2	2
Education and Training							■	▨					■		■		■		■	▨			5	2
Energy																							0	0
Financial Assistance							■																1	0
Health							■																1	0
Human Rights							■																1	0
Illegal Immigration																							0	0
Illicit Drugs																							0	0
Industrial Cooperation							■	▨					■	■	■						■	▨	4	3
Information Society											■	■					■		■				3	1
Mining							■	▨							■								2	1
Money Laundering																							0	0
Nuclear Safety							■																1	0
Political Dialogue																							0	0
Public Administration							■										■						2	0
Regional Cooperation													■	■	■		■						3	1
Research and Technology							■	▨							■		■		■	▨			4	2
SME							■	▨			■	▨			■		■						4	2
Social Matters							■	▨															1	1
Statistics																							0	0
Taxation							■																1	0
Terrorism							■																1	0
Visa and Asylum							■				■	■	■	■	■	■	■	▨					5	4
WTO-X Total	3	0	2	2	1	0	20	12	2	2	8	8	6	4	13	2	15	4	10	8	7	7	87	49

Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Korea-Turkey, Japan-Vietnam, and China-Peru PTAs.[25] As for IPRs, all of Korea's PTAs have included the provisions during the period of study, while there were some exceptions for China and Japan. China did not include IPR provisions in its PTAs with ASEAN, Singapore and Peru. Also, IPR provisions are not included in the Japan-Chile, Japan-Indonesia and Japan-India PTAs.

Both Korea and Japan actively include competition policy provisions, while China does not. For China, competition policy provisions only start appearing in PTAs from China's FTA with Peru, and then with Costa Rica. However, these provisions were not legally enforceable. More recent PTAs with Iceland and Switzerland include competition policy provisions which are legally enforceable, though they exclude dispute settlement. Provisions related to movement of capital are also included in the majority of Korea's and Japan's PTAs, while China includes such provisions only in PTAs with Peru, Costa Rica and Iceland. Movement of capital provisions are included in all of Korea's PTAs, except for those with ASEAN and Turkey. They are also included in all of Japan's PTAs, except for its PTA with ASEAN.

In Korea's PTAs, we find that Korea takes a strong position on regulatory issues in environmental laws and labor market regulation; not only are provisions related to these fields included in about two-thirds of all agreements, but they are also legally enforceable. Also, provisions related to promotion of the industry, such as audiovisual,[26] appear commonly in Korea's PTAs. Like Korea's PTAs, environmental laws provisions appear frequently in Japan's PTA, but most of them are legally non-enforceable. Japan also actively includes provisions related to regulatory issues, such as visa and asylum, as well as the promotion of industries, such as SME.

China contains relatively less WTO-X areas in its PTAs than Korea and Japan. However, China tends to cover a wider variety of new WTO-X areas, especially in recently signed agreements, most of which are related to developmental issues such as regional, cultural, industrial cooperation while including only a few regulatory issues such as environmental laws. Contrary to Korea's and Japan's PTAs, provisions related to movement of capital do not appear much in China's PTAs.

Overall, the remainder of WTO-X provisions account for a very small portion. Three WTO-X provisions, civil protection, money laundering and statistics, are not covered in at least one of the three countries' agreements. In addition, we find that each country has certain areas that are included in at least one of their PTAs, which are not included in any PTA of the other two counterparts: political dialogue for Korea, approximation of legislation, illegal immigration and illicit drugs for Japan, and human rights, nuclear safety and terrorism for China.

Examining the characteristics of PTAs varying by partner, we find that PTAs with Latin American countries, such as Peru or Chile, appear to prominently include WTO-X provisions. In contrast, plurilateral agreements signed by Korea, Japan or China

---

[25] It is interesting to observe that the Korea-EFTA FTA does not include the TRIMs provision, while it includes a legally enforceable investment provision (but without dispute settlement).

[26] Korea-Australia FTA. Article 7.12: Audiovisual Co-Production, Recognising that audiovisual, including film, animation and broadcasting program co-productions can significantly contribute to the development of the audiovisual industry and to the intensification of cultural and economic exchange between them. . . .

with three or more countries, such as FTAs with ASEAN or the EFTA, contained only a small number of WTO-X provisions.

While all three countries' agreements show a tendency to cover more of the new WTO-X areas in recently signed agreements, the number of legally enforceable WTO-X provisions contained in all three countries' agreements is still quite low. The legal inflation rate for WTO-X provisions is 43.7% for China's PTAs in average, and 32.8% each in the case of Korea's and Japan's PTAs.

16.4.4 The Balance Between WTO+ and WTO-X Areas in PTAs

In order to detect the center of gravity for the three countries' agreements in terms of coverage, we plot the number of covered WTO+ areas (vertical axis) against the number of covered WTO-X areas (horizontal axis) for each of the three countries' agreements in Fig. 16.4. The dots indicating the balance between WTO+ and WTO-X areas move back and forth depending on PTA partners. Figure 16.4 shows that the more recent PTAs are located in the northeast area (e.g. Korea-Canada (2015), Japan-Australia (2015), China- Costa Rica (2011)), but the earlier ones are located in the southwest or northwest areas (e.g. Korea-EFTA (2006), Japan-Malaysia (2006), and China- Hong Kong (2003)). In other words, each of the three countries' agreements tend to include more WTO+ and WTO-X provisions in more recent PTAs.

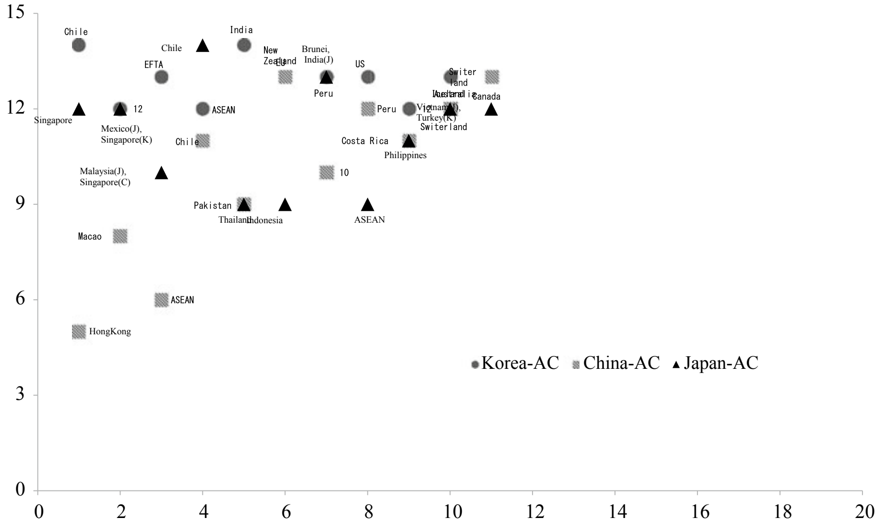


Fig. 16.4 The balance between WTO+ and WTO-X areas
Note AC: "Area Covered"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

The upward pattern is most pronounced in China's agreements. This may reflect China's increased confidence in liberalizing in more diverse areas based on its remarkable economic growth, and China's tendency to sign FTAs with more advanced economies. The agreements with Korea and Japan show a slight variation depending on their partner countries. While there has been a small increase in the number of WTO+ provisions, more recent PTAs covered more WTO-X areas in Korea's and Japan's agreements.

Figure 5 plots the balance between WTO+ and WTO-X provisions with legal enforceability. The figure shows a clearer upward tendency compared to Fig. 16.4; more recent PTAs are located in the northwest area (e.g. Korea- Canada (2015), Japan-Australia (2015), and China-Iceland (2014)). While the vertical distribution of the PTAs in Fig. 5 is similar to Fig. 16.4, the horizontal distribution becomes shorter, implying that many of the WTO+ provisions are legally enforceable, while the WTO-X provisions are not. The clearer upward tendency also indicates that more recent PTAs tend to include more legally enforceable provisions, instead of including provisions that are not legally binding.

16.4.5 Legal Inflation by Groups of Areas

We aggregate all WTO+ and WTO-X areas into five groups, depending on the nature of each area in accordance with WTO agreements, the classification of which is in parallel with Horn et al. (2010). As shown in Tables 16.9 and 16.10, Group 1 includes obligations that deal with policy instruments related to trade in goods and investment, which are applied at the border and regarded as the traditional mandate of WTO agreements; Group 2 includes GATS and TRIPs, which are considered as new WTO areas compared to GATT; Group 3 includes migration-related regulations; Group 4, domestic trade-related regulations, includes obligations related to behind-the-border regulations; Group 5 includes all remaining areas, mainly containing symbolic and declarative meanings for cooperation and development issues.[27]

The legal inflation rate of the five groups for Korea, Japan, China, the EU, and the US PTAs worldwide is shown in Tables 16.9 and 16.10. Horn et al. (2010) originally compare the legal inflation rate between the US and the EU, and find that the number of deeper provisions included in PTAs is larger for the EU, but they tend to be legally inflated. On the other hand, they find that US PTAs include less WTO+ or WTO-X provisions than the EU PTAs but they tend to be more legally enforceable than in the EU.

The inflation rate for all of China's, Japan's and Korea's PTAs appear to be low compared to the EU (34%) PTAs and the world's PTA average (27%). The inflation rate is 17% for Korea, 18% for Japan and 24% for China. However, compared to the 9% legal inflation rate of the US, the inflation rates are quite high.

[27] Horn et al. (2010), pp. 33–34.

Table 16.9 Details for legal inflation for groups of areas

	Korea PTAs			Japan PTAs			China PTAs			EU PTAs			US PTAs			World PTAs		
	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation
1. Trade and Investment related Obligations	92	88	4%	113	109	4%	78	71	9%	330	315	5%	103	103	0%	1,810	1,675	7%
2. GATS/TRIPs/IPR	33	32	3%	38	36	5%	23	21	9%	90	81	10%	38	37	3%	472	405	14%
3. Migration-related Regulations	5	5	0%	11	9	18%	6	5	17%	62	43	31%	2	1	50%	150	108	28%
4. Domestic Trade-related Regulations	80	70	13%	79	58	27%	44	36	18%	280	213	24%	97	88	9%	1,375	1,059	23%
5. Others	54	25	54%	50	28	44%	46	17	63%	478	171	64%	19	6	68%	1,125	369	67%
Total All Areas:	264	220	17%	291	240	18%	197	150	24%	1,240	823	34%	259	235	9%	4,932	3,616	27%

Note "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Table 16.10 Legal inflation by five groups of areas

	KoreaPTAs			Japan PTAs			China PTAs			EU PTAs			US PTAs			World PTAs		
	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation
Group1. Trade and Investment related Obligations																		
FTA Ind.	11	11	0%	14	14	0%	11	11	0%	46	46	0%	13	13	0%	279	275	1%
FTA Agr.	11	11	0%	14	14	0%	11	11	0%	46	46	0%	13	13	0%	278	274	1%
Customs	11	10	9%	14	14	0%	9	9	0%	45	44	2%	12	12	0%	252	228	10%
Export Taxes	10	10	0%	9	8	11%	7	6	14%	42	42	0%	12	12	0%	219	213	3%
AD	11	11	0%	12	11	8%	11	10	10%	45	45	0%	11	11	0%	211	189	10%
CVM	11	11	0%	12	11	8%	11	11	0%	39	39	0%	11	11	0%	178	162	9%
TRIMs	8	8	0%	12	12	0%	5	3	40%	10	9	10%	11	11	0%	90	86	4%
Investment	10	7	30%	13	12	8%	10	7	22%	26	14	46%	10	10	0%	153	108	29%
Movement of Capital	9	9	0%	13	13	0%	3	3	0%	31	30	3%	10	10	0%	150	140	7%
Total (1)	92	88	4%	113	109	4%	78	71	9.0%	330	315	5%	103	103	0	1810	1675	7%
Group 2. GATS/TRIPs																		
GATS	11	10	9%	14	14	0%	9	8	13%	24	20	17%	13	12	8%	181	141	22%
TRIPs	11	11	0%	13	13	0%	6	6	0%	35	34	3%	13	13	0%	159	154	3%
IPR	11	11	0%	11	9	18%	8	7	0%	31	27	13%	12	12	0%	132	110	17%
Total (2)	33	32	3%	38	36	5%	23	21	9%	90	81	10%	38	37	3%	472	405	14%
Group 3. Migration-related Regulations																		
Illegal Immigration	0	0	-	1	0	100%	0	0	-	18	13	28%	0	0	-	19	13	32%
Social Matters	0	0	-	1	0	100%	1	1	0%	28	18	36%	1	0	100%	42	22	48%
Visa andAsylum	5	5	0%	9	9	0%	5	4	20%	16	12	25%	1	1	0%	89	73	18%
Total (3)	5	5	0%	11	9	18%	6	5	17%	62	43	31%	2	1	50%	150	108	28%
Group 4. Domestic Trade-related Regulations																		
SPS	11	11	0%	9	9	0%	8	8	0%	25	18	28%	11	11	0%	186	146	22%
TBT	11	11	0%	11	10	9%	8	7	13%	30	18	40%	11	10	9%	196	151	23%
STE	9	9	0%	7	6	14%	5	4	20%	32	31	3%	9	7	22%	146	137	6%
State Aid	7	6	14%	7	6	14%	7	5	29%	36	36	0%	10	10	0%	183	161	12%
Public Procurement	9	8	11%	12	10	17%	2	2	0%	31	23	26%	13	12	8%	157	119	24%
Anti Corruption	2	0	100%	5	1	80%	0	0	-	6	3	50%	9	6	33%	34	16	53%
CompetitionPolicy	10	10	0%	13	12	8%	4	2	50%	40	39	3%	7	7	0%	208	184	12%
Environmental Laws	7	6	14%	8	3	63%	5	4	20%	32	15	53%	12	12	0%	114	56	51%
Labour Market Regulation	6	6	0%	2	0	100%	4	3	25%	17	13	24%	12	12	0%	72	50	31%
Consumer Protection	5	3	40%	2	1	50%	1	1	0%	17	7	59%	2	1	50%	46	22	52%
Data Protection	3	0	100%	3	0	100%	0	0	-	14	10	29%	1	0	100%	33	17	48%
Total (4)	80	70	13%	79	58	27%	44	36	18%	280	213	24%	97	88	9%	1375	1059	23%

(continued)

Table 16.10 (continued)

	Korea PTAs			Japan PTAs			China PTAs			EU PTAs			US PTAs			World PTAs		
	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation	AC	LE	Legal Inflation
Group 5. Others																		
Agriculture	4	3	25%	4	3	25%	3	1	67%	29	9	69%	3	3	0%	87	33	62%
Apporx. of Legislation	0	0	-	1	0	100%	0	0	-	22	13	41%	0	0	-	33	15	55%
Audio Visual	6	4	33%	0	0	-	1	0	100%	15	5	67%	0	0	-	24	10	58%
Civil Protection	0	0	-	0	0	-	0	0	-	4	2	50%	0	0	-	4	2	50%
Innovation Policies	2	0	100%	1	0	100%	2	1	50%	3	0	100%	2	0	100%	29	7	76%
Cultural Cooperation	5	4	20%	0	0	-	5	1	80%	29	13	55%	0	0	-	46	18	61%
Economic Policy Dialogue	1	1	0%	2	2	0%	2	2	0%	19	11	42%	0	0	-	67	21	69%
Education and Training	3	0	100%	5	3	40%	5	2	60%	21	12	43%	0	0	-	55	20	64%
Energy	5	2	60%	5	4	20%	0	0	-	28	10	64%	1	1	0%	57	28	51%
Financial Assistance	1	0	100%	2	2	0%	1	0	100%	27	12	56%	0	0	-	40	14	65%
Health	2	1	50%	1	1	0%	1	0	100%	17	12	29%	2	2	0%	31	17	45%
Human Rights	0	0	-	0	0	-	1	0	100%	17	1	94%	0	0	-	24	3	88%
Illicit Drugs	0	0	-	1	0	100%	0	0	-	16	2	88%	0	0	-	25	6	76%
Industrial Cooperation	1	0	100%	2	1	50%	4	3	25%	29	11	62%	0	0	-	61	22	64%
Information Society	7	3	57%	6	3	50%	3	1	67%	20	2	90%	3	0	100%	90	12	87%
Mining	2	2	0%	1	0	100%	2	1	50%	12	4	67%	0	0	-	22	12	45%
Money Laundering	0	0	-	0	0	-	0	0	-	17	1	94%	0	0	-	19	3	84%
Nuclear Safety	0	0	-	0	0	-	1	0	100%	10	6	40%	1	0	100%	14	6	57%
Political Dialogue	3	1	67%	0	0	-	0	0	-	18	2	89%	1	0	100%	42	6	86%
Public Administration	3	0	100%	1	0	100%	2	0	100%	9	2	78%	3	0	100%	53	10	81%
Regional Cooperation	1	0	100%	6	2	67%	3	1	67%	26	5	81%	3	0	100%	88	13	85%
Research and Technology	5	2	60%	4	2	50%	4	2	50%	29	11	62%	0	0	-	68	24	65%
SME	3	2	33%	7	4	43%	4	2	50%	13	2	85%	0	0	-	42	16	62%
Statistics	0	0	-	0	0	-	0	0	-	23	9	61%	0	0	-	51	29	43%
Taxation	0	0	-	1	1	0%	1	0	100%	14	10	29%	0	0	-	38	16	58%
Terrorism	0	0	-	0	0	-	1	0	100%	11	4	64%	0	0	-	15	6	60%
Total (5)	54	25	54%	50	28	44%	46	17	63%	478	171	64%	19	6	68%	1125	369	67%

Note AC: "Area Covered"; LE: "Legally Enforceable"
Source Own calculations based on World Bank Data. Content of Deep Trade Agreements. (http://data.worldbank.org/data-catalog/deep-trade-agreements). (updated October, 2016)

Including China, Japan and Korea, it is common that the legal inflation rate is low for Group 1 and Group 2. The inflation rate for Group 2 among all three countries' PTAs is lower than that of the world's PTAs. Relative to the EU and the world average, PTAs of China and Korea all exhibit lower legal inflation rates for Group 4, while the PTAs of Japan show relatively higher legal inflation rates. The inflation rate for Group 5, including all remaining areas, is significantly high among all three countries' agreements (54% for Korea, 44% for Japan, and 63% for China). This is because Group 5 mostly contains provisions related to cooperation and development issues (regional, industrial, cultural cooperation) as well as promotion of industry/technology (technical assistance, audiovisual) and exchange of information (data protection, information society).

16.5 Closing Remarks

In this paper, based on the methodology of Horn et al. (2010) and the dataset of the World Bank, we analyze all preferential trade agreements (PTAs) of China, Japan and Korea which date of entry into force is prior to February 2015. In comparing the provisions in the WTO+ areas and WTO-X areas for the three sets of PTAs, we find that China, Japan and Korea all have a large number of legally enforceable WTO+ provisions, which are covered by the current WTO mandate. While all three countries' agreements show a tendency to cover more of the new WTO-X areas in recently signed agreements, the number of legally enforceable WTO-X provisions contained in all three countries' agreements is still quite low. It is interesting to note that China has the highest inflation rate for WTO+ areas (8.2%), followed by Japan (5%) and Korea (2.8%), and similarly for WTO-X areas, China (43.7%) has the highest inflation rate, followed by Korea and Japan (32.8% each).

Considering the trends of the three countries' trade agreements, it will be necessary for the Korean government to develop strategies for certain areas that counterparties have dealt with to varying levels up to now. According to the results of this study, the SME provisions are more often included in Japan's PTAs, than China's and Korea's. In contrast, Japan is passive in dealing with labor issues in their PTAs, which provisions are frequently incorporated in Korea's PTAs. In the case of China, fields that are contrary to the socialist market economy system (e.g. public procurement, state trading enterprises, labor and environment) have generally been excluded from trade negotiations so far.[28] However, China has begun to include various areas, including the above sensitive areas, in recent PTAs.

Among the three countries, China's PTAs have evolved most dramatically with time. Since its first PTA signed with a developed country, New Zealand, in 2008, China has taken a more comprehensive and vigorous approach toward PTAs. For example, agreements with Iceland and Switzerland, signed in 2013, provide wider coverage in goods, services, and investments. China's PTAs expand from merely trade issues in goods and services to a broader coverage including investment, com-

[28] Park and Kwon (2013), p. 9.

petition policy, intellectual property rights, environment, economic and technical cooperation, and even potentially trade remedies and government procurement provisions. For example, environmental issues began to be gradually included in China's PTAs that have recently entered into force. In a recent FTA Korea has concluded with China, the environment chapter is included and evaluated as the most comprehensive of all of China's FTAs.[29]

However, despite its resemblance to the best FTA practices in Asia, such as seen in the Korea-US and Korea-EU FTAs, the level of liberalization of the Korea-China FTA is not as comprehensive in terms of coverage and depth. Based on this study, it is shown that the depth and the comprehensiveness of China's PTA is converging to the level of PTAs signed by Korea and Japan, and thus there is ample possibility to pursue a deeper agreement between the three countries.

Acknowledgements This paper was prepared as KIEP policy reference 17-01. We gratefully acknowledge financial support from KIEP. All remaining errors are our own.

Appendix

WTO-plus areas	
FTA Industrial	Tariff liberalization on industrial goods; elimination of non-tariff measures
FTA Agriculture	Tariff liberalization on agriculture goods; elimination of non-tariff measures
Customs	Provision of information; publication on the Internet of new laws and regulations; training
Export Taxes	Elimination of export taxes
SPS	Affirmation of rights and obligations under the WTO Agreement on SPS; harmonization of SPS measures
TBT	Affirmation of rights and obligations under WTO Agreement on TBT; provision of information; harmonization of regulations; mutual recognition agreements
STE	Establishment or maintenance of an independent competition authority; nondiscrimination regarding production and marketing condition; provision of information; affirmation of Art XVII GATT provision
AD	Retention of Antidumping rights and obligations under the WTO Agreement (Art. VI GATT)
CVM	Retention of Countervailing measures rights and obligations under the WTO Agreement (Art VI GATT)
State Aid	Assessment of anticompetitive behaviour; annual reporting on the value and distribution of state aid given; provision of information
Public Procurement	Progressive liberalisation; national treatment and/or non-discrimination principle; publication of laws and regulations on the Internet; specification of public procurement regime
TRIMs	Provisions concerning requirements for local content and export performance of FDI
GATS	Liberalisation of trade in services
TRIPs	Harmonisation of standards; enforcement; national treatment, most-favoured nation treatment

29 Schott et al. (2015), p. 14.

WTO-X areas	
Anti-Corruption	Regulations concerning criminal offence measures in matters affecting international trade and investment
Competition Policy	Maintenance of measures to proscribe anticompetitive business conduct; harmonisation of competition laws; establishment or maintenance of an independent competition authority
Environmental Laws	Development of environmental standards; enforcement of national environmental laws; establishment of sanctions for violation of environmental laws; pubblications of laws and regulation
IPR	Accession to international treaties not referenced in the TRIPs Agreement
Investment	Information exchange; Development of legal frameworks; Harmonisation and simplification of procedures; National treatment; establishment of mechanism for the settlement of disputes
Labour Market Regulation	Regulation of the national labour market; affirmation of International Labour Organization (ILO) commitments; enforcement
Movement of Capital	Liberalisation of capital movement; prohibition of new restrictions
Consumer Protection	Harmonisation of consumer protection laws; exchange of information and experts; training
Data Protection	Exchange of information and experts; joint projects
Agriculture	Technical assistance to conduct modernisation projects; exchange of information
Approximation of Legislation	Application of EC legislation in national legislation
Audio Visual	Promotion of the industry; encouragement of co-production
Civil Protection	Implementation of harmonised rules
Innovation Policies	Participation in framework programmes; promotion of technology transfers
Cultural Cooperation	Promotion of joint initiatives and local culture
Economic Policy Dialogue	Exchange of ideas and opinions; joint studies
Education and Training	Measures to improve the general level of education
Energy	Exchange of information; technology transfer; joint studies
Financial Assistance	Set of rules guiding the granting and administration of financial assistance
Health	Monitoring of diseases; development of health information systems; exchange of information
Human Rights	Respect for human rights
Illegal Immigration	Conclusion of re-admission agreements; prevention and control of illegal immigration
Illicit Drugs	Treatment and rehabilitation of drug addicts; joint projects on prevention of consumption; reduction of drug supply; information exchange
Industrial Cooperation	Assistance in conducting modernisation projects; facilitation and access to credit to finance
Information Society	Exchange of information; dissemination of new technologies; training
Mining	Exchange of information and experience; development of joint initiatives
Money Laundering	Harmonisation of standards; technical and administrative assistance
Nuclear Safety	Development of laws and regulations; supervision of the transportation of radioactive materials

(continued)

(continued)

WTO-X areas	
Political Dialogue	Convergence of the parties' positions on international issues
Public Administration	Technical assistance; exchange of information; joint projects; Training
Regional Cooperation	Promotion of regional cooperation; technical assistance programmes
Research and Technology	Joint research projects; exchange of researchers; development of public-private partnership
SME	Technical assistance; facilitation of the access to finance
Social Matters	Coordination of social security systems; non-discrimination regarding working conditions
Statistics	Harmonisation and/or development of statistical methods; training
Taxation	Assistance in conducting fiscal system reforms
Terrorism	Exchange of information and experience; joint research and studies
Visa and Asylum	Exchange of information; drafting legislation; training

References

Balassa, B. 1961. *The theory of economic integration*. London: George Allen & Unwin.

Ferrantino, M.J. 2012. Using supply chain analysis to examine the costs of non-tariff measures (NTMs) and the benefits of trade facilitation. Staff Working Paper, ERSD-2012-02. World Trade Organization.

Hofmann, C., A. Osnago, and M. Ruta. 2017. Horizontal depth: A new database on the content of preferential trade agreements. Policy Research Working Paper, No. WPS 7981. Washington, D.C.: World Bank Group.

Horn, H., P.C. Mavroidis, and A. Sapir. 2010. Beyond the WTO? An anatomy of EU and US preferential trade agreements. *The World Economy* 33 (11): 1565–1588.

Kim, S.Y. 2012. Negotiating the Nexus: Production networks and behind-the-border commitments in regional trade agreements. Unpublished Manuscript.

Kim, S.Y. 2015a. Deep integration and regional trade agreements. In *The Oxford handbook of the political economy of international trade*, ed. Lisa L. Martin, 360–379. Oxford: Oxford University Press.

Kim, S.Y. 2015b. Regionalisation in search of regionalism: Production networks and deep integration commitments in Asia's PTAs. *Trade cooperation: The purpose, design and effects of preferential trade agreements*, 134–165. Cambridge: Cambridge University Press.

McCaffrey, S.C. 2006. *Understanding international law*. New York: Lexis Nexis.

Matsushita, M., T.J. Schoenbaum, P.C. Mavroidis, and M. Hahn. 2015. *The world trade organization: Law, practice, and policy*. Oxford: Oxford University Press.

Mitchell, A.D., and N.J.S. Lockheart. 2016. Legal requirements for PTAs under the WTO. *Bilateral and regional trade agreements: Commentary and analysis*, 81–114. Cambridge: Cambridge University Press.

Orefice, G., and N. Rocha. 2014. Deep integration and production networks: An empirical analysis. *The World Economy* 27 (1): 106–136.

Orefice, G. 2015. International migration and trade agreements: The new role of PTAs. *Canadian Journal of Economics/Revue canadienne d'économique* 48 (1): 310–334.

Osnago, A., N. Rocha, and M. Ruta. 2016. Deep agreements and global value chains. *Global Value Chain Development Report 2016*. Background Paper Conference. World Bank.

Park, S.Y., and S.H. Kwon. 2013. Current status of TPP and position of major countries. *Foreign Economic Focus*, 2013-13. The Bank of Korea.

Schott J.J., E.J. Jung, and C. Cimino-Isaacs. 2015. An assessment of the Korea-China free trade agreement. Peterson Institute for International Economics. Policy Brief, No. PB 15-24.

Salidjanova, N. May 28, 2015. China's trade ambitions: Strategy and objectives behind China's Pursuit of free trade agreements. *Staff Research Report.* U.S.-China Economic and Security Review Commission.

United Nations. 1969. *Vienna convention on the law of treaties*, Treaty series, vol. 1155, 333. United Nations.

World Bank dataset. 2016. Content of deep trade agreements. http://data.worldbank.org/data-catalog/deep-trade-agreements. Accessed on 30 Oct 2016.

World Trade Organization. 2011. *World Trade Report 2011*.

World Trade Organization. 2016. Responding to least developed countries' special needs in intellectual property. https://www.wto.org/english/tratop_e/trips_e/ldc_e.htm. Accesssed on 11 Nov 2016.

World Trade Organization database. 2006. Regional trade agreements information system (RTA-IS). http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx. Accessed on 30 Oct 2016.

Chapter 17
Effects of Multidimensional Poverty on Health Indicators in Japan: Income, Time, and Social Relations

Wei Wang and Kunio Urakawa

Abstract In recent years, many studies have been tackling the measurement of multidimensional poverty, reflecting the diversification and complexity of poverty even in developed countries as well as developing countries. These studies treat on several dimensions of poverty such as education and living environment as well as income, but few of them consider the time dimension. This study attempts to investigate the impact of multidimensional poverty including time poverty on key health indicators (self-rated health (SRH), psychological distress (K6)) in Japan. By using individual data from Japanese Study on Stratification, Health, Income, and Neighborhood (J-SHINE) [2010, 2012], we measured multidimensional poverty index, based on the method of Alkire and Foster (Journal of Public Economics 95:476–487, 2011). We mainly set three dimensions of poverty (income, time, and social relations) and investigated the impact on health statuses, controlling other important variables. Results obtained from the analysis confirmed the practical relevance of multidimensional poverty for predicting health indicators.

17.1 Introduction

An origin of the multidimensional poverty approach, Sen's capabilities approach, proposed the characterization of poverty from the viewpoint of the degree of freedom that people can achieve (Sen 1992). Based on the capabilities approach, many studies measured deprivation on widely diverse dimensions such as access to education, living standards, social relations, and health conditions, rather than merely examining income. In terms of the methodology for identification, Atkinson (2003) discussed

W. Wang
Shanghai University of Engineering Science, Shanghai, China

K. Urakawa (✉)
Kyushu University, Fukuoka, Japan
e-mail: urakawa@econ.kyushu-u.ac.jp

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_17

a ‘counting approach’ that summarizes the number of dimensions on which people are deprived.

In recent work, Alkire and Foster (2011) provided a multidimensional approach (hereafter, AF method) that embodied Sen’s capabilities theory and which was motivated by Atkinson’s ‘counting method’ of measuring deprivation. The poverty measurement by AF method can be conceptualized as the following two main steps: identification of who is poor, and aggregation of information related to poverty across society (Alkire and Santos 2013). For the identification of poverty, both the indicators’ cutoffs z_j on each dimension and the poverty cut-off k, which is the number of weighted deprivations required to be considered multi-dimensionally poor, are considered. For that reason, the AF methodology is said to follow a dual cut-off method of identification (Alkire and Santos 2013). For measurement of the aggregated poverty indicator, most FGT measures can reflect the width and depth of deprivation from multiple dimensions.

Actually, AF method has been applied to multidimensional poverty analysis to assess various trends of poverty over time in countries and regions throughout the world (Batana 2013; Battiston et al. 2013; Santos 2013; Yu 2013). In these studies, researchers selected important dimensions for measuring poverty based on (1) human psychological needs (e.g., food and water, educational attainment, housing conditions, security) and (2) philosophical human value (e.g., income and wealth, well-being, social inclusion, skills) (Alkire 2002). However, few such studies have drawn attention explicitly to the time dimension. Capabilities toward poverty reduction that Sen emphasized are defined as individual freedom and opportunities to achieve the kind of life that they have reason to value (Sen 1992). In this context, securing living hours for the realization of various opportunities should be regarded as a candidate poverty dimension. Time is a finite resource along with money. It can play a salient role in fulfilling basic needs of life (Vickery 1977).

Previous reports of the relevant literature have specifically described situations of time poverty. They have demonstrated that not a few people or groups nowadays are adversely affected by a lack of time. For example, Harvey and Mukhopadhyay (2007) reported a high time-poverty rate for single-parent households in Canada. After analyzing conditions in Japan, Ishii and Urakawa (2014) showed that not only single-parent households but also double earner (both full-time employed) couples with small children are greatly deprived of time resources.

Importantly, the measurement of poverty incorporating the viewpoint of living hours is expected to complement government efforts at health improvement of people. From European data, Nolan and Marx (2009) observed that low income alone is insufficient to predict who is experiencing poor health and access to health services. Rather, recent studies have demonstrated clear correlation between time poverty and health. For example, Bittman (2002), using Time Use Survey in Australia, demonstrated that women who have less discretionary time are more likely to have poor health.[1] These studies fundamentally addressed income and time dimensions. They

[1] However, Kalenkoski et al. (2011) showed a positive relation between time poverty and subjectively evaluated health levels using the American Time Use Survey. Regarding this result, they pointed

did not compare the independent effects of respective dimensions and the composite effects multiple dimensions. Regarding relations between health and multidimensional poverty, Oshio and Kan (2014) investigated them in Japan using AF methodology and found that multidimensional poverty is more useful for predicting poor self-rated health and high psychological distress. However, this analysis was conducted on a single-year basis. It did not treat the time dimension.

In Japan, the lack of leisure hours caused by long working and commuting time and other factors is regarded as strongly related to severe health problems (Iwasaki et al. 2006). As represented by the term 'Karoshi,' much Japanese research shows an association between working hours and cerebrovascular/ cardiovascular disease (Ohtsu et al. 2013). To ascertain the determinants of health more clearly, an analysis incorporating the viewpoint of the securing a minimum amount of leisure hours should be conducted. Therefore, in this study using Japanese micro-panel data, we examine how multi-dimensional poverty, including that of the time dimension, is associated with health.

17.2 Data and Measurement

17.2.1 Data

Data used for this study were obtained from the Japanese Study of Stratification, Health, Income, and Neighborhood (J-SHINE) conducted by the University of Tokyo. J-SHINE is an ongoing longitudinal panel data originally intended for about 4000 adults, aged 20–55 years, who were selected randomly from urban and suburban municipalities in the greater Tokyo metropolitan area in Japan.[2] J-SHINE, which included various fields of questions such as income, employment status, living hours, and health conditions, was conducted to clarify complex associations between socioeconomic factors and health. From the study, we excluded students and limited respondents to the samples who answered both questionnaires in 2010 and 2012. The final estimation samples comprised 2314 individuals in 2010 and 2012. We mainly addressed six household categories to assess circumstances of multidimensional poverty.

- Couples with children who are all older than 6
- Couples with one child younger than 6
- Couples with two or more children younger than 6
- Couples without children
- Single parent with children
- One-person household

out that it is likely attributable to the fact that employed people have, on average, better health status than those who are not employed, and are also more likely to be time-poor.

[2]Regarding the detailed characteristics of J-SHINE data, refer to Takada et al. (2014).

Table 17.1 presents the demographic structures of samples in 2010 and 2012. The ratios of respective household types in the data were presented as follows: Couples with children who are all older than 6 [40.1% in 2010, 46.8% in 2012], Couples with one child younger than 6 [13.8% in 2010, 11.5% in 2012], Couples with two or more children younger than 6 [20.1% in 2010, 15.7% in 2012], Couples without children [18.5% in 2010, 18.7% in 2012], Single parent with children [3.0% in 2010, 2.9% in 2012], and One-person households [3.8% in 2010, 3.8% in 2012].

Regarding academic background, about 40% of the householders were respondents who had graduated from a four-year college or graduate school. Regarding employment status, about 43% were engaged in work as full-time employed, where Full time employed 1 means regular employees and Full time employed 2 means contract or dispatched full time but limited-term workers.

17.2.2 Poverty Dimension

Our poverty measure in this paper followed the multidimensional methodology reported by Alkire and Foster (2011). The AF methodology consists of an identification step and an aggregation step. At the identification step, we selected dimensions (indicators) and two cutoffs to identify who poor people using a 'dual cutoff' method. At the aggregation step, we mainly addressed the headcount ratio. For simplicity, all weights were assigned equally to each dimension.

We specifically examined three dimensions of poverty: 1. income, 2. time and 3. social inclusion. Then we investigated the effects on two indicators of health: self-rated health (SRH) and psychological distress (K6). By assessing poverty from three dimensions of income, time, and social inclusion, we clarified the following: (1) What households are likely to fall into time and other dimensions of poverty? (2) Does multiple poverty that includes a time concept prevent people from achieving good health conditions? (3) Are there any gender differences in the effects of multidimensional poverty on health?

In applying a dual cutoff approach, let D(•) denote the number of deprivations in poverty dimensions. The three union sets were defined as D(1) = 1 indicating income poverty, D(2) = 1 indicating time poverty, and D(3) =1 denoting social connection poverty. The full intersection set, D(1, 2, 3) =3 D (1, 2, 3) = 3 includes individuals who were deprived in all three poverty dimensions. In addition, D(1, 3) = 2 denotes individuals who have both income poverty and social connections poverty. Therefore, multidimensional poverty was defined as D(1, 2, 3) $\geqslant k$, where k =1, 2 or 3. Here, k represents the number of deprivations necessary to be judged multidimensionally poor. An individual was regarded as multidimensional poor if the person's D(•) was above a certain aggregated level of cutoff. For example, D(1, 2, 3) ⩾2 denotes individuals who were deprived in at least two of three poverty dimensions. The dual cutoff set also includes unidimensional poverty (if k = 1) and full intersection set (if k = 3). Indicators and the certain level of cutoff for three dimensions were defined as described below.

Table 17.1 Basic statistics of variables

	2010 (N = 2314)		2012 (N = 2314)	
	Mean (%)	Std. Dev	Mean (%)	Std. Dev
Sex				
Male	42.7	0.49	42.7	0.49
Female	57.3	0.49	57.3	0.49
Partner				
Yes	92.5	0.26	92.7	0.26
No	7.5	0.26	7.3	0.26
Age				
20s	13.6	0.34	8.0	0.27
30s	42.4	0.49	38.0	0.49
40s	44.0	0.50	54.1	0.50
Household type				
Couple with children who are all older than 6	40.1	0.49	46.8	0.50
Couple with one children younger than 6	13.8	0.34	11.5	0.32
Couple with more than two children younger than 6	20.1	0.40	15.7	0.36
Couple without children	18.5	0.39	18.7	0.39
Single-parent with children	3.0	0.17	2.9	0.17
One-person household	3.8	0.19	3.8	0.19
Householders' educational level				
High school graduate	22.9	0.42	23.2	0.42
Two-year college / technical college	36.6	0.48	36.6	0.48
Four-year college / graduate school	39.6	0.49	39.6	0.49
Working conditions				
Company/organization executives	2.5	0.16	2.6	0.16
Full time employed 1	43.3	0.50	43.3	0.50
Full time employed 2	5.6	0.23	5.6	0.23
Part time employed	18.0	0.38	18.2	0.39
Self-employed	6.2	0.24	6.4	0.24
No work	24.1	0.43	23.8	0.43

Note Full time employed 1 represents regular employed workers;
Full time employed 2 represents contract or dispatched full-time but limited term workers
Source Authors' calculations based on J-SHINE 2010–2012

(i) Dimension 1: Income (Household income)

The poverty line of income can be chosen in two ways: an absolute level and a relative level. OECD defines relative income poverty by measuring the ratio of the number of people who fall below the poverty line. The poverty line was set as half of the median of equivalent household income of the total population. According to this definition of income poverty, the average poverty rate of Japan for entire population was 16% in 2012, ranking high (sixth) among OECD 34 countries (OECD 2016, p.57). Importantly, Japan's poverty rates surpass OECD average for all of three age-groups (age 0–17, age 18–65, age 65 and over).

For the study, the reported household annual income by J-SHINE was divided by the square root of the number of family members, to adjust for household size. According to the calculation of the data, the poverty line was defined as JPY 1.84 million (equivalent to about USD 15,543) for each household per year. Therefore, a person for whom the equivalent annual income was less than JPY 1.84 million will be regarded as having been deprived on the income dimension.

(ii) Dimension 2: Time (Living hours)

Definitions of time poverty vary among studies. This study defined the time-poverty line as mainly following the method applied by Ishii and Urakawa (2014), which considered the disposable time and minimum time necessary for housework.

The estimation procedure was summed in three main steps: First, it defined the essential time (T_e) spent for basic activities (sleeping, eating, personal care (excretion, bathing, getting dressed, etc.)), which included the minimum leisure time required (one hour per day during Monday–Friday and three hours per day on Saturday and Sunday) in basic activities. Second, it defined the minimum time necessary for housework (T_1), consisting of housework, nursing care, child care, and shopping, based on a 2011 Survey on Time Use and Leisure Activities conducted by the Ministry of Internal Affairs and Communications (MIC). Third, the working and commuting time (T_w) were calculated from the data for each responded household.

As the time spent for basic activities (T_e), the study used the average time spent in an entire week by men and women, separately, who were aged 20–64 years. The time for basic activities included 7.5 hours per day for men and 7.2 hours per day for women for sleeping, 1.1 hours per day for men and 1.5 hours per day for women for taking personal care, and 1.5 hours per day for men and 1.6 hours per day for women for eating. Following the example of earlier studies, this study further included the minimum leisure time required in the basic activity time.

Regarding the minimum time required for housework (T_1), it represents the least time necessary for housework without outsourcing the series of household chores such as cooking, washing, taking care of children, elderly people, or disabled people, and shopping (dining out or ordering food delivery, purchasing prepared food, using housework-related services in the market economy, etc.), for each household type to be analyzed, the study applied the average time spent for housework at households with at least one adult who did not work outside the home. More specifically, as a reference, the study used the time spent for housework at households with a husband working outside the home and wife not working outside the home for households comprising a married couple and children and households only of a married couple. For one-person households and single-parent households, the study applied the housework time of households without a member working outside the home.[3]

[3]Because many men living alone are likely to outsource much of their housework by, for instance, dining outside instead of cooking at home, the minimum housework time required at female one-person households was substituted for that at male one-person households.

Table 17.2 Minimum time required for basic activities and housework by household types

		Essential time (T_e)				Minimum time required for the housework T_1					
	7 days (V)		Leisure (weekday)	Leisure (weekend)	$T_m(V - T_e)$	Housework	Nusrsing/Caring	Childcare	Shopping	Total	Disposal time $T_a(T_m - T_1)$
(hours)	week	week	day	day	week	day	day	day	day	week	week
Couple with children											
with children who are all older than 6	336	165.5	2.0	6.0	170.5	5.5	0.2	0.4	1.2	50.9	119.6
with one child youger than 6	336	165.5	2.0	6.0	170.5	4.0	0.1	5.0	1.1	71.3	99.2
with two or more children younger than 6	336	165.5	2.0	6.0	170.5	3.7	0.1	6.2	1.0	77.0	93.5
Harvery and Mukhopadhyey (2007)	336	175.0	4.0	4.0	161.0	–	–	–	–	74.6	86.4
Couple without children	336	165.5	2.0	6.0	170.5	4.3	0.1	0.1	1.1	39.4	131.1
Vickery (1977)	336	162.8	2.0	5.0	173.2	–	–	–	–	74.6	86.4
Single parent with children	168	83.2	1.0	3.0	84.8	3.5	0.1	1.1	1.0	39.3	45.5
Harvery and Mukhopadhyey (2007)	168	87.5	2.0	2.0	80.5	–	–	–	–	52.0	28.5
One-person household	168	82.6	1.0	3.0	85.4	2.3	0.1	0.0	0.6	21.2	64.2
Vickery (1977)	168	81.4	1.0	2.5	86.6	–	–	–	–	31.0	55.6

Source Authors' calculations based on Survey on Time Use and Leisure Activities 2011

Consequently, the disposal week time (T_a) is calculated as Total time (V) - (T_e) - (T_1). According to these procedures, an individual is regarded as suffering from time-poverty if the disposal time is less than the working and commuting time (T_w). Table 17.2 presents the time spent for basic activities and the minimum time required for housework for six household types.

(iii) Dimension 3: Social relations (Social exclusion)

Social exclusion is a multidimensional general concept that refers not only to material, economic, or health deprivation, but also to deprivation from social relations and participation in society. For Abe (2012), the lack of social relations and social participation has been brought up in relation to poverty in Japan. Moreover, Ikeda et al. (2013) emphasized that good social relationships can contribute to good health conditions for Japanese people. The concrete contents of questionnaires related to social exclusion are the following.
[Question]
"Now we will ask you a question about social exclusion and social support. If you encounter some trouble, how often you will contact or communicate with your family members, partners, neighbors, or friends?"
[List]

(1) Often. (2) Sometimes. (3) A little. (4) Never.
(5) I have no family member, partner, neighbor, or friend I can talk with.

We produce a dummy variable for social exclusion in considering the additional dimension of multidimensional poverty, to which we allocated a 1 if an individual's respondent was "Never" or "I have no family member, partner, neighbor, or friend I can talk with."

17.2.3 Health Indicators

Regarding health indicators, we considered self-rated health (SRH) and Kessler 6 (K6). Binary variables for two health indicators were constructed. The questionnaire related to responders' self-rated health conditions is formulated with five choices: good, somewhat good, average, somewhat poor, and poor. If the respondent answered poor or somewhat poor, then the respondent was treated as poor SRH.

For measuring psychological distress, Kessler 6 (K6) scores were used in this study. The K6 comprises six-item psychological questions about depressive and anxiety symptoms that an individual has experienced during a 30-day reference period: (1) nervousness, (2) hopelessness, (3) restlessness or fidgeting, (4) severe depression, (5) lack of motivation, and (6) worthlessness rated on a five-point scale (0 = not at all, 1, 2, 3, and 4 = all of the time). The summation of K6 scores gives a final score of 0–24. Here, respondents whose K6 scores were greater than seven were regarded as persons confronting high risk of psychological distress, according to the identification of earlier Japanese research (Sakurai et al. 2011).

Table 17.3 presents descriptive statistics of health and poverty measures. About 10% of respondents reported poor self-rated health (SRH) in both years (2010, 2012). About 20% of respondents were regarded as people whose levels of K6 are high. Regarding categories of poverty, we considered the rates of respondents only from each type of poverty and multiple poverty. Here, income poverty means D(1) = 1 & D(1, 2, 3) = 1. People who are both deprived of income and other dimensions are excluded. In 2010, people in time poverty are about 11%. This figure was followed by 8% of income poverty and 2% of social exclusion. As a whole, no large differences of health and poverty measures during two years were found, but in 2012, the number of people in non-poverty had increased by about 4 percentage points.

17.3 Data Analysis

17.3.1 Headcount Ratios by Dimension

Based on settings of multidimensional poverty, we first specifically examined the characteristics of three unidimensional types of poverty (D(1) = 1, D(2) = 1, and D(3) = 1) by different categories. Table 17.4 presents the situation of unidimensional poverty status by six categories: gender, partner, age class, household type, householders' educational level, and respondents' working conditions.

According to the table, the poverty ratios of three dimensions decreased from 2010 to 2012 for both men and women. In addition, respondents with no partner face high risk of falling into the categories of low income, lack of leisure time, and social exclusion. Regarding the age group, respondents in their twenties and thirties reported higher income poverty in 2010 than in 2012. Regarding household type categories,

Table 17.3 Health and poverty measures

	2010 (N = 2314)		2012 (N = 2314)	
	Mean (%)	Std. Dev	Mean (%)	Std. Dev
Well-being				
Poor self-rated health	10.4	0.30	9.6	0.29
K6 (high)	21.2	0.41	22.8	0.42
Poverty category				
Income poverty D(1) = 1 & D(1, 2, 3) = 1	7.7	0.27	6.7	0.25
Time poverty D(2) = 1 & D(1, 2, 3) = 1	10.5	0.31	7.5	0.26
Social relations poverty D(3) = 1 & D(1, 2, 3) = 1	2.3	0.15	2.3	0.15
Multiple poverty D(1,2,3) >= 2	1.6	0.13	1.4	0.12
No poverty	77.9	0.42	82.1	0.38

Source Authors' calculations based on J-SHINE 2010-2012

Table 17.4 Poverty rates of each dimension by individuals' attributes (2010 and 2012)

	Poverty dimension					
	Income		Time		Social relations	
	2010 (%)	2012 (%)	2010 (%)	2012 (%)	2010 (%)	2012 (%)
All Persons	8.8	7.8	11.8	8.4	3.3	3.2
Sex						
Male	7.6	7.0	14.7	10.4	5.9	5.5
Female	9.7	8.5	9.6	6.9	1.4	1.5
Partner						
Yes	7.4	6.5	10.6	7.1	1.5	1.6
No	26.6	24.3	26.6	24.3	26.0	23.1
Age						
20s	13.1	9.8	13.4	7.6	4.5	4.4
30s	9.8	9.2	14.3	10.7	2.6	2.8
40s	6.6	6.6	8.8	6.9	3.6	3.3
Household type						
Couple with children who are all older than 6	5.8	5.4	4.2	3.8	1.4	2.0
Couple with one children younger than 6	7.5	9.0	22.6	16.9	2.2	0.8
Couple with more than two children younger than 6	13.4	10.7	21.3	16.8	0.7	0.6
Couple without children	4.2	4.4	3.7	1.4	1.9	2.1
Single-parent with children	42.9	37.9	35.7	36.4	12.9	12.1
One-person household	15.9	15.7	23.9	19.1	36.4	31.5
Householders' educational level						
Junior-high / High school graduates	15.1	11.5	10.7	6.3	3.2	3.7
Two-year college and technical school	9.3	7.8	10.9	7.4	2.0	2.2
Four-year college and graduate school	4.9	5.7	13.3	10.5	4.4	3.8
Working conditions						
Management executive/Operator	6.8	5.0	11.9	15.0	3.4	3.3
Full time employed1	4.4	4.5	20.7	13.4	4.4	4.5
Full time employed2	14.7	14.0	14.7	10.9	7.0	3.9
Part time employed	12.0	8.8	7.9	6.6	3.1	3.3
Self-employed	17.4	15.7	3.5	6.1	2.1	1.4
No work	10.9	9.9	0.0	0.0	0.7	0.9

Source Authors' calculations based on J-SHINE 2010-2012 single parent households with children reported the highest income poverty ratio (42.9% in 2010 and 37.9% in 2012) and the highest time poverty ratio (35.7% in 2010 and 36.4% in 2012). One-person households and couples with children younger than six also showed high time-poverty rates. These results correspond to the findings reported by Ishii and Urakawa (2014).

Next, examining effects of the household head's educational level, people with low educational attainment (junior-high or high school graduates) reported higher risk of low income. Nevertheless, people with higher educational background (who had graduated from four-year college or graduate school) had high poverty ratios in the time dimension and social relations. In addition, regarding working conditions, regular full-time employed workers showed the lowest rates of income poverty, but reported the highest time poverty rates among all working conditions, which indicates a tradeoff between money and time. Self-employed workers also showed high income poverty ratios. Contract or dispatched full-time but limited-term workers showed high rates for social exclusion, particularly in 2010.

17.3.2 Descriptive Statistics for Poverty and Health

Table 17.5 presents the proportions of people with poor health (poor SRH and psychological distress (K6 ⩾7K6 ⩾)) for respective poverty type and gender in 2010 and 2012. Regarding poverty types, we considered three unidimensional poverty types (D(1) = 1, D(2) = 1, and D(3) = 1) and two multidimensional poverty types (D(1, 2, 3) ⩾1 and D(1, 2, 3) ⩾ 2).

For poor SRH, we first observed that among unidimensional poverty of three types, 17.2% (in 2010) and 18.2% (in 2012) of people who fall into income poverty (D(1) = 1) perceived themselves as having poor SRH, compared to 10.4% (in 2010) and 9.6% (in 2012) of the entire sample. In total, the proportions of men with poor SRH is higher than those of women in 2010 and in 2012. The rates of people who fall into time poverty (D(2) = 1) and are in poor SRH are lower than the cases of income poverty (9.2% in 2010, and 9.8% in 2012), the proportion of women with poor SRH and time poverty is much higher (11.8% in 2010, 13.2% in 2012) than those of men (6.9% in 2010, and 6.8% in 2012). Regarding social relations poverty (D(3) = 1), the proportions of people who reported poor health were 18.4% (in 2010) and 21.6% (in 2012) in total. That figure showed the highest influence compared to other two poverty dimensions. Regarding the multidimensional poverty, we observed that 29.0% (in 2010) and 21.9% (in 2012) of people who fall into at least two of three poverty dimensions (D(1, 2, 3) ⩾2) assessed their health as poor. Therefore, results suggest that people with multidimensional poverty status tend to fall into poorer health than those affected by one-dimensional poverty.

Related to the results of psychological distress (K6⩾7), results show that low levels of income and social connections are particularly related to a higher proportion of psychological distress. Among income-poor people, the rates of people with a high score of K6 were almost 30% in 2010 and 2012. Regarding people who fall into multidimensional poverty, results show that multiple poverty statuses increased the risk of high K6 in several situations. For example, people who had at least two of the three poverty dimensions (D(1, 2, 3) ⩾2) reported that 50% of them marked high K6 score in 2012, which was very high compared to the average proportion of people who have psychological distress (22.8%).

17.4 Panel Data Analysis

In the study, panel logit analysis is conducted to estimate how each binary variable showing a poverty regime affects health conditions. Table 17.6 shows the estimated results of their relation. In econometric analyses, each dummy variable that represents poor SRH and high Kessler 6 scores was used as the explained variable. For key

Table 17.5 Poverty and poor health: a descriptive analysis

[2010]		Proportion (%) of					
		Poor SRH			K6>= 7		
Dimension of poverty	Definition of poverty	All	Male	Female	All	Male	Female
No poverty	D(1, 2, 3) = 0	9.8	8.9	10.5	20.3	20.2	20.4
1. Income	D(1) = 1	17.2	18.7	16.3	29.4	29.3	29.5
2. Time	D(2) = 1	9.2	6.9	11.8	18.4	23.5	12.6
3. Social relations	D(3) = 1	18.4	15.5	27.8	40.8	36.2	55.6
Multidimensional poverty	D(1, 2, 3) >=1	12.3	11.3	13.3	24.4	26.9	22.0
	D(1, 2, 3) >= 2	29.0	20.0	38.9	36.8	35.0	38.9
All		10.4	9.5	11.0	21.2	21.9	20.7
[2012]		Proportion (%) of					
		Poor SRH			K6>= 7		
Dimension of poverty	Definition of poverty	All	Male	Female	All	Male	Female
No poverty	D(1, 2, 3) = 0	8.5	7.6	9.1	21.8	21.1	22.2
1. Income	D(1) = 1	18.2	15.9	19.6	32.6	37.7	29.5
2. Time	D(2) = 1	9.8	6.8	13.2	21.7	19.4	24.8
3. Social relations	D(3) = 1	21.6	25.9	10.0	41.9	44.4	35.0
Multidimensional poverty	D(1, 2, 3) >= 1	14.5	13.5	15.5	27.7	28.4	27.1
	D(1, 2, 3) >= 2	21.9	17.7	26.7	50.0	58.8	40.0
All		9.6	8.8	10.1	22.8	22.7	22.9

Source Authors' calculations based on J-SHINE 2010–2012

Table 17.6 Panel logit model of poor SRH and high K6

Model	Poor SRH (RE Model)		High K6 (RE Model)	
	Coefficient	z-value	Coefficient	z-value
Partner				
No	0.83***	3.08	0.68**	2.48
Yes (ref)				
Number of children				
One	-0.44**	-2.05	-0.69***	-3.36
More than two	-0.26	-1.39	-0.88***	-4.79
Zero (ref)				
Working conditions				
Business executive	-0.35	-0.74	-0.67	-1.40
Full time employed 1	-0.69***	-3.52	-0.34*	-1.80
Full time employed 2	-0.05	0.15	0.26	0.80
Part time employed	-0.19	-0.84	-0.04	-0.20
Self employed	-0.82**	-2.35	-0.04	-0.14
No work (ref)				
Poverty dimension				
Unidimensional poverty type				
Income poverty	0.60***	2.57	0.39*	1.77
Time poverty	0.15	0.58	-0.15	-0.62
Social relations poverty	0.77**	1.98	0.39	1.06
Multidimensional poverty type				
Deprived in at least two dimensions (D(1, 2, 3)>=2)	1.17***	2.63	0.79*	1.75
No poverty (ref)				
Number of observations	4628		4628	
Number of groups	2314		2314	
Wald chi2 (12)	56.46		56.23	
Prob > chi2	0.000		0.000	
Log-likelihood	-1415.17		-2232.80	

Note ***,**, and * are statistically significant at 1%, 5% and 10% level

independent variables, five variables of multiple poverty were constructed as follows: three unidimensional poverty types (only income poverty, only time poverty, only social relations poverty, and multidimensional poverty (income poverty and time poverty, income poverty and social exclusion, time poverty and social exclusion, and full dimensional poverty). Other control covariates were the following: a spouse dummy (1 if spouse exists, otherwise 0), a number of children dummy (one child, and more than two children and no child (reference)), a working conditions dummy (business executive, full time employed 1 full time employed 2, part time employed, self-employed, and no work (reference). We decided which model (fixed effects model or random effects model) should be employed according to Hausman test results. As a result, random effects models were applied for two econometric models.

Regarding the estimation results, the association with poor SRH was highly significant at the 1% level for income poverty, and at the 5% level for social relations poverty. In addition, multiple poverty ($D(1, 2, 3) \geqslant 2$) is highly associated with poor SRH. We were able to confirm a clear correlation between multiple poverty and people's poor health. Additionally, we observed that some variables such as single status, meaning people who have no partner, were also linked with poor SRH. As with results on poor SRH, regarding the influences on K6 score, we observed that each unidimensional poverty type of low income and social exclusion, and multiple poverty situations were related closely with poor mental health of people. In addition, people with no partner or children readily fell into the status of psychological distress. The effects of time poverty on health was not so strong compared to the other two poverty dimensions (income and social exclusion), but multiple poverty situations negatively affect people's health. For that reason, public policies aimed at the broader concept of poverty including the time dimension as well as monetary dimension should be reinforced.

17.5 Conclusion

For the present study, we investigated the effects of multidimensional poverty on key health indicators in Japan using panel data from Japanese study of stratification, Health, Income, and neighborhood (J-SHINE). By considering the multidimensionality poverty index, we can identify people in poor SRH and psychological distress more adequately than the mere usage of a one-dimensional poverty such as income. The study included time as a key poverty dimension for estimation, and investigated how multiple poverty including the time dimension prevented people from maintaining their health condition.

We first observed that the proportion of income poverty for women was higher than that for men in 2010 and 2012. However, in terms of the dimensions of time and social relations, the poverty ratios for men were much higher than for women. Gender differences are strongly related to the fact that male people work mainly as regular employees and women mainly work as non-regular workers or leave the labour market to care for their children. Similar to the results reported by Ishii and Urakawa (2014), we confirmed that lone parent households are associated with higher risks of time and income poverty, comparing to other household types. In addition, households with children younger than six years old presented high poverty ratios related to the time dimension. Moreover, we observed that one-person households marked high poverty ratios for the social relations dimension.

In the latter part, using the results of panel logit estimation, we confirmed the validity of multiple poverty for identifying people with poor SRH and psychological distress. For both SRH and psychological distress, we noticed that the multiple dimensions of poverty were more useful to predict people's poor health, which was consistent with results obtained from cross-section analysis of Oshio and Kan (2014). Regarding one-dimensional poverty types, monetary variables are key determinants of poor SRH and mental health, but social exclusion is also an important factor for predicting people's health.

Finally, several policy implications are described based on results obtained from the current study. First, many JSHINE data respondents are young people in Tokyo areas. Therefore, the estimation results probably clarified the severe situation of working and commuting time of workers and social exclusion of one-person households in Tokyo. People living in time poverty were particularly observed in the case of full-time working couples with preschool children. Policymakers should reconsider tightening regulations on overtime work, reflecting the trend of spreading long working hours at many firms. In addition, the working environment might be regarded as a risk factor of increasing the probability of social exclusion, because male and full time employees reported high ratios of deprivation of social inclusion. Takekawa et al. (2014) urged that there must be clear awareness of social exclusion when discussing policy options for poverty alleviation, particularly for working generations. Atkinson and Marlier (2010) reported that promotion of social inclusion will be helpful to create a society that is safer and more stable, which should be regarded as a necessary condition for sustainable economic growth and development. Consequently, to enhance their health condition, it is necessary to consider socially supportive policies such as improving the quality of working conditions for working generations.

Second, policy support for housework and child care particularly should be strengthened, especially for single-parent households, which must confront high risks of both income and time poverty. It has been difficult for single parents to work for a long time because of housework and childcare burdens, but their wage levels are lower than in other OECD countries. Not only in Japan, but also in some other economically developed countries, a phenomenon exists by which couples in more affluent regions have fewer children (Tachibanaki 2010), partly because of the increase of economic well-being and time costs of having children. Policies prompting enforcement of work–leisure balance should be implemented to help people who have children to realize a better work-life balance.

The current studies have been affected by several limitations. First, we have not investigated which combination of poverty dimensions will affect health indicators more because of data constraints. It is necessary to continue our studies by expanding the sample size or using another survey. Second, for simplification, we assigned equal weight to each dimension of poverty. However, it is necessary to investigate how different weights can be expected to affect the estimation results. We should consider how poverty is related to health conditions and health behaviors.

References

Abe, A. 2012. Poverty and social exclusion of women in Japan. *Japanese Journal of Social Security Policy* 9 (1): 61–82.

Alkire, S. 2002. Dimensions of human development. *World Development* 30: 181–205.

Alkire, S., and J.E. Foster. 2011. Counting and multidimensional poverty measurement. *Journal of Public Economics* 95: 476–487.

Alkire, S., and M.M. Santos. 2013. A multidimensional approach: Poverty measurement & beyond. *Social Indicators Research* 112: 239–257.

Atkinson, A.B. 2003. Multidimensional deprivation: Contrasting social welfare and counting approaches. *Journal of Economic Inequality* 1: 51–65.

Atkinson, A.B., and E. Marlier. 2010. *Analyzing and measuring social inclusion in a global context*, 010. United Nations (UNDESA), New York: Department of Economic and Social Affairs.
Batana, Y. 2013. Multidimensional measurement of poverty in Sub-Sahara Africa. *Social Indicators Research* 112: 337–362.
Battiston, D., G. Cruces, L.F. Lopez-Calva, M.A. Lugo, and M.E. Santos. 2013. Income and beyond: Multidimensional poverty in Six Latin American Countries. *Social Indicators Research* 112: 291–314.
Bittman,. 2002. Social participation and family welfare: The money and time costs of leisure in Australia. *Social Policy and Administration* 36 (4): 408–425.
Harvey, A., and A.K. Mukhopadhyay. 2007. When twenty-four hours is insufficient: Time poverty of working parents. *Social Indicators Research* 82: 57–77.
Iwasaki, K., M. Takahashi, and A. Nakata. 2006. Health problems due to long working hours in Japan: Working hours, worker's compensation (Karoshi) and preventive measures. *Industrial Health* 44: 537–540.
Ikeda, A., I. Kawachi, H. Iso, M. Iwasaki, M. Inoue, and S. Tsugane. 2013. Social support and cancer incidence and mortality: The JPHC study cohort II. *Cancer Causes and Control* 24 (5): 847–860.
Ishii, K., and K. Urakawa. 2014. Time-adjusted poverty among working households in Japan: Two-dimensional poverty line approach. *MITA Business Review* 57 (4): 97–121. (in Japanese).
Kalenkoski, C.M., K.S. Hamrick, and M. Andrews. 2011. Time poverty thresholds and rates for the US population. *Social Indicators Research* 104: 129–155.
Nolan, B., and I. Marx. 2009. Economic inequality, poverty and social exclusion. In *Oxford Handbook of Economic Inequality*, ed. W. Salverda, B. Nolan, and T.M. Smeeding. Oxford: Oxford University Press.
OECD (2016) OECD Factbook 2015–2016. OECD Publishing.
Ohtsu, T., et al. 2013. A cross-sectional study of the association between working hours and sleep duration among the Japanese working population. *Journal of Occupational Health* 55: 307–311.
Oshio, T., and M. Kan. 2014. Multidimensional poverty and health: Evidence from a nationwide survey in Japan. *International Journal for Equity in Health* 13 (1): 1–11.
Sakurai, K., A. Nishi, K. Kondo, K. Yanagida, and N. Kawakami. 2011. Screening performance of K6/K10 and other screening instruments for mood and anxiety disorders in Japan. *Psychiatry Clinic Neuroscience* 65: 434–441.
Santos, M.E. 2013. Tracking poverty reduction in Bhutan: Income deprivation alongside deprivation in other sources of happiness. *Social Indicators Research* 112: 259–290.
Sen, A.K. 1992. *Inequality Re-Examined*. Oxford: Clarendon Press.
Tachibanaki, T. 2010. *The new paradox for Japanese women: Greater choice, greater inequality*. Japan: International House of Japan Inc.
Takada, M., N. Kondo, and H. Hashimoto. 2014. Japanese study on stratification, health income, and neighborhood: Study protocol and profiles of participants. *Journal of Epidemiology* 24 (4): 334–344.
Takekawa. S., et al. 2014. Infrastructure development of social inclusive society. *Science Council of Japan*. (in Japanese)
Vickery, C. 1977. The time-poor: A new look at poverty. *Journal of Human Resources* 12 (1): 27–48.
Yu, J. 2013. Multidimensional poverty in China: Findings based on CHNS. *Social Indicators Research* 112: 315–336.

Chapter 18
A Real Options Analysis on Speculative Bubbles in Housing Market: Considering Seoul Metropolis Case

Hojeong Park

18.1 Introduction

As mega cities around the world have experienced rapidly rising prices in real estate markets since the mid-2000s, the issue on the scale of speculative bubbles has become controversial (Goodman and Thibodeau 2008; Kashiwagi 2014a; Kouwenberg and Zwinkels 2015). Seoul Metropolis, the capital of South Korea, is no exception to the recent booms in housing markets and accordingly to the bubble controversy. Since the house value generally accounts for more than half the household wealth in South Korea (Kim et al. 2015), it is crucially important to comprehend the scales of speculative bubbles in the housing market.

Recently, in response to the rising house price, there is growing voice which calls for the cancellation of greenbelt zones near to Seoul to increase the house supply. The speculative demand in the housing market creates adverse effects on the sustainable urban management through the extensive urban sprawl. However, before the lift of greenbelt zone is discussed, it is necessary to understand the nature of speculative behaviour and to identify the resulting bubbles. This paper is an attempt for this.

Although there are numerous ways to measure the scale of speculative bubbles, it is common to represent them as a portion of the asset price which exceeds the fundamental value of the asset. Scheinkman and Xiong (2003) analyse a continuous-time model in which agents' heterogeneous beliefs in the fundamental values on housing could generate bubbles in market when each agent has different overconfidence in the valuation. This approach is distinct from Harrison and Kreps (1978) in which heterogeneous belief is formed by the disagreement about the probability distribu-

H. Park (✉)
Korea University, Seoul, South Korea
e-mail: hjeongpark@korea.ac.kr

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_18

tion of dividend streams.[1] Granziera and Kozicki (2015) consider a rational bubble model by adopting rational expectation in which agents extrapolate the future from the latest realizations. They prove the model performance by empirically analyzing the U.S. housing boom and bust cycle of the 2000s.

Another strand of literature relies on the search model in studying speculative behavior in the asset markets. For example, Duffie et al. (2002) develop a search model to analyze the actual short-sale process and its implication for asset prices. Kashiwagi (2014b) develops a model of the rental and home ownership markets in a unified framework of search and matching theory. The paper analytically shows that rent is volatile, although the degree of change in house price is larger than that in rents. The literature using real options model to examine the speculative bubbles is somewhat recent development in the real options field. Turvey (2002) and Buchsteiner and Zavodov (2012) find the real options framework useful to study bubbles in the property market. Ekstrom and Tysk (2009) prove the usefulness of Black-Scholes equation in modelling speculative bubbles. Kivedal (2012) analyzes the ratio between the housing prices and the rental price to identify the possibility of bubbles in the US housing market, since, as noted by Himmelberg et al. (2005), the presence of persistent variations in housing prices from the rental price may indicate a bubble in the housing market.

This paper attempts to empirically examine the bubbles with a focus on the values of residential buildings in Seoul by employing a real option model. Scheinkman and Xiong (2003) provide an optimal-stopping model which analyzes an option to sell the asset with bubbles through trading among agents with heterogeneous beliefs. Similar to them, the model presented in this paper is based on the optimal-stopping framework but provides a more tractable solution by assuming a monopolistic setting of the investment decision. Furthermore, the paper presents a real options model to consider the role of mortgage in examining the scale of speculative bubbles. This is an empirically important issue since the mortgage accounts for on average more than 50% of household loans in South Korea. The presence of stochastic rental income variable and mortgage variable leads to a two-state model which makes the analysis more complicated compared to the one-state model. As manifested in Dixit and Pindyck (1994), two-state model generally relies on numerical analysis, since it is not straightforward to derive a closed-form solution in multiple states. This paper thus attempts to provide analytical solutions which do not necessarily require high-dimensional numerical analysis but give rather transparent and tractable working scheme. In the next section, the regional scales of speculative bubbles in Seoul are empirically estimated after developing a real options model to account for agent's investment decision in housing market.

[1]There are an increasing number of studies that adopt agents' heterogeneous beliefs or trading types between pessimists and optimists. Wigniolle (2014) studies the existence of rational bubbles with optimism and pessimism to identify deterministic and stochastic bubbles. Brock and Hommes (1998) consider heterogeneous beliefs between fundamentalists and chartists in the asset pricing model, while (Park, 2015) applies a similar model to assess the possibility of chaos emergence in the housing market.

The remainder of the paper is organized as follows: Sect. 18.2 develops a real option model to show the existence of speculative bubbles in housing investment decision. Section 18.3 analyses empirical data on apartment prices in Seoul Metropolis and calibrates the scale of bubbles. Conclusion is discussed in Sect. 18.4.

18.2 Model

There is no formal definition on the bubble as we observe that different papers adopt different perspectives on bubble. However, the basic concept on the bubble is rather clear as noted by Stiglitz (1990) who advocates a relatively psychological approach: "if the reason that the price is high today is only because investors believe that the selling price will be high tomorrow, then a bubble exists."

This notion of bubble is closely related to the rational bubble with investors who believe that they will be able to sell their assets in the future at a higher price than its fundamental value (Flood and Hodrick 1990). Such bubble definition can be embedded in the put option concept as shown in Dixit and Pindyck (1994).

Consider an economic agent who makes a decision whether to purchase a house when the rental income, $R(t)$, follows a geometric Brownian motion:

$$dR\left(t\right) = \alpha R\left(t\right)dt + \sigma R\left(t\right)dz\left(t\right). \tag{18.1}$$

Parameters α and σ denote the drift and volatility time-invariant parameter, respectively. The increment of the standard Wiener process is characterized by $E\left(dz\right) = 0$ and $V\left(dz\right) = dt$ with $R\left(0\right) = R$. Denote τ as the tax rate on the rental income. The mortgage payment is rm where r is the mortgage interest rate and m is the total mortgage. Then, similar to (Baker et al. 1991), the net income (after-tax return) for each instant of time is given by $(1-\tau)\,R\left(t\right) - rm$. Let $D(t)$ denote the household debt stock at time t. Upon the mortgage introduction, the incremental debt change is given by $dD(t)/dt = m$. The value function for the housing investment with the discount rate ρ is

$$W(R, D) = E_0 \int_0^{\infty} \left[(1-\tau)\,R - rm\right] e^{-\rho t} dt \tag{18.2}$$

which is subject to (18.1). Note that E_0 is the expectation operator at time 0. Prior to the investment in the housing, the decision maker holds an option value that is a function of $R\left(t\right)$. In what follows, the objective is to identify the first optimal passage time $\tau^* = \inf\{t | R(t) \geqslant R^*\}$, implying that it is optimal to purchase a house when its rental income exceeds the optimal threshold R^*. The time-notation is dropped unless it is necessary for clarity. Since there is no dividend-like income prior to the holding of house, the option value $V(R)$ is analogous to perpetual American option being expressed as

$$\rho V(R) = \alpha R V_R(R) + \frac{1}{2}\sigma^2 R^2 V_{RR}(R) \tag{18.3}$$

by applying the Ito's lemma. The solution to (18.3) is $V(R) = AR^{\beta}$ in which A denotes an option constant and β is defined by the characteristic equation $\alpha\beta + 0.5\sigma^2\beta(\beta - 1) - \rho = 0$. The dynamic programming approach expands the equation (18.2) to the following Hamilton–Jacobi–Bellman (HJB) equation:

$$\rho W(R, D) = (1 - \tau)R - rm + mW_D(R, D) + \alpha R W_R(R, D) + \frac{1}{2}\sigma^2 R^2 W_{RR}(RD). \tag{18.4}$$

For the case of D= 0 without initial debt stock, it is immediate to see that the HJB equation (18.4) degenerates to a one-factor model for which the solution could be trivially obtained. The boundary condition is $W(R, D) = (1 - \tau)\, R/(\rho - \alpha)$, intuitively indicating that the fundamental value of housing in the absence of house mortgage is the present value of perpetual rental income. An analytically tractable solution is obtained by setting $W(R, D) = B(e^{\gamma D}R)^{\varphi}$ which provides $W_D = \varphi B(e^{\gamma D}R)^{\varphi-1}\gamma e^{\gamma D}R$, $W_R = \varphi B(e^{\gamma D}R)^{\varphi-1}e^{\gamma D}$ and $W_{RR} = \varphi(\varphi - 1)$ $B(e^{\gamma D}R)^{\varphi-2}e^{2\gamma D}$. This finally leads to the solution to (18.4):

$$W(R, D) = \frac{(1-\tau)\, R}{\rho-\alpha} - \frac{rm}{\rho} - \delta e^{\frac{\gamma D}{m}} + B(e^{\gamma D}R)^{\varphi} \tag{18.5}$$

where δ is a scaling parameter. The first and second terms in the RHS, $(1 - \tau)\, R/(\rho - \alpha) - rm/\rho$, represents the present value of net rental income after paying the rental tax and mortgage payback. Note that in conventional valuation approach the first and second terms provides an information on the price of real asset. The third term is the negative appreciated value with respect to the debt stock based on the payback period. The scaling parameter is determined according to the above-specified boundary condition as $D \to 0$. The last term in the RHS, $B(e^{\gamma D}R)^{\varphi}$, denotes the speculative bubble. φ is the positive solution of the characteristic equation, $0.5\sigma^2\varphi^2 + \varphi\left(\alpha - m\gamma - 0.5\sigma^2\right) - \rho = 0$ to satisfy the boundary condition of $W(R, D) = (1 - \tau)\, R/(\rho - \alpha)$.

The negative root is not attainable since the presence of $\varnothing < 0$ does not satisfy the boundary condition of $W(R, D) = (1 - \tau)R/(\rho - \alpha)$ by allowing infinite value as $D \to 0$. The presence of the positive root, $\varnothing > 0$, indicates a component of $W(RD)$ attributable to the speculative bubble. The decision makers might value the asset above its fundamentals if they think that they are able to resell the assets later for sufficiently large capital gains. This is conceptually identical to the definition of speculative bubble formulated in Stiglitz (1990). The term of $B(e^{\gamma D}R)^{\varphi}$ with $\varnothing > 0$ is usually ruled out especially when considering an investment in real assets such as factory, railroad and so on. It is not, however, clear that we can simply rule out the speculative component in investment in housing because many investors in housing markets are simply motivated by an anticipation that there are opportunities of reselling the house for capital gain. Hence, the term of $B(e^{\gamma D}R)^{\varphi}$ with $\varnothing > 0$

remains in this analysis for the identification of speculative bubbles. To obtain the optimal investment threshold $R^*$and the option constants A and B, the appropriate boundary conditions when the housing cost is K are presented as follows:

$$AR^{\beta} = \frac{(1-\tau)\,R}{\rho-\alpha} - \frac{rm}{\rho} - \delta e^{\gamma D/m} + B\left(e^{\gamma D}R\right)^{\varphi} - K \tag{18.6}$$

$$\beta AR^{\beta-1} = \frac{(1-\tau)}{\rho-\alpha} + \varphi B(e^{\gamma D}R)^{\varphi-1}e^{\gamma D} \tag{18.7}$$

$$0 = -\frac{\delta\gamma}{m}e^{\gamma D/m} + \varphi B\left(e^{\gamma D}R\right)^{\varphi-1}\gamma e^{\gamma D}R. \tag{18.8}$$

The Eq. (18.6) represents the value-matching condition in which the option value of investment in housing equals the fundamental value of housing. The high-contact conditions are (18.7) and (18.8) with respect to R and D, respectively (Dumas 1991). It is possible to obtain the exact identifications on A, B and R^*, since there are three equations that must be simultaneously solved. By solving (18.6) through (18.8), the optimal threshold R^* is provided as follows:

$$R^* = \left[K + \frac{rm}{\rho} + \delta e^{\gamma D/m}\frac{m\gamma\varphi-\rho}{m\gamma\varphi}\left(\frac{\beta-\varphi}{\beta}\right)\right]\left(\frac{\rho-\alpha}{1-\tau}\right)\left(\frac{\beta}{\beta-1}\right) \quad . \tag{18.9}$$

The comparative statics provides intuitively consistent outcomes: for example, the negative factors in the investment decision such as the investment cost K or the rental income tax τ reduces the investment incentives. This is confirmed by an increase of the level of threshold R^* which delays the investment timing. Finally, the speculative bubble is identified as a function of the rental income:

$$S\left(R\right) = B\left(R^*\right)\left(e^{\gamma D}R\right)^{\varphi} = \frac{\rho\delta e^{\gamma D}}{m\gamma\varphi}\left(\frac{R}{R^*}\right)^{\varphi}. \tag{18.10}$$

There exists an inverse relationship between the speculative bubble and the mortgage level ($\partial S/\partial m < 0$), implying that the gains from speculative investment would diminish as the mortgage ratio increases. Now we are in a position to empirically examine the scale of speculative bubbles by looking into housing market in Seoul Metropolis.

18.3 Numerical Illustration

This section provides an empirical analysis of speculative bubbles in the housing market in Seoul. Notably, similar to many mega cities in developed regions, the housing cost in Seoul has recently experienced rapid increase especially since the early 2010s. In response to the soaring housing price, the government introduced several regulatory policies to control the market, particularly targeting the presumed speculative hot zones. Two policies of LTV (loan-to-value) and DTI (debt-to-income) are most influential tools in controlling the mortgage financing size. The LTV ratio that is defined as the ratio of a loan to the value of an asset purchased plays a role to curtail economically unhealthy loans. Similarly, the DTI ratio that provides a measure for an individual's financial ability to manage monthly payment under mortgage can limit the houseowners to rely on more mortgage than their financial capability. The ratios of LTV and DTI in three districts of Gangnam, Seocho and Songpa that were classified as speculative hot zones were substantially lowered compared to other districts.

The database of House Price provided by the Ministry of Land, Infrastructure and Transport is used to calibrate the speculative bubbles across districts.[2] The data covers the year 2015 when the growth rates of house price in Seoul recorded high. Detail information is available on the full address, price and size of house, contract date, status of monthly payment or lump-sum deposit contract (Jeon-se), down payment, the year when the apartment was built and the story level. As the apartment is the major target for the housing investment in South Korea, the following analysis strictly focuses on the apartment complex data while the data on single-family residential house is not included. The contracts between houseowner and tenants could use the lump-sum deposit (Jeon-se) or monthly-payment system. Every lump-sum contract is converted to the monthly-payment system to be consistent with the developed model above.

For the total 4,546 observations, the average unit price is 6.45 million Korean won per m^2 and the standard deviation is 2.90 million Korean won. The annual rental income is calibrated by considering the down payment and monthly payment. For some of the data, only the house price is available since the transaction is based on buy-and-sell and thus there is no monthly payment information. In this case, the monthly payment is calibrated via the interest rate for the potential rental income. Based on Lee and Choi (2014), the annual growth rate of rental income, α, is assumed to be 0.09 while allowing a range between 0.07 and 0.09 for sensitivity analysis. The discount rate r is 0.1. Note that the LTV rules for the maximum mortgage rate in 2015 required 30% in the speculative hot zones and 79% for the rest of districts. Hence, the level of m is set to be consistent with each zone's LTV rule. The housing investment cost K that should be paid by the investor upon the transaction is determined from the unit price less the mortgage. Unfortunately, due to the protection of private information, the data on the mortgage period for the individual house and the initial debt status is not available. Hence, it is assumed that the initial debt is zero and 0-year basis contract is applied on the basis of the traditional practice in the mortgage market ($D/m \approx 10$). Similar to Mason (2001), γ is set to be $r - \alpha$ and $\delta = rm/\rho$ to imply that the term $\delta e^{\gamma D/m}$ represents the opportunity cost of the mortgage as long as the mortgage is not completely cleared during the contract period. The bubble scale is defined as the bubble ratio to the house price. In the following simulation,

[2]For more information on the data, please visit: http://rt.molit.go.kr.

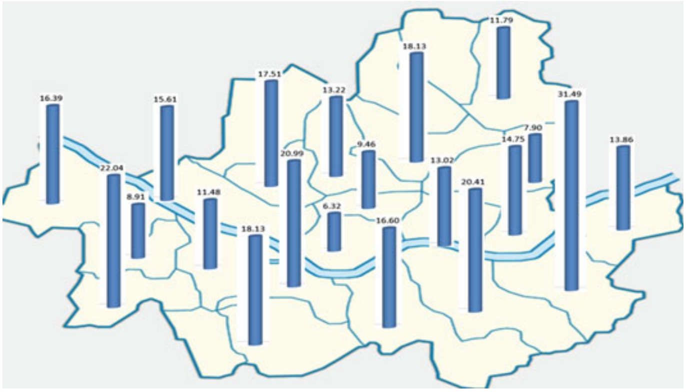


Fig. 18.1 The bubble scale map (2015)

the bubble scale for the individual house is computed using (18.9) and (18.10) after which the average bubble scale for each district is also computed.

The result supports prior intuition: the larger bubble scale is observed in the speculative hot zones compared to other districts. When the drift rate is $\alpha = 0.09$, the bubble scales in the hot zones are 20.41% in Gangnam, 16.60% in Seocho, and 31.49% in Songpa. The bubble scales in the non-hot zones are moderately lower than those in the hot zones but not significantly: for example, 13.86% in Gangdong, 18.13% in Gwanak and 14.75 in Gwangjin (see Fig. 18.1).

One notable feature in the speculative bubble map is the wide distribution of bubbles not only in the speculative hot zones but also in the non-hot zones. The northern regions over the Han River (Gang-buk) are no exceptions to bubbles in real estate markets. Such rise in the apartment price may be explained by the shortage of apartment supply in the region. However, the diffusion of psychological fear on the real estate market is another critical factor for the speculative purchase and thereby creates the bubbles. To examine the speculative bubbles in a more local level for the speculative hot zones, the zip code information is transformed to the information on latitudes and longitude. The resulting GIS (Geographic Information System) mapping is illustrated in Fig. 18.2 which shows the speculative bubbles in Gangnam, Seocho and Songpa. The circle size represents the relative scale of bubble ratios in each observation. Figure 18.2 substantiates the evidence that the bigger size bubbles are more likely clustered in Gangnam and Songpa compared to Seocho.

Figure 18.3 shows the relationship between the ratio of bubble to apartment price (S/P) and the ratio of rental value to the apartment price (R/P) with the drift rate, $\alpha = 0.07$ and $\alpha = 0.09$, respectively. The bubble ratios are positively associated with the rent value ratio, implying that the small gap between the house price and rental value would create more speculative investment in the housing market. This has a particular meaning in South Korea, because the so-called 'gap-investment' has become widely popular in recent years: the investor purchases the house and makes a Jeon-se contract which enables him to buy the asset even with small budget. The

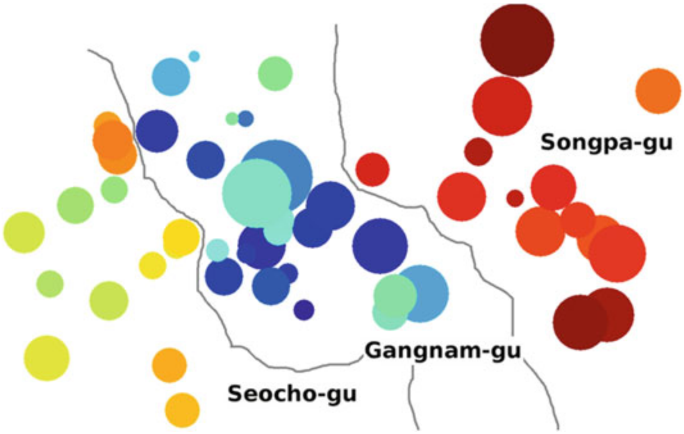


Fig. 18.2 GIS map on the bubble scale in the hot zones

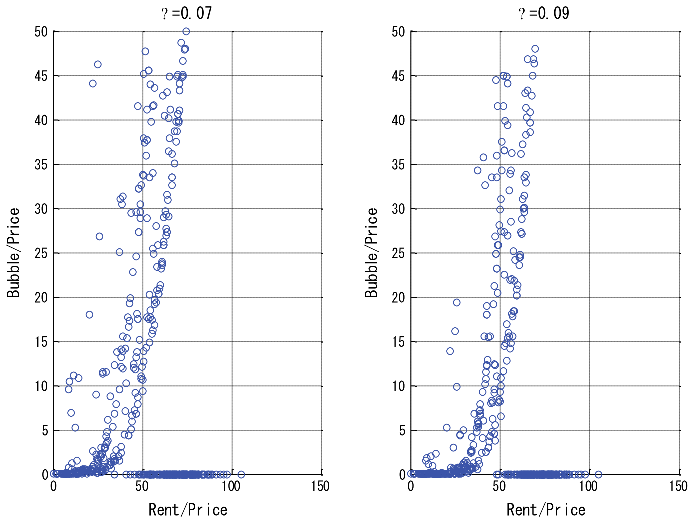


Fig. 18.3 The relationship between rent/price and bubble/price

gap-investment as a form of buy-and-lease strategy is one of culprits that disturbed the real estate market in South Korea, in particular, during the period of low interest rate.

There may be, however, controversial issues on the relationship between the bubble and the ratio of rent value to the price. One could claim that the speculative bubble is more likely to be present when the rent value closes to the apartment price, as the fear of consequent price increase would result in further speculative incentives to hold the asset. On the other hand, when there exists strong speculative incentive in the bullish housing market, the apartment price would go up while the rent

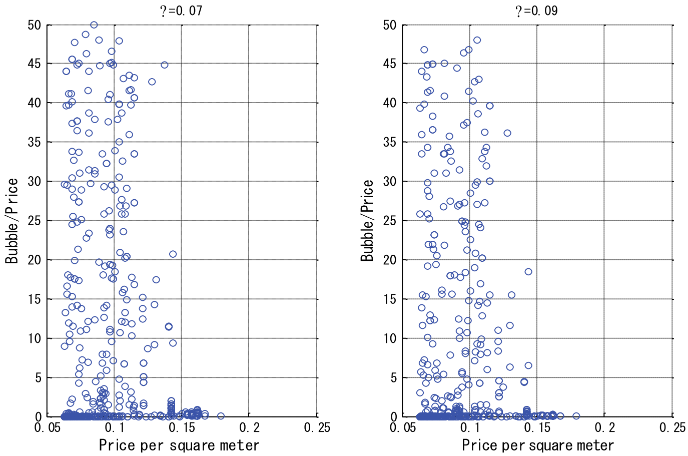


Fig. 18.4 The relationship between unit price and bubble/price

Table 18.1 Regression results on the bubble ratios

	α= 0.07			α= 0.09		
Variables	Coeff.	t-value	p-value	Coeff.	t-value	p-value
Constant	8.7432	3.1105	0.0020	12.7090	3.4278	0.0000
Zone_dummy	5.7265	4.5330	0.0000	4.5511	2.7621	0.0059
House_size	−0.2575	−9.5238	0.0000	−0.2490	−7.0739	0.0000
Floor	0.0277	0.2921	0.7703	−0.0398	−0.3228	0.7417
Apt_age	0.0522	0.8389	0.0091	0.1021	1.2601	0.2082
Rent_ratio	0.4907	16.1596	0.0000	0.5414	13.6952	0.0000
	$R^2 = 0.3784$			$R^2 = 0.2894$		

value relatively remains as it is. This implies that the rent value is largely determined by the real demand while the apartment price is partially influenced by speculative demand which is not necessarily related with the real demand. Hence, the relationship between the speculative ratio and the rent value to the asset price is not bound to straightforward conclusion. Figure 18.4 which shows the relation between S/P and price per square meter indicates no critically identifiable pattern between them.

Table 18.1 reports the regression results on the bubble ratio with several explanatory variables. The set of variables includes the Zone_dummy with the value 1 (0) when the zone belongs to the speculative hot zone (the non-hot zone) regulation. The variable House_size represents the size of apartment and the variable Floor is the story level of the apartment. Apt_age indicates the degree of aging of the apartment, the variable introduced hypothetically considering that the older apartment is more likely to be subject to the reconstruction planning, thereby creating speculative incentives for holding.

Table 18.2 Quantile regress results on the bubble ratios

Quantiles	0.1	0.4	0.7	0.9
Constant	0.0836 (0.01557)	2.6485 (0.1923)	13.7633 (1.8120)	7.9018 (3.6294)
Zone_dummy	0.0011 (0.0003)	0.0321 (0.0051)	0.0986 (0.0526)	0.0750 (0.1037)
House_size	−0.0074 (0.0041)	−0.3895 (0.0569)	−2.8175 (0.5128)	−2.3902 (1.0063)
Apt_age	−0.0080 (0.0005)	−0.0078 (0.0058)	−0.0009 (0.00008)	0.0063 (0.0447)
Rent_ratio	0.00001 (0.0000)	0.0008 (0.0023)	0.5194 (0.0310)	0.9844 (0.0600)

The value inside ( ) denotes the standard error

It is intuitive to see that the zone dummy and the ratio of rent to the house price are positively related to the bubble ratio, that is consistent with the evidence in the previous discussions. The reason for that the house size has a negative association with the bubble ratio reflects the recent demand pattern changes in favor of the medium or small size apartment under aging society and the development of single households. The influence of level of story on the bubble is not statistically verified. Lastly, the age of apartment is positively associated with the bubble ratio, since the complex of aged apartments is the first target for the remodelling or reconstruction city schedule.

As seen in Fig. 18.3, the distribution of bubble has some outliers which require more robust estimation conditional on the quantile. Thus, the quantile regression that is useful in treating uneven distributions is conducted with respect to the zone dummy and rent ratio variables. As defined in Angrist and Pischke (2009), the conditional quantile function (CQF) at quantile τ given explanatory variable x_i is defined as $Q_\tau\,(y_i \mid x_i) = F_y^{-1}(\tau|x_i)$ where $F_y(\tau|x_i)$ is the distribution function for y_i at y, conditional on x_i. Table 18.2 reports the results of quantile regressions of the bubble scale with respect to the zone dummy and rent ratio variables for the drift rate α= 0.07. The finding is that the group of high level of speculative bubbles is well explained by the employed explanatory variables, with results being consistent with those of Table 18.1.

Figures 18.5 and 18.6 illustrate the sensitivity analysis of $R^*$and $S(R)$, respectively, with respect to the investment cost K and the discount rate ρ. Hypothetical parameter values are adopted to conduct the simulations. As noted above, R^* is the critical threshold above which it is economically optimal to purchase the apartment even while considering the resales option as a form of speculative bubbles. Figure 18.5 shows that the investment threshold increases as the cost of investment and the cost of capital increase, implying that the investment hysteresis resulting from the cost delays the purchase timing. This finding is consistent with conventional real options literature.

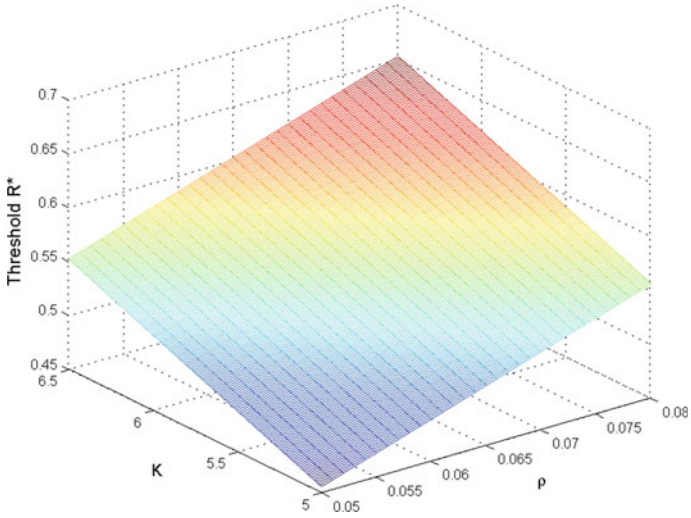


Fig. 18.5 Sensitivity analysis of R^* with respect to K and ρ

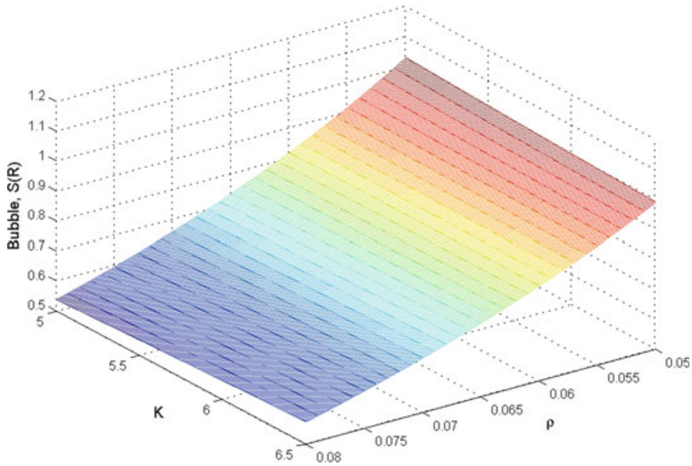


Fig. 18.6 Sensitivity analysis of $S(R)$ with respect to K and ρ

Figure 18.6 shows the relative insensitivity of speculative bubble with respect to K whereas the interest rate plays a significant role in determining the bubble. The result indicates that the bubble is pervasive when the interest rate is low because of the low cost of mortgage finance. Figure 18.7 provides a simulated illustration on the sensitivity of the ratio of bubble to house price with respect to the investment cost and mortgage size. It is intuitive to see that the bubble size relative to the house price increases as the size of mortgage finance increases. This is a supportive finding for the need of strengthening LTV and DTI limits to curtail speculative housing demand.

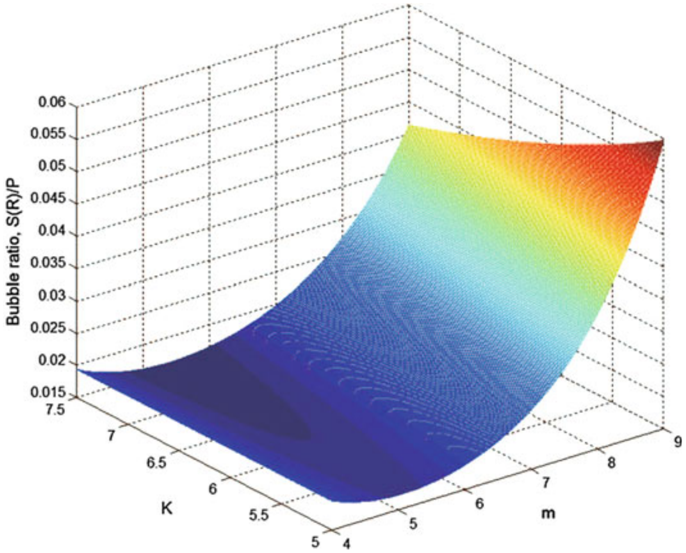


Fig. 18.7 Bubble ratio, $S(R)/P$ with respect to K and ρ

18.4 Conclusion

There is growing demand for the research on speculative bubbles in housing markets as we recently observe boom and bust cycles in the market. The study on the speculative market behaviour carries a critical significance related not only to the real estate market but also to sustainable urban management, because the rising housing costs raises a voice to develop the green belts regions for the purpose of house supply. There would not be many controversial issues in speculative studies had there been unequivocal and standard definitions of bubbles. Nevertheless, as noted in Turvey (2002), a real option modelling is regarded as a proper approach to identifying the speculative bubbles in the real estate market.

This paper presents an individual's dynamic investment model in the housing market when the speculative motives in housing investment is significant. The speculative bubble defined in the paper is conceptually based on a put option. The derived value function consists of the component of present value of the real asset classically defined as well as the speculative component as a form of resale option. Using the 2015 Transaction Price DB by the Ministry of Land, Infrastructure and Transport, the regional speculative bubbles were calibrated. The bubble scales in the speculative hot zones are relatively higher: 20.41% in Gangnam, 16.60% in Seocho, and 31.49% in Songpa. Additional interesting finding is that the speculative bubble in Seoul is substantially prevalent not only in the three hot zones but also in the rest of regions. The shortage of apartment supply is of course one of the reasons for the rising speculative motives in Seoul. However, somewhat fearful preemptive decision in housing investment particularly when the interest rate is low and the limit of mortgage finance is lax is more critical to the widespread diffusion of speculative bubbles. Simulated illustrations on the sensitivity of the ratio of bubble to the house price with respect to the investment cost and mortgage size indicates that the ratio of bubble to the apartment price is positively associated with the size of mortgage finance. Significant presence of speculative bubbles suggests the need of more restrictive LTV and DTI rules to curtail speculative housing demand. However, the government policy was rather in an opposite track: the deregulation policy in 2016 relaxed LTV rules in an attempt to boost economy. It is obvious to see that such policy failure brought about further incentives on the gap-investment, i.e., buy-and-lease strategy at the low interest rate. This implies that the house supply through the greenbelt zone lift may not be sufficient unless more restrictive mortgage rule is adopted to curb the rising housing prices.

The paper casts a caveat regarding the methodological framework for the bubbles study since it is still questionable as to the model approach in identifying the bubble that may be not empirically fully identifiable. The finding in the paper, nevertheless, will be one part of contributions to the relevant literature for better understanding and designing the real estate policy. The paper needs to be further extended to the future research by consolidating more recent data on the housing market to address the impact study of policies that were introduced to relax the LTV/DTI rules.

References

Angrist, J.D., and J.-S. Pischke. 2009. *Mostly harmless econometrics*. Princeton: Princeton University Press.

Baker, Timothy G., Edward H. Ketchabaw, and Calum Greig Turvey. 1991. An income capitalization model for land value with provisions for ordinary income and long-term capital gains taxation. *Canadian Journal of Agricultural Economics* 39: 69–82.

Brock, W.A., and C.H. Hommes. 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control* 22: 1235–1274.

Buchsteiner, H., and K. Zavodov. 2012. Bubbles in open economics: Theory and empirical detection. Real Option 2012 London Conference.

Dixit, A., and R. Pindyck. 1994. *Investment under Uncertainty*. Princeton: Princeton University Press.

Duffie, D., N. Garleanu, and L.H. Pedersen. 2002. Securities lending, shorting, and pricing. *Journal of Financial Economics* 60: 307–339.

Dumas, B. 1991. Super contact and related optimality conditions. *Journal of Economic Dynamics and Control* 15: 675–695.

Ekstrom, E., and J. Tysk. 2009. Bubbles, convexity and the Black-Scholes equation. *Institute of Mathematical Statistics* 19 (4): 1369–1384.

Flood, R., and R. Hodrick. 1990. On testing for speculative bubbles. *The Journal of Economic Perspectives* 4 (2): 85–101.

Goodman, A., and T.G. Thibodeau. 2008. Where are the speculative bubbles in US housing market? *Journal of Housing Economics* 17: 117–137.

Granziera, E., and S. Kozicki. 2015. House price dynamics: Fundamental and expectations. *Journal of Economic Dynamics and Control* 60: 152–165.

Harrison, M., and D.M. Kreps. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* 92: 323–336.
Himmelberg, C., C. Mayer, and T. Sinai. 2005. *Assessing high house price: Bubbles, fundamentals, and misperceptions*. National Bureau of Economic Research.
Kashiwagi, M. 2014. A search-theoretic model of the rental and homeownership markets. *Journal of Housing Economics* 26: 33–47.
Kashiwagi, M. 2014. Sunspots and self-fulfilling beliefs in the U.S. housing market. *Review of Economic Dynamics* 17 (4): 654–676.
Kim, Moonyeun, and Youngman Lee. 2015. An analysis of timing and hazard rate of arrears and default in residential mortgages. *The Korea Spatial Planning Review* 87: 17–32.
Kivedal, B.K. 2012. Testing for rational bubbles in the housing market. Norwegian University of Science and Technology, working paper No. 10.
Kouwenberg, R., and R.C.J. Zwinkels. 2015. Endogenous price bubbles in a multi-agent system of the housing market. *PLos One* 10 (6): e0129070.
Lee, O.D., and J.I. Choi. 2014. Analysis of increase rate of housing prices of Seoul and six other metropolitan cities. *Korean Management Consulting Review* 14 (5): 313–333.
Mason, C.F. 2001. Nonrenewable resources with switching costs. *Journal of Environmental Economics and Management* 42: 65–81.
Shiratsuka, S. 2005. The asset price bubble in Japan in the 1980s: Lessons for financial and macroeconomic stability. *Real Estate Indicators and Financial Stability*. Basel: Bank for International Settlements.
Scheinkman, J.A., and W. Xiong. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* 111 (6): 1183–1219.
Park, H. 2012. Real option analysis for effects of emission permit banking on investment under abatement cost uncertainty. *Economic Modelling* 29: 1314–1321.
Park,. 2015. Adaptive belief system on dynamic inconsistency of property tax and complexity in housing market with heterogeneous agent. *Applied Economics Journal* 17 (4): 5–24.
Pindyck, R.S. 2002. Optimal timing problems in environmental economics. *Journal of Economic Dynamics and Control* 26: 1677–1697.
Stiglitz, J. 1990. Symposium on bubbles. *The Journal of Economic Perspectives* 4 (2): 13–18.
Turvey, C. 2002. Can hysteresis and real options explain the farmland valuation puzzle. Working paper, University of Guelph.
Wang, K., and Y. Zhou. 2006. Equilibrium real option exercise strategies with multiple players: The case of real estate markets. *Real Estate Economics* 34 (1): 1–49.
Wigniolle, B. 2014. Optimism, pessimism and financial bubbles. *Journal of Economic Dynamics and Control* 41: 188–208.

Part IV
Resource and Environmental Economics

Chapter 19
CO_2 Emissions, Energy Consumption, GDP, and Foreign Direct Investment in ANICs Countries

Suyi Kim

19.1 Introduction

Hong Kong, Singapore, South Korea, and Taiwan are called the Asian Newly Industrialized Countries (ANICs), which were famous for maintaining exceptionally high growth rates (in excess of 7% per a year) and attaining rapid industrialization between the early 1960s (mid-1950s for Hong Kong) and the 1990s. These countries have a comparative advantage in producing different types of products at different stages in their economic development. As such, their economic success stories have served as role models for numerous developing countries.

Particularly, energy consumption has rapidly increased in proportion to economic growth while these countries have achieved outstanding economic growth. This rapid energy consumption has also caused a rapid increase in the greenhouse gases (GHG). Additionally, the inward or outward foreign direct investment (FDI) has also increased rapidly due to the change in the comparative advantage of the manufacturing industry and their domestic economic development policy. Hence, these FDIs have contributed to the economic growth of these areas. However, research on economic growth, energy consumption, GHG and FDI for the ANICs is relatively lacking. In particular, the causal relationship between energy consumption, GHG and FDI in the process of achieving rapid economic growth is in question. These causal relationships could supply implications for other countries ahead of their future economic development. Specifically, Korea has a GHG emissions reduction target of 37% compared to Business as Usual by 2030. Singapore intends to reduce its Emission Intensity by 36% by 2030 from its 2005 levels, and stabilize its emissions with the aim of peaking around 2030. Taipei has a GHG emission reduction target of reducing community-wide CO_2-equivalent emissions by 25% from 2005 to 2030, and community-wide CO_2-equivalent emissions by 60% from 1990 to 2050. Hong

S. Kim (✉)
Hongik University, Sejong, South Korea
e-mail: suyikim@hongik.ac.kr

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_19

Kong has a plan to reduce community-wide CO_2 equivalent emissions intensity by 55% per GDP HKD by 2020 from the base year 2005. Consequently, these emission reduction targets may influence the economic growth and energy consumption of the region through the causal relationship between these variables.

The causality relationships between CO_2 emissions, energy consumption, GDP and FDI have been a subject of debate in empirical and theoretical literature for the past twenty years. A first category of literature focuses on the causal relationships between CO_2 emissions and economic growth which has been investigated under the environment Kuznets curve (EKC) that demonstrates a U-shaped relationship between economic growth and environmental pollution emissions. The pollution levels increase with national income and subsequently decrease after the national income has achieved a certain income. A series of empirical studies have verifies the EKC hypothesis. Hettige et al. (1992), Cropper and Cropper and Griffths (1994), Selden and Song (1994), Grossman and Krueger (1995), Martinez-Zarzoso and Bengochea-Morancho (2004) support the EKC hypothesis. Holtz-Eakin and Selden (1995), and Shafk (1994) demonstrate a linear relationship between national income and environmental pollution levels.

The second category of literature stream investigates the relationship between energy consumption and economic growth. This relationship has important implications for energy policy (Jumbe 2003; Shiu and Lam 2004; Yoo 2005; Chen et al. 2007; Mozumder and Marathe 2007; Squalli 2007; Apergis and Payne 2009a). Ozturk (2010) categorizes the causal relationship between energy consumption and economic growth into four types: neutrality, conservative, growth, and feedback hypothesis. Recent studies comprehensively consider the causal relationship between energy consumption and economic growth, and between GHG emissions and economic growth. For example, Ang (2007) examined the dynamic causal relationship between CO_2 emissions, energy consumption and GDP in France and employed a cointegration and vector error correction model (VECM) for the period 1960–2000, resulting in long-run causal relationships from economic growth to energy consumption and also from economic growth to CO_2 emissions. As such, economic growth exerts a causal influence on the expansion of energy use and CO_2 emissions in the long run. In the short run, the causality runs from the energy use growth to output growth. Since Ang's (2007) work, there have been several studies on the dynamic causal relationship, and the results may appear different for different target countries (Halicioglu 2008; Zhang and Cheng 2009; Apergis and Payne 2009b; Soytas and Sari 2009; Pao and Tsai 2010; Arouri et al. 2012).

Table 19.1 presents an overview of these comprehensive studies on the causal relationship between CO_2 emissions, energy consumption and economic growth. For example, Halicioglu (2008) demonstrated the existence of bidirectional Granger causalities between CO_2 emissions and energy consumption and between CO emissions and income. Zhang and Cheng (2009) indicated the existence of unidirectional Granger causalities running from GDP to energy consumption and running from energy consumption to carbon emissions in the long run for China, but neither carbon emissions nor energy consumption cause economic growth. For the BRIC countries, Pao and Tsai (2010) demonstrated the existence of bidirectional strong causalities

Table 19.1 Overview of comprehensive studies on the causal relationship among CO_2 emissions, economic growth, and energy consumption

Authors	Period	Country	Methodology	Causality relationship
Ang (2007)	1960–2000	France	Cointegration, Vector Error Correction Model	GDP → EC (long run) GDP → CO_2 (long run) EC → GDP (short run)
Halicioglu (2008)	1960–2005	Turkey	Cointegration Granger Causality ARDL	CO_2 ↔ EC CO_2 ↔ GDP
Zhang and Cheng (2009)	1960–2007	China	Granger Causality	GDP → EC EC → CO_2
Apergis and Payne (2009b)	1971–2004	7 Central America (Costa Rica, El Salvador, Guatemala, Honduras, Panama)	Panel Vector Error Correction Model	EC → CO_2 (short run) GDP → CO_2 (short run) EC ↔ GDP (short run) EC ↔ CO_2 (long run)
Soytas and Sari (2009)	1960–2000	Turkey	Granger Causality	CO_2 → EC
Pao and Tsai (2010)	1971–2005	BRIC countries (Brazil, China, India, Russia)	Panel Cointegration Granger Causality Panel VECM	CO_2 ↔ EC (long run) GDP ↔ EC (long run) EC → GDP (short run) CO_2 → GDP (short run)
Arouri et al. (2012)	1981–2005	12 Middle East and North African Countries (MENA)	Bootstrap panel unit root tests, Cointegration	GDP → CO_2 (short run) EC → CO_2 (short run)

between energy consumption and emissions, and between energy consumption and output in the long run, along with unidirectional short run causalities from emissions and energy consumption to output, respectively. For 12 Middle Eastern and North African Countries (MENA), Arouri et al. (2012) demonstrated the existence of short run causality from energy consumption to CO_2 emissions and also from GDP to CO_2 emissions. Salahuddin (2015) investigated the relationship between carbon dioxide emissions, economic growth, electricity consumption and financial development in the Gulf Cooperation Council (GCC) countries. Shahbaz et al. (2014) explored the relationship between economic growth, electricity consumption, and environmental degradation in the case of the United Arab Emirates (UAE).

The last category of the literature stream investigates the relationships between CO_2 emissions, energy consumption, GDP and FDI, comprehensively. The impact of FDI on the host country's environment has also been a subject of debate. As such, two conflicting hypotheses have been presented in the previous studies: the pollution haven and the halo effect hypotheses. According to the latter, the presence of foreign investors will spur positive environmental spill-overs to the host country because MNCs multinational companies (MNCs) have more advanced technologies than their domestic counter parts and tend to disseminate cleaner technology which is less harmful to the environment. In contrast, the pollution haven hypothesis postulates that MNCs will lock more into countries where environmental regulations are less strict. As such, this strategy might harm the environment in the host country if the issue is not taken seriously. However, the results are both theoretically and empirically mixed.

Merican et al. (2007) investigate the causal link between FDI and pollution by employing an autoregressive distributed lag (ARDL) approach. According to the results, FDI increases emissions in Malaysia, Thailand, and the Philippines, while there seems to be an inverse relationship between the variables in Indonesia. Hoffmann et al. (2005) found unidirectional causality running from FDI to CO_2 emissions. Pao and Tsai (2010) examine the causal links between CO_2 emissions, energy consumption, FDI, and GDP in the BRIC countries (Brazil, Russia, India and China) using a multi variate Granger causality approach. According to their results, there is bi-directional causality between emissions and FDI, and a one-way causality running from output to FDI.

Hence, this paper analyses the causal relationships between GHG emissions, energy consumption, economic development, and FDI in ANICs countries. To the best of our knowledge, this is the first study to explore the causal links between the four variables in these countries using VECM from 1971 to 2011. This issue is particularly important because these countries experienced high economic growth over past 40 years. Therefore, understanding the causality relationships between the variables will help policy-makers in designing appropriate policies. The remainder of the study is structured as follows: Sect. 19.2 presents the methodologies employed in this study, Sect. 19.3 reports on data, and Sect. 19.4 discusses the empirical findings, and finally, Sect. 19.5 concludes the paper.

19.2 Empirical Model

Following Kivyiro and Arminen (2014), the long run relationship between CO_2 emissions, energy consumption, output and FDI is modelled as indicated by Eq. (19.1) below. According to the EKC hypothesis, there is an inverted U shaped relationship between environmental pollution and output. We can apply this relationship between CO_2 emissions and GDP, and express it mathematically by including the squared value of GDP per capita in the set of regressors.

$$CO_{2,it} = \alpha_0 + \alpha_1 EN_{it} + \alpha_2 FDI_{it} + \alpha_3 Y_{it} + \alpha_4 Y_{it}^2 + V_i + \epsilon_{it} \quad (19.1)$$

where $i(i = 1, 2, \ldots, N)$ denotes the country and $t(t = 1, 2, \ldots, T)$ denotes the period. CO_2 denotes the CO_2 emissions per capita, EN is the energy consumption per capita, and FDI denotes foreign direct investment which can be divided into inward FDI (IFDI) and outward FDI (OFDI). Y denotes GDP per capita, and Y^2 denotes GDP per capita squared. All the variables are in their natural logarithmic form. Individual fixed country effects are denoted by V, and ϵ denotes the stochastic error term.

First, we test whether these time series have unit roots. If so, we use panel cointegration techniques to investigate the relationship. Panel estimation techniques are attractive because models estimated from cross-sections of time series have more degrees of freedom and efficiency than models estimated from individual time series. This is particularly useful if the time series dimension of each cross-section is short. Panel cointegration techniques have recently been used by a number of authors to investigate the relationship between energy consumption and output (e.g. Hoffmann et al. 2005; Chen et al. 2007; Mahadevan and Asafu-Adjaye 2007; Narayan 2005; Mehrara 2007; Lee and Chang 2008; Lee et al. 2008; Narayan and Smyth 2008; Apergis and Payne 2009b; Sadorsky 2009a, Sadorsky 2009b; Apergis and Payne 2010; Sadorsky 2011).

If each time series are I (1) and these variables are cointegrated, a panel VECM can be used to estimate causality, as in Engle and Granger (1987). Finding cointegration between groups of variables is very important because it ensures that an error correction mechanism exists according to which changes in the dependent variable are modelled as a function of the level of the equilibrium in the cointegration relationship and changes in other explanatory variables. Equation 19.1 can be written as the following VECM model.

$$\begin{aligned}
\Delta CO_{2,it} &= c_{1i} + \sum_{j=1}^{q} \gamma_{11ij} \Delta CO_{2,it-j} + \sum_{j=1}^{q} \gamma_{12ij} \Delta EN_{it-j} + \sum_{j=1}^{q} \gamma_{13ij} \Delta FDI_{it-j} \\
&\quad + \sum_{j=1}^{q} \gamma_{14ij} \Delta Y_{it-j} + \sum_{j=1}^{q} \gamma_{15ij} \Delta Y^2_{it-j} + \gamma_{16i} \varepsilon_{1t-1} + \upsilon_{1it}
\end{aligned}$$

$$\begin{aligned}
\Delta EN_{it} &= c_{2i} + \sum_{j=1}^{q} \gamma_{21ij} \Delta CO_{2,it-j} + \sum_{j=1}^{q} \gamma_{22ij} \Delta EN_{it-j} + \sum_{j=1}^{q} \gamma_{23ij} \Delta FDI_{it-j} \\
&\quad + \sum_{j=1}^{q} \gamma_{24ij} \Delta Y_{it-j} + \sum_{j=1}^{q} \gamma_{25ij} \Delta Y^2_{it-j} + \gamma_{26i} \varepsilon_{1t-1} + \upsilon_{2it}
\end{aligned}$$

$$\begin{aligned}
\Delta FDI_{it} &= c_{3i} + \sum_{j=1}^{q} \gamma_{31ij} \Delta CO_{2,it-j} + \sum_{j=1}^{q} \gamma_{32ij} \Delta EN_{it-j} + \sum_{j=1}^{q} \gamma_{33ij} \Delta FDI_{it-j} \\
&\quad + \sum_{j=1}^{q} \gamma_{34ij} \Delta Y_{it-j} + \sum_{j=1}^{q} \gamma_{35ij} \Delta Y^2_{it-j} + \gamma_{36i} \varepsilon_{1t-1} + \upsilon_{3it}
\end{aligned}$$

$$\begin{aligned}
\Delta Y_{it} &= c_{4i} + \sum_{j=1}^{q} \gamma_{41ij} \Delta CO_{2,it-j} + \sum_{j=1}^{q} \gamma_{42ij} \Delta EN_{it-j} + \sum_{j=1}^{q} \gamma_{43ij} \Delta FDI_{it-j} \\
&\quad + \sum_{j=1}^{q} \gamma_{44ij} \Delta Y_{it-j} + \sum_{j=1}^{q} \gamma_{45ij} \Delta Y^2_{it-j} + \gamma_{46i} \varepsilon_{1t-1} + \upsilon_{4it}
\end{aligned}$$

$$\begin{aligned}
\Delta Y^2_{it} &= c_{5i} + \sum_{j=1}^{q} \gamma_{51ij} \Delta CO_{2,it-j} + \sum_{j=1}^{q} \gamma_{52ij} \Delta EN_{it-j} + \sum_{j=1}^{q} \gamma_{53ij} \Delta FDI_{it-j} \\
&\quad + \sum_{j=1}^{q} \gamma_{54ij} \Delta Y_{it-j} + \sum_{j=1}^{q} \gamma_{55ij} \Delta Y^2_{it-j} + \gamma_{56i} \varepsilon_{1t-1} + \upsilon_{5it}
\end{aligned}$$

where Δ is the first difference operator, q is the lag length, Y is the natural log of real output, FDI is the natural log of FDI, CO_2 is the natural log of the CO_2 emissions, EN is the natural log of energy consumption, ε is the error correction term, and v is the random error term. The VECM is estimated using a seemingly unrelated regression (SUR) technique that allow for cross-sectional specific coefficient vectors and cross-sectional correlations in the residuals.

19.3 Data

The data on the ANICs are the annual time series covering the period from 1971 to 2011. CO_2 emissions per capita (measured in kilograms of oil equivalent per capita), energy consumption per capita (in metric tons per capita) and GDP per capita (constant 2005 USD) of Hong Kong, Korea (i.e., Republic of Korea), Singapore are taken from the World Bank (2019), World Development Indicators online database. The FDI series (in percentage of gross fixed capital) for all countries are taken from the United Nations Conference on Trade and Development (UNCTAD). The data on CO_2 emissions, energy consumption and GDP of Taiwan are from OECD's Indicators for CO_2 emissions from fuel combustion. Table 19.2 shows descriptive statistics of the actual variables used in this analysis, for each country.

Time series plots of the natural logs of CO_2 emissions per capita for each of the countries are shown in Fig. 19.1. The CO_2 emissions per capita for Hong Kong and Singapore have been increasing until the mid-1990s but have decreased and fluctuated since that. The CO_2 emissions per capita for Korea and Taiwan have been steadily increasing over time but the growth rate has decreased in recent years.

Time series plots of the natural logs of energy consumption per capita for each country are shown in Fig. 19.2. Energy consumption per capita has been increas-

Table 19.2 Descriptive statistics (before data transformation), 1971–2011

	CO_2		EN		Y	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Hong Kong	4.068	1.265	1527.415	498.267	17706.930	7843.877
Singapore	11.395	3.918	3747.677	1662.486	18262.360	9158.456
Korea	6.431	3.148	2544.210	1547.505	10470.000	6606.025
Taiwan	6.675	3.369	2654.244	1358.344	9469.980	5444.583
	IFDI		OFDI			
	Mean	S.D.	Mean	S.D.		
Hong Kong	42.070	45.775	42.663	50.414		
Singapore	39.561	29.144	18.156	22.819		
Korea	2.468	3.285	2.057	2.231		
Taiwan	2.529	1.862	4.864	4.773		

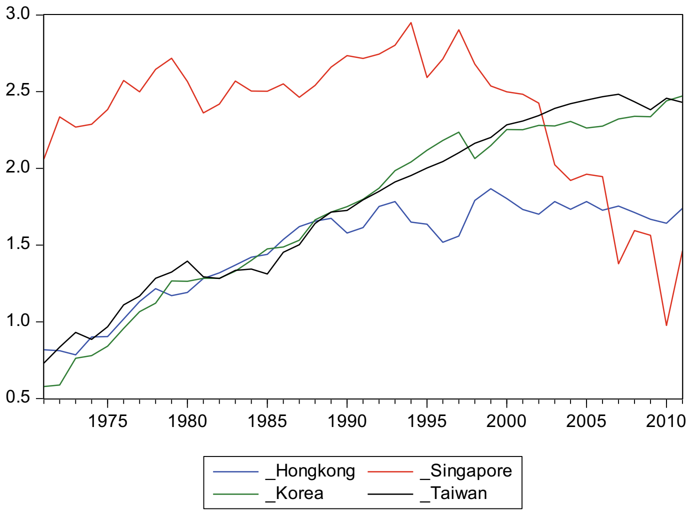


Fig. 19.1 Natural log of CO_2 emissions per capita (metric tons per capita)

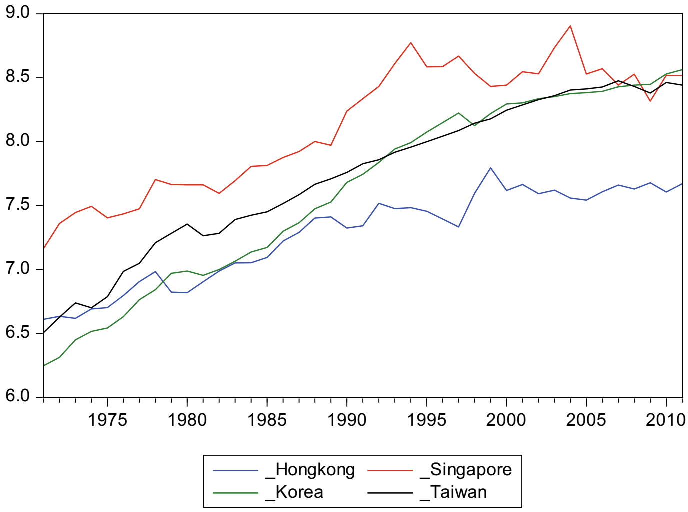


Fig. 19.2 Natural log of energy consumption per capita (kilogram of oil equivalent)

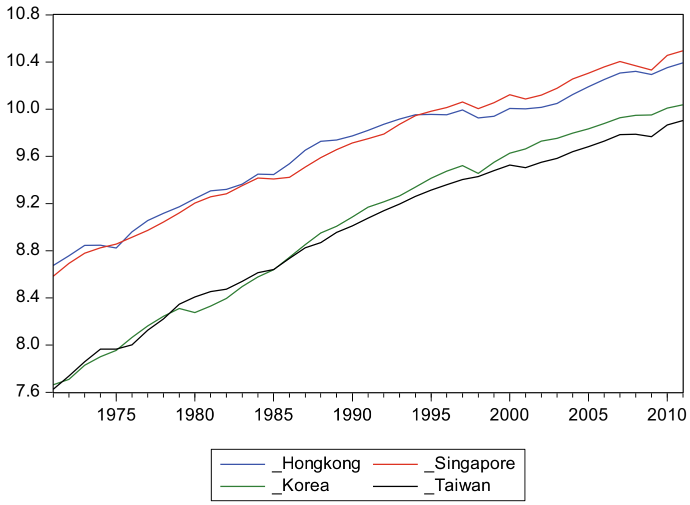


Fig. 19.3 Natural log of GDP per capita (constant 2005 USD)

ing over time, although the strength of this trend varies by country. For instance, energy consumption per capita in Korea, and Taiwan has increased steadily over time whereas in Singapore and Hong Kong it has been fluctuating since the mid-1990s.

Figure 19.3 shows time series plots of natural logs of the real GDP per capita for each country, and overall, GDP has been increasing over time. In the case of Hong Kong, Japan, Korea, and Singapore, the GDPs dropped temporarily in the mid-1990s because of the Asian financial crisis and once again temporarily at the end of the 2000s due to the global financial crisis. However, overall GDP has shown an increasing trend.

The time series plots of the natural logs of IFDI measured in percentage of gross fixed capital formation for each country are shown in Fig. 19.4. Despite fluctuating over time due to the economic conditions, the trends have been generally upward sloping with the largest in Hong Kong, followed by Singapore in recent years. The percentage of IFDI is relatively small in Korea and Taiwan. The time series plots of the natural logs of OFDI measured in percentage of gross fixed capital formation for each country are shown in Fig. 19.5. This trend has been generally upward sloping as well.

Table 19.3 shows the correlations among the panel data variables in natural log. Most of correlations are positive. The natural of CO_2 emissions is highly correlated with the natural log of energy consumption, followed respectively by correlations with the natural log of GDP, OFDI, and IFDI. The natural log of CO_2 emissions is more highly correlated with the natural of OFDI than with that of IFDI. The natural log of GDP is highly correlated with the natural log of OFDI, followed by that of IFDI, energy consumption and CO_2 emissions.

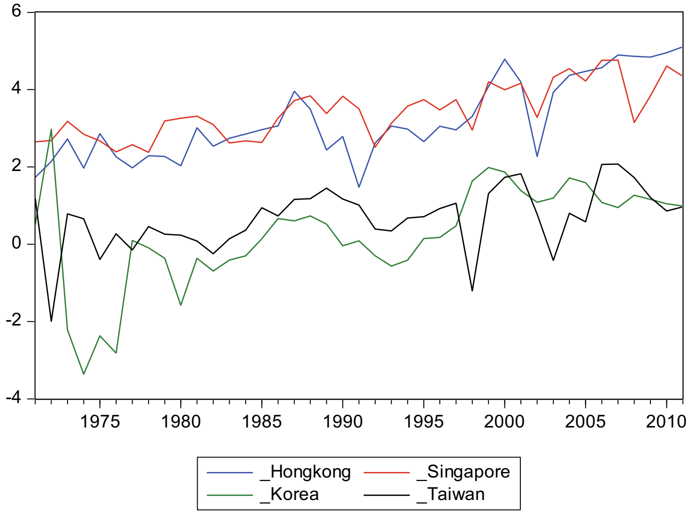


Fig. 19.4 Natural log of IFDI (percentage of gross fixed capital formation)

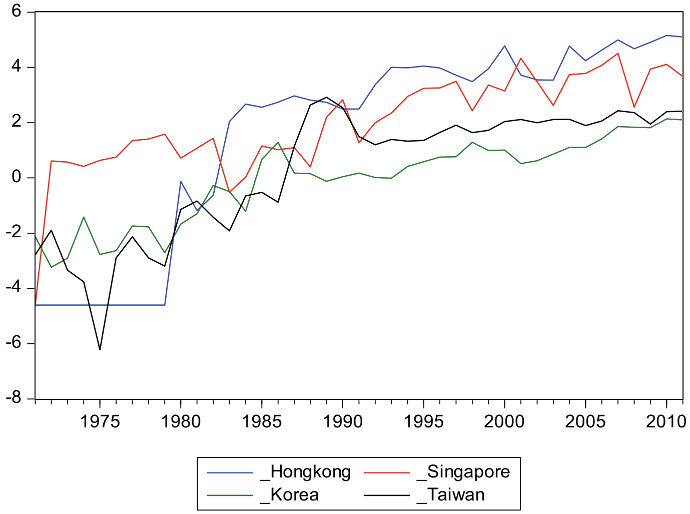


Fig. 19.5 Natural log of OFDI (percentage of gross fixed capital formation)

Table 19.3 Correlations for the panel data set (variables in natural log), 1971–2011

	CO_2	EN	Y	IFDI	OFDI
CO_2	1				
EN	0.810163	1			
Y	0.579752	0.75678	1		
IFDI	0.325305	0.33164	0.697433	1	
OFDI	0.544538	0.659935	0.824931	0.574287	1

19.4 Empirical Findings

19.4.1 Unit Root Tests

In this paper, we conduct four types of panel unit root tests that assume cross-sectional independence (Dickey and Fuller 1979; Phillips and Perron 1988; Levin et al. 2002; Im et al. 2003). For these tests, the null hypothesis is that there is a unit root while the alternative hypothesis is that there is no unit root. The result of these tests is that for each series in levels, the null hypothesis cannot be rejected at the 5% significance level. According to Table 19.4, in case of IFDI, the ADF and PP tests indicate that we can reject the null hypothesis at the 5% significance level. For each series in the first differences, the null hypothesis that there is no unit root can be rejected at the 1% significance level.

19.4.2 Cointegration Tests

We test whether these I (1) variables are cointegrated using the tests of Pedroni (1999; 2004). The Pedroni panel cointegration tests are to test the residuals from the following equation for unit root variables.

$$\hat{\varepsilon}_{it} = \rho_i \hat{\varepsilon}_{it-1} + \delta_{it}$$

Overall, Pedroni (1999; 2004) provides seven statistics for testing null hypothesis of no cointegration in heterogeneous panels. These tests can be classified as either within-dimension (panel tests) or between-dimension (group tests). For the within-dimension approach, the null of no cointegration ($\rho_i = 1$ for all i) is tested against the alternative ($\rho_i < 1$ for all i). The group means approach is less restrictive because it does not require a common value of æ under the alternative hypothesis ($\rho_i < 1$ for all i).

According to Table 19.5, the cointegration test results for a model with IFDI are mixed. Five of the within-dimension statistics indicate cointegration at the 5%

Table 19.4 Panel unit root tests

Method	CO_2		ΔCO_2		EN		ΔEN		Y		ΔY	
	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.
Null: Unit root (assumes common unit root process)												
Levin, Lin & Chu t*	−0.05	0.48	−5.31	0.00	0.71	0.76	−7.67	0.00	−0.79	0.21	−7.74	0.00
Null: Unit root (assumes individual unit root process)												
Im, Pesaran and Shin W-stat	2.03	0.98	−7.62	0.00	2.13	0.98	−6.72	0.00	1.93	0.97	−6.91	0.00
ADF—Fisher Chi-square	1.84	0.99	66.66	0.00	2.11	0.98	57.41	0.00	1.70	0.99	59.24	0.00
PP—Fisher Chi-square	1.35	0.99	95.89	0.00	1.93	0.98	101.05	0.00	2.36	0.97	74.64	0.00
Method	Y^2		ΔY^2		IFDI		ΔIFDI		OFDI		ΔOFDI	
	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.	Statistic	Prob.
Null: Unit root (assumes common unit root process)												
Levin, Lin & Chu t*	−0.38	0.35	−8.00	0.00	−0.20	0.42	−12.46	0.00	−1.48	0.07	−8.78	0.00
Null: Unit root (assumes individual unit root process)												
Im, Pesaran and Shin W-stat	1.74	0.96	−7.41	0.00					−0.07	0.47	−11.98	0.00
ADF—Fisher Chi-square	2.20	0.97	64.39	0.00	17.56	0.02	148.10	0.00	5.75	0.68	109.29	0.00
PP—Fisher Chi-square	2.61	0.96	81.78	0.00	24.76	0.00	660.45	0.00	16.71	0.03	121.74	0.00

Table 19.5 Panel cointegration tests for a model with IFDI

Alternative hypothesis: common AR coefs. (within-dimension)				
	Statistic	Prob.	Weighted statistic	Prob.
Panel v-statistic	3.793***	0.000	1.633*	0.051
Panel rho-statistic	−2.294**	0.011	−0.915	0.180
Panel PP-statistic	−3.590***	0.000	−2.148**	0.016
Panel ADF-statistic	−3.021***	0.001	−1.120	0.131
Alternative hypothesis: individual AR coefs. (between-dimension)				
		Statistic		Prob.
Group rho-statistic		−0.565		0.286
Group PP-statistic		−2.374***		0.009
Group ADF-statistic		−1.384*		0.083
Null hypothesis: no cointegration				
Trend assumption: deterministic intercept and trend				
Automatic lag length selection based on SIC with a max lag of 5				
Newey-West automatic bandwidth selection and Bartlett kernel				

*means 10% significance level
**means 5% significance level
***means 1% significance level

level, and one of them at the 10% level. In the between-dimension case, one of the statistics indicates cointegration at the 1% level and one at the 10% level. According to Table 19.6, the results for the model with OFDI are also mixed. Four of the within-dimension statistics indicate cointegration at the 1% level and one at the 5% level. In the between-dimension case, one of the statistics indicates cointegration at the 1% level and one at the 10% level.

19.4.3 VECM Tests

19.4.3.1 Short Run Dynamics

Short-run dynamics for equations with exports are estimated by Engle and Granger (1987). The vector auto regression lag length q is set at 2, determined using the Schwarz and the Hannan-Quinn information criteria. The results of the short-run Granger causality test are shown in Tables 19.7 and 19.8.

The main interest of this paper is a feedback relationship among CO_2 emissions, GDP, energy consumption, and IFDI and OFDI. Table 19.7 shows the short-run Granger causality results for VECM with IFDI. According to Table 19.4, there is some evidence of short-run causality from GDP to CO_2 emissions at the 5% significance level, from IFDI to GDP at the 1% significance level. Additionally, there is a short-run

Table 19.6 Panel cointegration tests for model with OFDI

Alternative hypothesis: common AR coefs. (within-dimension)				
	Statistic	Prob.	Weighted statistic	Prob.
Panel v-statistic	4.165***	0.000	1.579*	0.057
Panel rho-statistic	−2.614***	0.005	−0.868	0.193
Panel PP-statistic	−4.132***	0.000	−2.065**	0.020
Panel ADF-statistic	−3.299***	0.001	−1.062	0.144
Alternative hypothesis: individual AR coefs. (between-dimension)				
		Statistic		Prob.
Group rho-statistic		−0.576		0.282
Group PP-statistic		−2.460***		0.007
Group ADF-statistic		−1.411*		0.079
Null hypothesis: no cointegration				
Trend assumption: deterministic intercept and Trend				
Automatic lag length selection based on SIC with a max lag of 5				
Newey-West automatic bandwidth selection and Bartlett kernel				

*means 10% significance level
**means 5% significance level
***means 1% significance level

Table 19.7 Short-run Granger causality results for VECM with IFDI

From	To				
	ΔCO_2	ΔEN	ΔIFDI	ΔY	ΔY^2
ΔCO_2		9.02***	3.90	0.87	1.13
		(0.01)	(0.15)	(0.65)	(0.56)
ΔEN	3.69		0.03	0.74	0.60
	(0.15)		(0.98)	(0.69)	(0.74)
ΔIFDI	1.38	0.91		7.02***	6.18**
	(0.50)	(0.63)		(0.03)	(0.05)
ΔY	5.72*	0.05	1.22		0.96
	(0.06)	(0.97)	(0.54)		(0.61)
ΔY^2	(6.27)**	0.12	1.37	1.18	
	(0.04)	(0.94)	(0.50)	(0.55)	
ε_{1t-1}	13.4***	1.18	0.05	4.32**	1.73
	(< 0.01)	(0.27)	(0.82)	(0.04)	(0.19)

The table reports chi-sq statistics with p values in parenthesis
The chi-sq tests for short-run Granger causality have 2 degrees of freedom.
The system of equation is estimated using OLS with SUR technique.
*means 10% significance level,
**means 5% significance level,
***means 1% significance level.

Table 19.8 Short-run ranger causality results for VECM with OFDI

From	To				
	ΔCO_2	ΔEN	ΔOFDI	ΔY	ΔY^2
ΔCO_2		9.89***	3.59	1.19	2.18
		(< 0.01)	(0.16)	(0.38)	(0.33)
ΔEN	3.68		0.27	0.67	0.53
	(0.16)		(0.87)	(0.71)	(0.76)
ΔIFDI	0.79	0.83		1.49	1.69
	(0.67)	(0.65)		(0.47)	(0.43)
ΔY	4.18	0.003	6.49***		0.39
	(0.12)	(0.99)	(0.04)		(0.82)
ΔY^2	4.67*	0.03	6.19	0.49	
	(0.09)	(0.98)	(0.04)	(0.78)	
ε_{1t-1}	10.70 ***	1.30	5.95***	1.56	4.00 ***
	(< 0.01)	(0.25)	(0.01)	(0.21)	(0.05)

The table reports chi-sq statistics with p values in parenthesis.
The chi-sq tests for short-run Granger causality have 2 degrees of freedom.
The system of equation is estimated using OLS with SUR technique.
*means 10% significance level, **means 5% significance level, *** means 1% significance level.

causality from CO_2 emissions to energy consumption at the 1% significance level, but no short-run direct causality from IFDI to CO_2 emissions. However, we can conclude that there is short-run causality indirectly from IFDI to CO_2 emissions (IFDI cause GDP, and GDP causes CO_2 emissions). The error correction term is significant at the 1% level in the first equation in model (2). Therefore, there is long-run causality from energy consumption, GDP, IFDI to CO_2 emissions.

Table 19.8 shows short-run Granger causality results for VECM with OFDI. According to Table 19.5, there is some evidence of short-run causality from GDP to CO_2 emissions at the 10% significance level, and from OFDI to GDP at the 1% significance level. Additionally, there is short-run causality from CO_2 emissions to energy consumption at the 1% significance level, but no short-run direct causality from OFDI to CO_2 emissions. However, we can conclude that there is short run causality indirectly from OFDI to CO_2 emissions (OFDI causes GDP, and GDP causes CO_2 emissions). The error correction term is significant at the 1% level in the first equation of model (2). Therefore, there is long-run causality from energy consumption, GDP, OFDI to CO_2 emissions.

19.4.3.2 Long-Run Equilibrium

Long-run output elasticities are estimated using ordinary least squares (OLS), or fully modified OLS (FMOLS) (Pedroni 2001). The estimated coefficients are elasticities because the variables are measured in natural logarithms. The equations with IFDI

Table 19.9 Longrun equilibrium for equations

	Equations with inward FDI		Equations with outward FDI	
	OLS	FMOLS	OLS	FMOLS
EN	0.791***(0.179)	0.846***(0.325)	0.967***(0.173)	1.145***(0.308)
IFDI	0.014(0.027)	0.029(0.049)		
OFDI			0.065***(0.014)	0.079***(0.026)
Y	5.830***(0.846)	6.113***(1.604)	5.698***(0.786)	5.688***(1.480)
Y^2	−0.329***(0.042)	−0.348***(0.080)	−0.340***(0.039)	−0.351***(0.074)
R^2	0.808	0.799	0.831	0.822
Fixed effects F value	9.481***		22.963 ***	

*means 10% significance level, **means 5% significance level, *** means 1% significance level.

shown in Table 19.9 imply the following facts. According to the results of FMOLS, the long-run elasticity of GDP to CO_2 emissions is 6.113, which means that a 1% increase in output increases CO_2 emissions by 6.1% and the long-run elasticity of energy consumption to CO_2 emissions is 0.846, which means that a 1% increase in energy consumption increases CO_2 emissions by 0.85%, respectively. However, a 1% increase in IFDI increases CO_2 emissions by 0.029% which is not statistically significant. Furthermore, we can find EKC hypothesis respectively. In the long run, economic growth and energy use hold a major role in the increase of CO_2 emissions.

The equations with OFDI shown in Table 19.9 imply the following. According to the results of FMOLS, the long-run elasticity of GDP to CO_2 emissions is 5.688, which means that a 1% increase in output increases CO_2 emissions by 5.7% and the long-run elasticity of energy consumption to CO_2 emissions is 1.145, which means that a 1% increase in energy consumption increases CO_2 emissions by 1.15%, respectively. Moreover, a 1% increase in OFDI increases CO_2 emissions by 0.079% which is significant at the 1% level. Furthermore, we can also find the EKC hypothesis in this equation. In the long run, economic growth and energy use have a major role in the increase in CO_2 emissions and OFDI also contribute to the increase of CO_2 emissions.

19.5 Conclusions

The previous section analysed the short-run causality and long-run equilibrium for ANICs. First, there is some evidence of direct short-run causality from GDP to CO_2 emissions and indirect evidence from IFDI (or OFDI) to CO_2 emissions. However, in the long-run equilibrium, the effect of IFDI on CO_2 emissions was not statistically significant. This means that IFDI did not contribute to the increase of CO_2 emissions in ANICs in the long term. On the other hand, the effect of OFDI on CO_2 emissions was statistically significant with positive coefficients in the long term. Therefore, OFDI in the region did not contribute to the mitigation of greenhouse gas emissions, and these countries have comparative advantage in energy intensive industries over time. This supports the claim that the purpose of OFDI in the region was not for pollution haven or for securing competitiveness in all aspects of CO_2 emissions. Second, although CO_2 emissions have increased with the economic growth in the long run, the long-run results supported the EKC hypothesis. Therefore, there is the possibility of green growth, which brings GHG mitigation and economic growth at the same time in the region.

References

Ang, J.B. 2007. CO_2 emissions, energy consumption, and output in France. *Energy Policy* 35: 4772–4778.

Apergis, N., and J.E. Payne. 2009a. Energy consumption and economic growth in Central America: evidence from a panel cointegration and error correction model. *Energy Economics* 31: 211–216.

Apergis, N., and J.E. Payne. 2009b. CO_2 emissions, energy usage, and output in 5. *Central America. Energy Policy* 37: 3282–3286.

Apergis, N., and J.E. Payne. 2010. Energy consumption and growth in South America: Evidence from a panel error correction model. *Energy Economics* 32: 1421–1426.

Arouri, M. E. H., A. B. Youssef, H. M'henni, and C. Rault. 2012. Energy consumption, economic growth and CO_2 emissions in Middle East and North African Countries. Institute for the Study of labor Discussion Paper Series IZA DP No. 6412.

Chen, S.-T., H.-I. Kou, and C.-C. Chen. 2007. The relationship between GDP and electricity consumption in 10 Asian Countries. *Energy Policy* 35: 2611–2621.

Cole, M.A., R.J.R. Elliott, and P.G. Fredriksson. 2006. Endogenous pollution havens: Does FDI influence environmental regulations? *Scandinavian Journal of Economics* 108: 157–178.

Cropper, M., and C. Griffths. 1994. The interaction of population growth and environmental quality. *American Economic Review* 84: 250–254.

Dickey, D.A., and W.A. Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74: 427–431.

Engle, R.F., and C.W.J. Granger. 1987. Cointegration and error correction: Representation, estimation, and testing. *Econometrica* 55: 251–276.

Grossman, G.M., and A.B. Krueger. 1995. Economic growth and the environment. *Quarterly Journal of Economics* 110: 353–377.

Halicioglu, F. 2008. An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey. Munich Personal RePEc Archive Paper No. 11457.

Hettige, H., R.E.B. Lucas, and D. Wheeler. 1992. The toxic intensity of industrial production: Global patterns, trends, and trade policy. *American Economic Review* 82: 478–481.

Holtz-Eakin, D., and T.M. Selden. 1995. Stoking the fires? CO_2 emissions and economic growth. *Journal of Public Economics* 57: 85–101.

Hoffmann, R., C.G. Lee, B. Ramasamy, and M. Yeung. 2005. FDI and pollution: A Granger causality test using panel data. *Journal of International Development* 17: 311–317.

Im, K.S., M.H. Pesaran, and Y. Shin. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econmetrics* 115: 53–74.

Jumbe, C.B.L. 2004. Cointegration and causality between electricity consumption and GDP: Empirical evidence from Malawi. *Energy Economics* 26 (1): 61–68.

Kivyiro, P., and H. Arminen. 2014. Carbon dioxide emissions, energy consumption, economic growth and foreign direct investment: Causality analysis for Sub-Saharan Africa. *Energy* 74: 595–606.

Lee, C.C., and C.P. Chang. 2008. Energy consumption and economic growth in Asian economies: A more comprehensive analysis using panel data. *Resource and Energy Economics* 30: 50–65.

Lee, C.C., C.P. Chang, and P.F. Chen. 2008. Energy-income causality in OECD countries revisited: The key role of capital stock. *Energy Economics* 30: 2359–2373.

Levin, A., C.F. Lin, and C. Chu. 2002. Unit root tests in panel data: Asymptotic and Finite sample properties. *Journal of Econometrics* 108: 1–24.

Mahadevan, R., and J. Asafu-Adjaye. 2007. Energy consumption, economic growth and prices: A reassessment using panel VECM for developed and developing countries. *Energy Policy* 35: 2481–2490.

Martinez-Zarzoso, I., and A. Bengochea-Morancho. 2004. Pooled mean group estimation of an environmental Kuznets curve for CO_2. *Economics Letters* 82: 121–126.

Mehrara, M. 2007. Energy consumption and economic growth: The case of oil exporting countries. *Energy Policy* 35: 2939–2945.

Merican, Y., Z. Yusop, Z. MohdNoor, and S.H. Law. 2007. Foreign direct investment and the pollution in Five ASEAN Nations. *International Journal of Economics and Management* 1: 245–261.

Mozumder, P., and A. Marathe. 2007. Causality relationship between electricity consumption and GDP in Bangladesh. *Energy Policy* 35: 395–402.

Narayan, P. 2005. The saving and investment nexus for China: Evidence from cointegration tests. *Applied Economics* 37: 1979–1990.

Narayan, P.K., and R. Smyth. 2008. Energy consumption and real GDP in G7 countries: New evidence from panel cointegration with structural breaks. *Energy Economics* 30: 2331–2341.

OECD (Organization for Economic Cooperation and Development). Indicators for CO_2 emissions from fuel combustion. Accessed at http://www.oecd.org/data.

Ozturk, I. 2010. A literature survey on energy-growth nexus. *Energy Policy* 38: 340–349.

Pao, H.-T., and C.-M. Tsai. 2010. CO_2 emissions, energy consumption and economic growth in BRIC countries. *Energy Policy* 38: 7850–7860.

Pao, H.T., and C.M. Tsai. 2011. Multivariate Granger causality between CO_2 emissions, energy consumption, FDI (foreign direct investment) and GDP (gross domestic product): Evidence from a panel of BRIC (Brazil, Russian Federation, India, and China) countries. *Energy* 36: 685–693.

Pedroni, P. 1999. Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics* 61: 653–670.

Pedroni, P. 2004. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20 (3): 597–625.

Pedroni, P. 2001. Full modified OLS for heterogeneous cointegrated panels. Nonstationary Panels Panel cointegration and Dynamic Panels. *Advances in Econometrics* 15: 93–130. Emererals group Publishing Limited.

Phillips, P.C.B., and P. Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75: 335–346.

Sadorsky, P. 2009a. Renewable energy consumption, CO_2 emissions and oil prices in the G7 countries. *Energy Economics* 31: 456–462.

Sadorsky, P. 2009b. Renewable energy consumption and income in emerging economies. *Energy Policy* 37: 4021–4028.

Sadorsky, P. 2011. Trade and energy consumption in the Middle East. *Energy Economics* 33: 739–749.

Selden, T.M., and D. Song. 1994. Environmental quality and development: Is there a Kuznets curve for air pollution emissions? *Journal of Environmental Economics and Management* 27: 147–162.

Shafk, N. 1994. Economic development and environmental quality: An econometric analysis. *Oxford Economic Papers* 46: 757–773.

Shafk, N., and S. Bandyopadhyay. 1992. Economic growth and environmental quality: Time series and cross country evidence. World Bank Policy Research Working Paper No. 904.

Shahbaz, M., R. Sbia, and I. Ozturk. 2014. Economic growth, electricity consumption, urbanization and environmental degradation relationship in United Arab Emirates. *Ecological Indicators* 45: 622–631.

Shiu, A., and P. Lam. 2004. Electricity consumption and economic growth in China. *Energy Policy* 32: 47–54.

Soytas, U., and R. Sari. 2009. Energy consumption, economic growth, and carbon emissions; Challenges faced by an EU candidate member. *Ecological Economics* 68: 2706–2712.

Squalli, J. 2007. Electricity consumption and economic growth: Bounds and causality analyses of OPEC members. *Energy Economics* 29 (6): 1192–1205.

UNCTAD (United Nations Conference on Trade and Development). UNCTADSTAT. Accessed at http://unctadstat.unctad.org/EN/.

World Bank. 2019. World development indicators. Accessed at http://www.worldbank.org/data.

Yoo, S. 2005. Electricity consumption and economic growth: Evidence from Korea. *Energy Policy* 33: 1627–1632.

Zhang, X.-P., and X.-M. Cheng. 2009. Energy consumption, carbon emissions, and economic growth in China. *Ecological Economics* 68: 2706–2712.

Chapter 20
Consumer Heterogeneity and Trade in Shared Renewable Resource Trading

Takeshi Ogawa

20.1 Introduction

Shared or common renewable resources are becoming to be more important in recent days. On the one hand, Pacific bluefin tuna tends to be disappearing, so it is included in Red List. On the other, Atlantic bluefin tuna is recovering after critical endangered situation because of ICCAT's agreements. Especially in Japan, transboundary renewable resources (like crabs between Japan and Russia and so on) have also focused on because of some border decision problems.

Also, these shared, common, or transboundary resources tends to be difficult to overcome the difference of product areas even if the resource is essentially common. The most important example is eel. In Japan, both eels fished/made in Japan and eels fished/made in China exist in Japanese supermarkets, but these prices are obviously different, that is, eels fished/made in Japan are higher than eels fished/made in China even if both resources are essentially completely common glass eels: at least the kind is Japanese eel, and the area in ocean is the same. However, for example, the way of fishing eels, how to grow eels in incomplete aquaculture type's eel, and so on, the production area has affected in price decision.[1]

Moreover, even if not all people have considered the production area, some people consider the production area as the key factor. The first part is local production for local consumption movement. The second part is longing to imported goods. Additional factors exist because of region-specific circumstances. For example, some (not all) Japanese people hate foods made in China because some Japanese mass-

[1]The tendency to emphasize differences in production locale was also revealed by MAFF (2005) in its Food Consumption Monitor Survey of 2004. For other cases, see for example, Wessels (2002), and Oishi et al. (2010) and so on.

T. Ogawa (✉)
Department of International Economics, School of Economics, Senshu University,
2-1-1-9710 Higashi-Mita, Tama-Ward, Kawasaki 214-8580, Japan

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_20

medias stressed the accident food is made in China when some food accident made in China happened. For example, Japanese seafoods are worried for some (not all) non-Japanese (included Chinese) people after Fukushima's nuclear accidents.

Seeing past researches, the renewable resources and trade, general equilibrium model combined between traditional Clark model in fishery economics[2] and traditional Ricardian model in international trade are started with Brander and Taylor (1997a, b, 1998). The making transboundary types of Brander=Taylor model is Takarada et al. (2013) (completely common resource) and Rus (2012) (incompletely common resource). Takarada (2010) and Takarada et al. (2012) introduced resource management into Takarada et al. (2013) type's completely common resource model. Especially, Takarada (2010) introduced technical regulations (like way of catch) and so on.[3] However, these researches do not include consumer heterogeneity.

Quaas and Requate (2013), Quaas and Stöven (2014) consider love of variety, and variety of goods with continuous goods model, but these papers does not consider consumer heterogeneity within country. Ogawa (2014, 2017) are the most similar in this article, which is considered consumer heterogeneity with separated local resources, that is, consumer heterogeneity with Brander and Taylor (1998), but they do not treat shared, common, or transboundary resources. This article considers two-country, two-good, general equilibrium consumer heterogeneity with shared, common, or transboundary renewable resources.

This article considers conditions of trade and production pattern in the case of both countries' incomplete specialization. Introducing consumer heterogeneity, the technology conditions moves the threshold of strength of preference of production location. The key changing point determines only price condition, but with zero-profit condition of both industries, the condition moves technology parameters and resource stock, but the resource stock is common, the resource stock disappears.

The next section considers the model and autarky. The third section treats trade, and the final section is conclusion.

20.2 Model and Autarky

Based on Takarada et al. (2013), we consider two-country (home and foreign*), two-good (resource good H and non-resource manufacture M), 1-factor (labor L) model with common renewable resource S, general equilibrium trading model. A function $G(\bullet)$ is a gain function which shows the recovery rate of resource per time when S is given, assumed quasi-concave, twice-continuable, and single-peaked on $(0,S)$, continued on $[0,S]$, $G(0) = G(S) = 0$, and $\frac{d}{dS}\frac{G(S)}{S} < 0$ following to Ogawa et al. (2012). The last inequality is for uniqueness.[4] We suppose that non-resource

[2]See for example, Clark (2010).

[3]For the basic comparing research between input control, output control, and technical regulation introducing Brander–Taylor types small open economy model, see Ogawa et al. (2012).

[4]See for example, Clark (2010).

manufactures are numeraire, constant return to scale, and normalized to labor, that is, $M_P = L_M$, when M_P is a production amount, and L_M is a labor input for non-resource manufactures. L is a labor endowment, so home country's labor endowment constraint is $L_H + L_M = L$, where L_H is labor input of resource good. Assume that each person's initial endowment is normalized as 1. Foreign country is similar (denoted asterisk *). Resource good's harvesting function is Schaefer types, that is, $H_P = qSL_H$, where H_P is resource good's harvest, and q is a parameter of technical level. The relative price of resource good is p, Common resource's dynamic equilibrium equation is $\dot{S} = G(S) - H_P - H_P^*$, where S_0 is an initial positive common resource stock. Following to Ogawa (2014, 2017), once trade has started, home country is resource-good net-imported country (that is, non-resource manufacture exporting country), and foreign country is resource-good net-exporting country (that is, non-resource manufacture importing country).

Consumer's utility u is considered as log-linear to sub-utility of resource v and non-resource manufactures, that is, $u = \beta \ln v + (1-\beta) \ln m$, where $0 < \beta < 1$ is the percentage of expenditure for resource good (which is the same for both countries), and m is the consumption of non-resource manufactures. Following to Ogawa (2014, 2017), we consider sub-utility v as $v = bH + (1-b)\,h$ for simplicity, where H is the consumption of resource good made in home country, h is the one made in foreign country *, $0 < b < 1$ is a strength of preference for resource goods made in home country. Thus, (1) when $b \fallingdotseq 0$, the person prefers resource good made in foreign country extremely, (2) when $b = 0.5$, the person has no difference between resource good made in home country and that made in foreign country (choose only price is low or not), and (3) when $b \fallingdotseq 1$, the person prefers resource good made in foreign country extremely, for example. For distribution, $f(\bullet)$ is a probalility density function of a strength of preference b in (0,1), and is assumed as follows for local production for local consumption movement, and no difference of production location.

Assumption: For any small $\varepsilon >$,

$$\int_0^{\varepsilon} f(b)db > 0, \quad \int_{0.5-\varepsilon}^{0.5+\varepsilon} f^{(*)}(b^{(*)})db^{(*)} > 0,$$

$$\int_{1-\varepsilon}^{1} f^*(b^*)db^* > 0, \quad \int_0^1 f^{(*)}(b^{(*)})db^{(*)} = L^{(*)}$$

The last equation is for labor constraint. The first and the third inequality of Assumption is home-bias, that is, local production for local consumption extremely. In the most important case of this model, that is, eels fished/made in Japan and eels fished/made in China, this situation holds. Consider that home country is Japan, and foreign country is China. For example in Japan, in very sad case, some (not all) person takes the trash out when s/he noticed that the food is made in China. Also for example in China, there exists a term when Japanese food's import is banned, especially made in areas near Fukushima after Fukushima's nuclear accident. The second assumption is that there exist people felt no difference of production location.

The budget constraint for home country is $pH + p^*h + m = I$, where I is income, and the one for foreign country is $pH^* + p^*h^* + m^* = I^*$.

Firstly, in the autarky case is considered. Home country's utility maximization is as follows:

$$\max_{H>0,m>0}(u =)\beta \ln v + (1-\beta)\ln m$$
$$s.t. \qquad v = bH + (1-b)\,h, \quad h = 0, \quad pH + p^*h + m = I \tag{20.1}$$

The results are $H = \frac{\beta I}{p}$ $m = (1-\beta)\,I$ The results of foreign country are $H^* = \frac{\beta I^*}{p^*}$, $m^* = (1-\beta)\,I^*$, similarly. In the autarky case, the imported resource good does not exist.

$$H = \frac{\beta I}{p},\ m = (1-\beta)I$$

With $M_P = L_M$, non-resource manufactures' profit maximization problem and 0-profit condition are as follows:

$$\max_{0\le L_H \le L} M_p - wL_M \quad \therefore \quad w \ge 1; \quad w = 1 \quad where\ L_M > 0.$$

where w is the wage. With $H_P = qSL_H$, resource good's profit maximization problem and 0-profit condition are as follows:

$$\max_{0\le L_H \le L} pH_P - wL_H, \quad \therefore pqS \leqslant w; \quad pqS = 1\ \ where\ \ L_H > 0.$$

In the autarkic condition, $w = 1$ holds because both goods are needed for both countries' consumer. Thus, pqS=1 holds. Moreover, because each one's labor endowment as 1, I =1, m = 1$-\beta$, and $H = \frac{\beta}{p}$. Therefore, $L_M = (1-\beta)L$, H_P=$\frac{\beta L}{p}$ $= qSL_H$, $L_H = \beta L$ hold. With similar calculation, foreign country's autarky can be derived. In this case, the dynamic equation

$$\frac{\dot{S}}{S} = \frac{G(S)}{S} - q\beta L - q^*\beta L^* \lesseqgtr 0 \Longleftrightarrow \frac{G(S)}{S} \lesseqgtr q\beta L + q^*\beta L^*$$

determines common resource stock S, so other variables can be determined.

In the next section, A is used as autarky. S_A satisfies steady state, that is, $\frac{G(S_A)}{S_A} = q\beta L + q^*\beta L^*$ Moreover, $p_A^{(*)} = \frac{1}{q^{(*)}S_A}$ and $u_A = \beta \ln b + \beta \ln \beta + \beta \ln q + \beta \ln S_A + (1-\beta)\ln(1-\beta)$ Similarly, $u_A^* = \beta \ln(1-b) + \beta \ln \beta + \beta \ln q^* + \beta \ln S_A + (1-\beta)\ln(1-\beta)$.

20.3 Trade Equilibrium for Trade Pattern

In this section, trading equilibrium is considered. Each consumer's utility maximization can be treated as 2-step types. Firstly, each consumer of home country maximizes sub-utility v as positive expenditure e for resource good, that is,

$$\max_{H\geq 0, h\geq 0} (v =) \, bH + (1-b)\, h \quad s.t. \quad pH + p^* h \leqslant e$$

Setting λ as multiplier, this linear types' maximization can be calculated as corner solution as follows:

$$\Phi := bH + (1-b)\, h + \lambda \left(e - pH - p^* H\right)$$

$$b - \lambda p \leqslant 0, \; H \geqslant 0, \quad (b - \lambda p)\, H = 0.$$

$$(1-b) - \lambda p^* \leqslant 0, \quad h \geqslant 0, \quad \left(1 - b - \lambda p^*\right) h = 0.$$

$$\lambda > 0, \; pH + p^* h = e$$

Thus, in the home country,

(i) $\frac{b}{p} > \frac{1-b}{p^*} \Leftrightarrow b > \frac{p}{p+p^*}$ means $h = 0, \;\; H = \frac{e}{p}, \;\; \lambda = \frac{b}{p}, \;\; v = \frac{be}{p}$.
(ii) $\frac{b}{p} < \frac{1-b}{p^*} \Leftrightarrow b < \frac{p}{p+p^*}$ means $H = 0, \;\; h = \frac{e}{p^*}, \;\; \lambda = \frac{1-b}{p^*}, \;\; v = \frac{(1-b)e}{p^*}$.

In these situations, $\ln\left(\frac{be}{p}\right) = \ln b - \ln p + \ln e$, and $\ln\left[\frac{(1-b)e}{p^*}\right] = \ln(1-b) - \ln p^* + \ln e$, hold, where the variables (not parameters) are only the part of $\ln e$ for consumers Secondly, each consumer's utility maximization can be simplified as follows:

$$\beta \ln e + (1-\beta) \ln m \;\; \text{s.t.} \;\; e + m = I. \qquad \therefore e = \beta I, \; m = (1-\beta)\, I$$

Thus,

(i) $b > \frac{p}{p+p^*}$ means $h = 0, \; H = \frac{\beta I}{p}, \;\; v = \frac{b\beta I}{p}$

$$u = \beta \ln \beta + (1-\beta) \ln (1-\beta) + \beta \ln b - \beta \ln p + \ln I,$$

(ii) $b < \frac{p}{p+p^*}$ means $H = 0, \; h = \frac{\beta I}{p^*}, \;\; v = \frac{(1-b)\beta I}{p^*}$,

$$u = \beta \ln \beta + (1-\beta) \ln (1-\beta) + \beta \ln (1-b) - \beta \ln p^* + \ln I.$$

However, if $L_M > 0$, then $I = 1$, otherwise $(L_M = 0, \; L_H = L)$ $I = w = pqS$. At least home country is non-resource manufacture exporting country, $L_M > 0$, thus, $I = 1$, Therefore,

(i) $b > \frac{p}{p+p^*}$ means $h = 0,\ H = \frac{\beta}{p},\ \ v = \frac{b\beta}{p}$,

$$u = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln b - \beta \ln p,$$

(ii) $b < \frac{p}{p+p^*}$ means $H = 0,\ h = \frac{\beta}{p^*},\ v = \frac{(1-b)\beta}{p^*}$,

$$u = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln(1-b) - \beta \ln p^*.$$

Similarly, foreign country's consumer chooses as follows:

$$\max_{H\geq 0, h\geq 0}(v^* =)b^*H^* + \left(1-b^*\right)h^*$$
$$s.t.\ \ pH^* + p^*h^* \leqslant e^*$$

(iii) $b^* > \frac{p}{p+p^*}$ means $h^* = 0,\ H^* = \frac{\beta I^*}{p},\ \ v^* = \frac{b^*\beta I^*}{p}$,

$$u^* = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln b^* - \beta \ln p + \ln I^*,$$

(iv) $b^* < \frac{p}{p+p^*}$ means $H = 0,\ h = \frac{\beta I^*}{p^*},\ v = \frac{(1-b^*)\beta I^*}{p^*}$,

$$u^* = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln(1-b^*) - \beta \ln p^* + \ln I^*.$$

However, if $L_M^* > 0$, then $I^* = 1$, otherwise ($L_M^* = 0,\ L_H^* = L^*$) $I^* = w^* = p^*q^*S$. So, there are two types' case. In this paper, the case that foreign country incompletely specializes are focused on: In this case, $I^* = 1$, and

(iii) $b^* > \frac{p}{p+p^*}$ means $h^* = 0,\ H^* = \frac{\beta}{p},\ \ v^* = \frac{b^*\beta}{p}$,

$$u^* = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln b^* - \beta \ln p,$$

(iv) $b^* < \frac{p}{p+p^*}$ means $H = 0,\ h = \frac{\beta}{p^*},\ v = \frac{(1-b^*)\beta}{p^*}$,

$$u^* = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln(1-b^*) - \beta \ln p^*.$$

However, pqS=1 and $p^*q^*S = 1$ means $\frac{p}{p*p^*} = \frac{q^*S}{q^*S+qS} = \frac{q^*}{q+q^*}$. The following proposition holds:

Proposition 20.1 (Threshold of Strength of Preference) *Consider both countries have incomplete specialization. The threshold of strength of preference which determines of buying whose country resource good is made, is only determined by only technical parameters.*

Proposition 20.1 has the following explanation. In the local resource, the stocks of resources are different between countries, so some parameters of resource are included in conditions like Brander and Taylor (1998). Like Takarada et al. (2013), if consumers feel indifferent between resource good made in home country and that made in foreign country, the only technical condition determines trade pattern in common resource. However, if consumer heterogeneity is included in the model, the choice of resource good for consumer may be different, but the resource stock is common. Thus, the technological conditions keep the threshold of strength of preference.

Therefore, the market clearing conditions are

$$\frac{\beta}{p}\left[\int_{\frac{q^*}{q+q^*}}^{1} f(b)db + \int_{\frac{q^*}{q+q^*}}^{1} f^*(b^*)db^*\right] = qSL_H \Leftrightarrow L_H = \beta\int_{\frac{q^*}{q+q^*}}^{1}\left[f(b) + f^*(b)\right]db$$

$$\frac{\beta}{p^*}\left[\int_{0}^{\frac{q^*}{q+q^*}} f(b)db + \int_{0}^{\frac{q^*}{q+q^*}} f^*(b^*)db^*\right] = q^*SL_H^* \Leftrightarrow L_H^* = \beta\int_{0}^{\frac{q^*}{q+q^*}}\left[f(b) + f^*(b)\right]db$$

Therefore,

$$\begin{aligned} L_M &= L - \beta\int_{\frac{q^*}{q+q^*}}^{1}\left[f(b) + f^*(b)\right]db > 0 \\ L_M^* &= L^* - \beta\int_{0}^{\frac{q^*}{q+q^*}}\left[f(b) + f^*(b)\right]db > 0 \end{aligned} \tag{20.2}$$

should be satisfied. The dynamic equation becomes as follows:

$$\frac{\dot{S}}{S} = \frac{G(S)}{S} - q\beta\int_{\frac{q^*}{q+q^*}}^{1}\left[f(b) + f^*(b)\right]db - q^*\beta\int_{0}^{\frac{q^*}{q+q^*}}\left[f(b) + f^*(b)\right]db$$

$$\therefore \frac{\dot{S}}{S} \lesseqgtr \Longleftrightarrow \frac{G(S)}{S} \lesseqgtr q\beta\int_{\frac{q^*}{q+q^*}}^{1}\left[f(b) + f^*(b)\right]db + q^*\beta\int_{0}^{\frac{q^*}{q+q^*}}\left[f(b) + f^*(b)\right]db$$

determines common resource stock S. Moreover, before and after trade, wages and incomes do not change, which means that demands of non-resource manufacture are the same in each country. Moreover, $L_{MA} + L_{MA}^* = M_{PA} + M_{PA}^* = M_P + M_P^* = L_M + L_M^*$, that is, total world outputs of non-resource manufactures are the same. Therefore, to determine the trade pattern as foreign country is non-resource manufacture importing country, the product of non-resource manufactures made in foreign country, and the labor input of non-resource manufactures of foreign country should also be smaller. Thus, the following condition should be satisfied:

$$L_M^* = L^* - \beta\int_{0}^{\frac{q^*}{q+q^*}}\left[f(b) + f^*(b)\right]db < (1-\beta)L^* \Longleftrightarrow L^* < \int_{0}^{\frac{q^*}{q+q^*}}\left[f(b) + f^*(b)\right]db \tag{20.3}$$

Additionally, the labor inputs of non-resource manufactures for home country should be larger with the same reason. Thus, the following condition should be satisfied:

$$L_M = L - \beta \int_{\frac{q^*}{q+q^*}}^{1} \left[f(b) + f^*(b)\right] db > (1-\beta) L \Longleftrightarrow L > \int_{\frac{q^*}{q+q^*}}^{1} \left[f(b) + f^*(b)\right] db. \tag{20.4}$$

In the paper above, only necessary conditions are argued, but this is the case classification, so the sufficiency is also satisfied. Trade pattern condition can be summarized as the following corollary:

Corollary 20.1 (Trade and Production Pattern for Incomplete Specialization) *The necessary and sufficient conditions that both countries have incomplete specialization and home country exports non-resource manufactures are as follows:*

$$L^* < \int_0^{\frac{q^*}{q+q^*}} \left[f(b) + f^*(b)\right] db < \frac{L^*}{\beta}, \; L > \int_{\frac{q^*}{q+q^*}}^{1} \left[f(b) + f^*(b)\right] db$$

Especially in the case that $f(b) = L$, $f^*(b^*) = L^*$, *the following conditions are calculated as follows:*

$$L^* < \frac{q^*}{q+q^*} \bullet \left(L + L^*\right) < \frac{L^*}{\beta}, \; L > \frac{q}{q+q^*} \bullet \left(L + L^*\right)$$

which can be rewritten as

$$\frac{L}{q} > \frac{L^*}{q^*}, \quad \frac{\beta q^*}{q+q^*} < \frac{L^*}{L+L^*} \tag{20.5}$$

In this situation, the following corollary holds.

Corollary 20.2 (Uniform Distribution) *In the uniform distribution case, the trade and production pattern conditions become as (20.5).*

When each consumer does not consider the difference between production location like Takarada et al. (2013), the trade pattern determines only technology parameter. However, in the uniform distribution case of consumer heterogeneity within each country, the demand of each resource good is affected as population, that is, labor endowment.

Using $p^{(*)} q^{(*)} S = 1$, the utility and so on are as follows:

(i) $b > \frac{q^*}{q+q^*}$ means $h = 0$, $H = \beta q S$, $v = b \beta q S$,

$$u = \beta \ln \beta + (1-\beta) \ln (1-\beta) + \beta \ln b + \beta \ln q + \beta \ln S,$$

(ii) $b < \frac{q^*}{q+q^*}$ means $H = 0$, $h = \beta q^* S$, $v = (1-b) \beta q^* S$,

$$u = \beta \ln \beta + (1-\beta) \ln (1-\beta) + \beta \ln (1-b) + \beta \ln q^* + \beta \ln S.$$

(iii) $b^* > \frac{q^*}{q+q^*}$ means $h^* = 0,\ H^* = \beta q S,\ \ v^* = b^* \beta q S,$

$$u^* = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln b^* + \beta \ln q + \beta \ln S,$$

(iv) $b^* < \frac{q^*}{q+q^*}$ means $H = 0,\ h = \beta q^* S,\ v = \frac{(1-b^*)\beta}{p^*},$

$$u^* = \beta \ln \beta + (1-\beta)\ln(1-\beta) + \beta \ln(1-b^*) + \beta \ln q^* + \beta \ln S.$$

The most important point is as follows: when (i) and (iv), the only difference between utility before trade and that after trade is only resource stock S. Because income and non-resource manufactures are the same between before and after trade, the difference of utility between before and after trade is caused as resource good consumption between before and after trade. However, from the original dynamic equation means that total outputs of resource good are higher, the resource decreases. If the resource stock becomes higher after trade, at least one country must decrease the product of resource good. This means the importance of resource stock management.

20.4 Conclusion

This article considers consumer heterogeneity and trade in shared renewable resource trading, especially in the case that both countries incomplete specialization case. This article focused on the conditions of trade and production pattern. Introducing consumer heterogeneity, the technology conditions moves the threshold of strength of preference of production location. The key changing point determines only price condition, but with zero-profit condition of both industries, the condition moves technology parameters and resource stock, but the resource stock is common, the resource stock disappears.

References

Brander, J.A., and M.S. Taylor. 1997a. International trade and open access renewable resources: The small open economy case. *Canadian Journal of Economics* 30: 526–52.

Brander, J.A., and M.S. Taylor. 1997b. International trade between consumer and conservationist countries. *Resource and Energy Economics* 19: 267–297.

Brander, J.A., and M.S. Taylor. 1998. Open-access renewable resources: Trade and trade policy in a two-country model. *Journal of International Economics* 44: 181–209.

Clark, C.W. 2010. *Mathematical bioeconomics: The mathematics of conservation*, 3rd ed. United States of America: Wiley.

Ministry of Agriculture, Forestry and Fisheries (MAFF). 2005. The labelling of fresh food and processed food, (written with Japanese), 2004 Food consumption monitor 2nd scheduled survey results. http://www.maff.go.jp/j/heya/hmoniter/pdf/h1602.pdf. Accessed on 15 Jan 2013.

Oishi, T., J. Ominami, N. Tamura, and N. Yagi. 2010. Estimates of the potential demand of Japanese consumers for eco-labeled seafood products, (written with Japanese). *Nippon Suisan Gakkaishi* 76: 26–33.

Ogawa, T., Y. Takarada, and W. Dong. 2012. International trade and management methods for renewable resources. In *Modern economic theory and policy issues: Professor Shigeki Yabuuchi's 60th anniversary collection, (written with Japanese)*, ed. K. Kondo, M. Kunisaki, and Y. Takarada. Keiso Shobo, Japan, Chapter 11: 160–178.

Ogawa, T. 2014. Consumer heterogeneity and gains from trade in renewable resource trading. *SSRN Working Paper* 2495465, pp. 1–31.

Ogawa, T. 2017. Consumer heterogeneity and gains from trade in renewable resource trading: No management case. In *New extension of economic theory, application, and empirical analysis*, ed. A. Matsumoto. Tokyo: Chuo University Press.

Quaas, Martin F. and Max T. Requate. 2013. Sushi or fish fingers? Seafood diversity, collapsing fish stocks, and multispecies fishery management. *The Scandinavian Journal of Economics* 115 (2): 381–422.

Quaas, Martin, and Max T. Stöven. 2014. New trade in renewable resources and consumer preferences for diversity. In Inaugural Dissertation zur Erlangung des akademischen Grades eines Doktors (Ph.D. Thesis) 2014, ed. Max T. Stöven, Kiel University - Christian-Albrechts-Universität zu Kiel, 64–83. http://macau.uni-kiel.de/servlets/MCRFileNodeServlet/dissertation_derivate00005463/diss_stoeven.pdf. Accessed on 31 Mar 2017.

Rus, H.A. 2012. Transboundary marine resources and trading neighbours. *Environmental and Resource Economics* 53 (2): 159–184. https://doi.org/10.1007/s10640-012-9554-6.

Takarada, Y. 2010. Shared renewable resource and international trade: Technical measures for fisheries management. *RIETI Discussion Paper Series*, 10-E-035. http://www.rieti.go.jp/jp/publications/dp/10e035.pdf. Accessed 16 Apr 2014.

Takarada, Y., T. Ogawa, and W. Dong. 2012. International trade and management of shared renewable resource. *Society of Economics Working Paper Series*, Series No. 48, Nanzan University.

Takarada, Y., W. Dong, and T. Ogawa. 2013. Shared renewable resources: Gains from trade and trade policy. *Review of International Economics* 25 (1): 1032–1047. https://doi.org/10.1111/roie.12089.

Wessells, C.R. 2002. The economics of information: Markets for seafood attributes. *Marine Resource Economics* 17 (2): 153–162.

Chapter 21
The Relationship Between Oil Price and Korean Industrial Production and Inflation

In Huh

21.1 Introduction

In the 1970s and 1980s, the global economy experienced a stagflation characterized by a decline in output and rising inflation as a result of the first and second oil shocks. In the first oil shock, the US real GDP grew at a rate of 10.3% in the first quarter of 1973, but declined to −2.1% in the third quarter. In 1974, the economy grew at an average annual rate of −1.9%. The decline in the production of the economy lasted about a year and a half after the oil shock. Meanwhile, US consumer prices rose by 3.4% in 1972, but rose to 8.9% in 1973 and 12.1% in 1974, despite the economic downturn. The impact of the second oil shock on the economy has also had a ripple effect of rising prices and a recession. The rise in oil prices due to two oil shocks is believed to be a stumbling block to economic growth due to a supply shock caused by production contraction caused by rising costs.

Crude oil prices remained stable at about US $20/bbl (Western Texas Intermediate: WTI) after the second oil shock. Oil prices rose after the second oil shock remained largely unchanged until November 1985, then plummeted to $10/bbl. It remained below $20/bbl by the end of the 1990s, except for the surge in oil prices due to the outbreak of the Gulf War in 1990. Therefore, the real oil prices are estimated to have declined for a considerable period of time. But in 2003, oil prices began to skyrocket and rose to $145/bbl in July 2010. In particular, during the first half of 2008, oil prices surpassed the forecasts of experts, and the real oil prices exceeded those of the second oil shocks era. The investment bank, Goldman Sachs even had predicted that it would rise to $200/bbl by the end of 2008 due to the shortage of oil supply. However, the problem of the US subprime mortgage crisis, which began at the end of 2007, was a burden on the US economy, and the financial crisis was evolving due to the merger of Bear Sterns, public funding for Fannie Mae and Freddie Mac. As

I. Huh (✉)
The Catholic University of Korea, Seoul, South Korea
e-mail: inhuh@catholic.ac.kr

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_21

Fig. 21.1 Prices of Western Texas intermediate

a result, the sluggish trade resulting from the prolonged global financial crisis led to sluggish production in the second half of 2008, and the price of international oil prices also fell sharply. As the global financial crisis resolved, investment sentiment recovered as well, and oil prices rebounded to US $70–80/bbl as of September 2010 (Fig. 21.1).

The rise in oil prices since 2003 did not have a large impact on the economy, unlike the oil price shocks. The US economy grew at a slower rate of 0.2% in 2001 and 1.9% in 2002 after the collapse of the IT bubble, while 3.7%, 3.4% and 3.2% in 2003, 2004, 2005 respectively. The US GDP growth were exceeding the estimated potential growth rate. Consumer price inflations in the US also remained at a low level of 3% in the early 2000s until 2005. This is a very different result from the rise in oil prices during the oil shocks. Conversely, oil prices fell from $145/barrel to $35/barrel due to the financial crisis, which stabilized prices but did not help recovery of production activity. We can conclude that the rise in oil prices in the 2000s was not accompanied by inflation, except for 2008, and falling oil prices led to price declines, but oil prices did not show a negative correlation with production activities.

Therefore, there is a lot of discussion about the fact that rising oil prices do not cause inflation and economic slowdown. According to Huh (2006) and other securities firms' reports, the reason for this is the emergence of emerging economies represented by BRICs. In particular, China's growth suggests that it supplies low-priced industrial goods to the world, easing inflationary pressures from rising oil prices. Second, the pace of oil price rises is relatively long, unlike the oil shock, so that economic agents have had time to cope with rising oil prices. Third, it is argued that the impact of rising oil prices is limited due to the lower importance of oil to the economy. As technology has been developed gradually in the direction of energy efficiency and alternative energy such as nuclear power generation has been developed, the proportion of crude oil to the economy has decreased and the impact on the economy has become smaller than before.

In addition, the oil shock and recent oil price hike have different causes. In the oil shocks, oil prices were rising due to supply limitations caused by political and economic reasons. The recent rise in oil prices is believed to be due to the large increase in demand of emerging economies. Moreover, at the time of the financial crisis, world economy shrank and crude oil demand also decreased, and oil price fell sharply. As the causes of oil price fluctuations are different, we can expect that the results of oil price fluctuations are also different.

Literatures on the crude oil market have been mainly carried out after the first and second oil shocks. The recognition of Hotelling (1931) that future generations may suffer when oil as nonrenewable resource left to the market system raises the need for a crude oil market analysis. Thus, the study of the oil market has focused primarily on how crude oil suppliers behave in view of the scarcity of finite resources (Krautkraemer 1998; Griffin 1985; Adelman 1986; Weitzman 1999; Hillman and Long 1985 and Borenstein and Shepard 2002). There are also studies about technology development in energy industry (Bruce 1980; Popp 2002; Newell et al. 1999). In this paper, we analyze the effects of oil price fluctuations by supply and demand factors which has not been done by other empirical studies about oil prices (Gisser and Goodwin 1986; Eastwood 1992; Hamilton 1983; Bomberger and Makinen 1993; Hooker 2002; Kim and Young 2002). Kilian (2009) analyzed the effects of oil prices on the macroeconomic variables of the US by classifying structural factors such as demand, supply and preliminary demand in the crude oil market. In this study, unlike Kilian 's analysis, we analyze the oil price fluctuation of supply and demand factors by using OPEC' s crude oil supply data, and analyze how oil price fluctuation affects production and inflation of Korea. Classification based on crude oil supply is simpler than classification using structural autoregression methodology, and has the advantage of distinguishing between supply and demand factors at the time of data observation.

In Sect. 21.2, we analyze the impact of oil price fluctuations on macroeconomic variables since 1990, and examine the effect of oil price fluctuations by supply and demand factors in Korean production and inflation. Section 21.3 concludes.

21.2 An Empirical Analysis of the Ripple Effect of Oil Price Rise

21.2.1 Data

For empirical analysis, price, output, and monetary variables were selected by Hamilton (1983). Other prior studies, including Hamilton (1983), use time series data related to prices and production, as the impact of oil prices on macroeconomics will generally shift production costs. Import price index and consumer price index were used as inflation variables. Industrial production index was used as output variable. In the analysis of Hamilton (1983), unemployment rate was used, but unemployment rate were excluded from the analysis because only statistics after 1999 were available due to the change of the standard of unemployment rate. As in Kilian (2009), it would be a good analysis to study countries with large economies in the world, such as the United States, and to use clear data on the causes of oil price fluctuations by demand and supply factors. However, it is difficult to expect that the results of the analysis of Kilian (2009) will be maintained in the case of a country that is not expected to have a large impact on the global oil market due to the smaller scale of the economy. Therefore, we used data from Korea to analyze the effect of oil price changes on Korean economy. In Korea as the small open economy, the economy is relatively dependent on the external sector. In particular, the external shocks that gave negative impact on Korea during the sample period were the IT bubble collapse, Asian currency crisis and Global financial crisis. Therefore, we can see that the Korean economic variables have moved along with global economic fluctuations. The currency variable was M1. M1 and import price index were obtained through the Economic Statistics System (ECOS) of the Bank of Korea, and the Industrial Production Index was obtained through the Korean Statistical Information Service (KOSIS) of the National Statistical Office. The oil price was the Dubai oil price of the Middle East which is closest to the Korean importing oil price. Oil prices were gained through statistical services Bloomberg. Each variable can be interpreted as a rate of change since it is differenced after taking the natural logarithm. The data were from February 1990 to July 2010 as monthly data. The data have a total of 246 observations (Table 21.1).

Monthly average growth rates of oil prices, consumer prices, import prices, and industrial production indices for the sample period are 0.6%, 0.3%, 04% and 0.7%, respectively. During the sample period, except for the Gulf War in 1991, there was no increase in the crude oil market due to a large supply shock. Oil prices rose 52% due to the outbreak of the Gulf War. Out of the 246 specimens, oil prices rose 143 times. During the sample period, the inflation rate was lower than the oil price increase on average, and industrial production index increased despite the rise in oil prices.

Table 21.1 Descript statistics

	Dubai (Crude oil price)	IP (Industrial production)	CPI (Consumer price index)	IPI (Import price index)	M1
Mean	0.00577	0.00668	0.00344	0.00397	0.00984
Median	0.01386	0.00106	0.00353	0.00410	0.00875
Max	0.52126	0.19582	0.02498	0.22441	0.11948
Min	−0.33545	−0.15445	−0.00606	−0.09574	−0.10193
Standard deviation	0.09170	0.06249	0.00472	0.02976	0.02455

Note All variables are the 1st differenced of natural logarithm

21.2.2 Impact of Oil Price Fluctuations

(1) Influence of General Oil Price Fluctuations
The Granger Causality test can be used to analyze the correlation between macro variables and oil price fluctuations and the direction of their effects. To see if a variable is a Granger cause

$$y_t = \beta_0 + \sum_{i=1}^{q} \beta_i y_{t-i} + \sum_{i=1}^{q} \gamma_i c_{t-i} + \epsilon_t \tag{21.1}$$

Estimate the Eq. (21.1) and test the null hypothesis, $\gamma_i = 0, \forall i$. If the null hypothesis is rejected, it is concluded that 'the variable x does not Granger-cause a variable y'. This is because we cannot be sure that the real causation so we only say 'Granger-cause'. The purpose of this paper is to analyze the production and inflation reaction due to oil prices fluctuation. Therefore, the analysis of the relationship between oil price fluctuations, production, and inflation by the Granger-causality is a valid analysis.

Table 21.2 shows the results of Granger-causality test between macro variables and oil price fluctuations. All of the variables except for M1 with q = 6 in Eq. (21.1) were statistically significant at 10% significance level. (H1, H3, H5) that production, inflation (CPI and IPI) do not Granger-cause oil prices. Data exist from February 1990, but if q = 6, it can be analyzed from July 1990 data. If q = 12, it can be analyzed from February, 1991.

However, the effect of oil price fluctuations on macroeconomic variables has a statistically significant effect on the production index only and the remainder are not statistically significant. In other words, Hypothesis (H2) that oil price does not Granger-cause production was rejected, but Hypothesis (H4, H6) that oil price does not Granger-cause inflations were not rejected. As in Kilian (2009), the impact of

Table 21.2 Granger-causality test between oil price and macro variables

Null hypothesis		1990:7 ∼ 2010:7, q = 6		1991:2 ∼ 2010:7, q = 12	
		F(6,240)	(p-value)	F(12, 234)	(p-value)
IP	H1: IP↛Dubai	2.00	(0.07)	1.74	(0.06)
	H2: Dubai↛IP	4.03	(0.00)	3.04	(0.00)
CPI	H3: CPI↛Dubai	0.87	(0.52)	2.30	(0.01)
	H4: Dubai↛CPI	0.91	(0.49)	1.05	(0.40)
IPI	H5: IPI↛Dubai	2.87	(0.01)	3.19	(0.00)
	H6: Dubai↛IPI	1.16	(0.33)	1.06	(0.40)
M1	H7: M1↛Dubai	0.24	(0.96)	0.81	(0.64)
	H8: Dubai↛M1	0.48	(0.82)	0.74	(0.71)

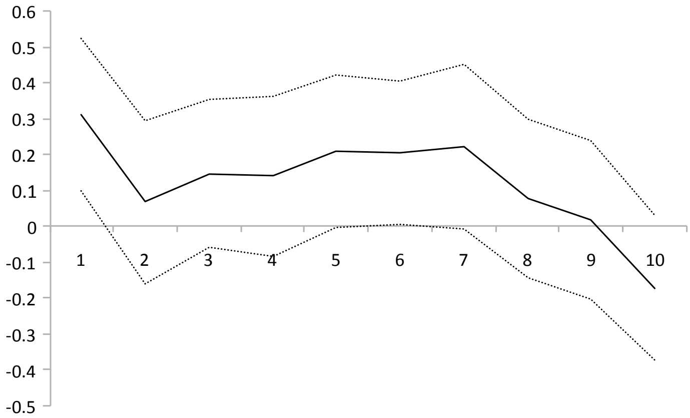


Fig. 21.2 Impulse response of oil prices to 1% increase of industrial production index

rising oil prices during this sample period on the Korean economy is different from that of the oil shock.

Among the above results, the fluctuations of the production index and the oil prices showed the significant results in both directions. However, the impulse response function is derived from vector autoregression (VAR) shows that the industrial production affects the oil prices. Figure 21.2 shows the impulse response curve of VAR analysis of the industrial production index and the oil prices when $q = 12$. As can be seen in the figure, the 1% increase in the production index had an effect of raising the oil price by about 0.3% and the effect continued for a considerable period of time. Moreover, the effect is statistically significant at 5% significance level,[1] So, we may assume that the fluctuation of oil prices during the sample period is caused by the demand of crude oil due to economic fluctuations.

As for consumer prices, there is a negative correlation between oil prices and consumer prices. In Fig. 21.3, we show the impulse response function of consumer price inflation and oil prices of VAR analysis when $q = 12$. As shown in the figure, oil prices fell after the rise in consumer price indexes except for the first month. Especially after seven months, the negative impulse of oil prices was statistically significant. Thus, the overall effect of consumer price inflation can be concluded to be related to the drop in oil prices. However, except for the effects after 7 months, these negative impulses were not statistically significant.[2] The correlation between consumer prices inflation and oil prices is small. In the case of rising oil prices, if the increase in demand for crude oil is caused by the demand shock which is increase in production, the correlation between oil price and inflation could be both positive and negative through cost side and demand side effect. Apparent negative correlation

[1]The confidence interval in Fig. 21.2 is calculated by Monte Carlo method.

[2]Although the overall effect was not statistically significant, there were statistically significant coefficients for each of the time-varying coefficients. So, the Granger Causality test yielded statistically significant result.

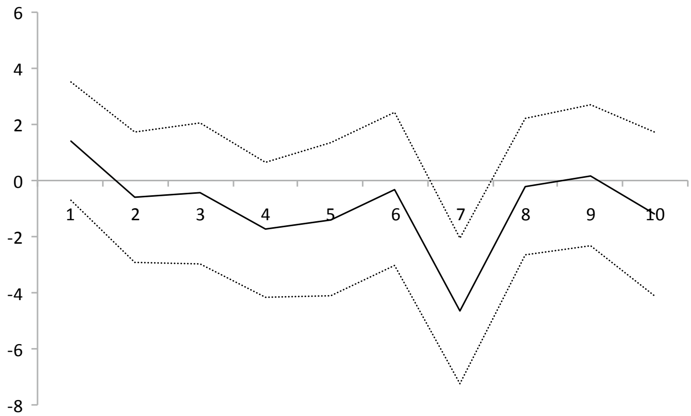


Fig. 21.3 Impulse response of oil prices to 1% increase of consumer price index

between oil price and inflation shows that it is highly likely that oil price fluctuations stemmed from demand factors.

During the sample period, the fluctuations of oil price were not largely affected by supply factors, such as the cartels of OPEC, indicating that the oil price changed mainly due to demand changes for crude oil. For example the global economic upturn continued until the financial crisis, which can be attributed to the rise in oil prices. It is hard to say that the fluctuation of consumer price inflation has provided the cause of oil price increase. However, the correlation between consumer price and oil price were low since the cost push effect from oil price increase were neutralized by the increase of production. Can be. In addition, production showed a positive correlation with oil prices, which supports the argument that oil prices was rising due to rising demand for crude oil. The empirical results up to now have argued that the fluctuation due to the crude oil demand factors led to these results. Therefore, it is necessary to analyze whether the fluctuation of oil price is made by demand factor or supply factor.

(2) Classification of Oil Price Fluctuation Factors

Whether the fluctuation of oil price is caused by fluctuation of crude oil supply or fluctuation of demand can be judged by seeing fluctuation of crude oil production. If there is a positive relationship between oil price changes and crude oil production changes, it can be judged that price fluctuations are caused by demand fluctuations rather than by changes in oil supply. Otherwise, we can judge price fluctuations by crude oil supply changes. In other words, price fluctuations due to fluctuations in demand means changes on the supply curve of crude oil (marginal cost curve when the supplier is monopoly), and price fluctuations due to supply fluctuations are the changes on the demand curve (or marginal profit curve). If demand and supply are changed simultaneously, the positive relation between oil production and price can be interpreted that demand fluctuation is greater than supply fluctuation, otherwise, supply fluctuation is larger. Therefore, it is possible to judge whether the fluctuation of oil price fluctuates according to mainly the supply factor or the demand factor by checking the oil production and price.

The data of OPEC's crude oil production has been available from January 1990 onwards in Bloomberg. According to the above-mentioned criteria, the oil price fluctuation by using OPEC's production and oil price is classified as follows: 138 fluctuations of 246 (56.6%) are from demand shocks, and 106 (43%) due to supply shocks. We could not identify the reason of fluctuation 2 times because there were no changes in the OPEC's production. Especially, the rise of oil price due to the demand of oil was the rise of 89, which is 62% of 143 oil price increases. On the other hand, only 49 (48%) out of 13 of oil prices decreases were due to demand factors. The oil price fluctuations were mainly caused by demand factors during the sample period. Demand factors were overwhelming in oil prices increases. This is in line with the opinions of the media and the experts that the surge in oil prices in the 2000s is due to demand increase of crude oil and the drop of oil price during the financial crisis is due to the sudden stop of demand for it. Therefore it supports the interpretation that oil price fluctuations as a demand factor during the sample period in previous section.

(3) Influence of Fluctuations in Oil Prices (Supply and Demand Factors)

In order to conduct the Granger-causality test after identifying the oil price fluctuations by the supply and demand factors, the estimation Eq. (21.1) must be modified to estimate the Granger-causality correctly. If you want to test whether the oil price GrangerCause macro-variables, estimate the following equation.

$$y_t = \beta_0 + \sum_{i=1}^{q} \beta_i y_{t-1} + \sum_{i=1}^{q} \gamma_i^d o_{t-1}^d + \sum_{i=1}^{q} \gamma_i^s o_{t-1}^s + \epsilon_i i \tag{21.2}$$

Where o^d is the oil price fluctuations caused by the demand changes, o^s is the oil price fluctuation caused by the supply changes.[3] In order to test if the oil price fluctuations driven by supply factor Granger-caused the variable y in the Eq. (21.2), the null hypothesis, $\gamma_i^s = 0, \ \forall i$ should be rejected. If the oil price fluctuations driven by demand factor Granger-caused the variable y, the null hypothesis $\gamma_i^d = 0, \ \forall i$ should be rejected. The null hypothesis H2, H4, H6, and H8 in Tables 21.3 and 21.4 are the results of Granger-causality test of oil price fluctuations by the supply and demand factor to macroeconomic variables.

In order to test if the macroeconomic variables Granger-caused the oil price fluctuations driven by supply and demand factors separately, we need to estimate different equations than Eqs. (21.1) and (21.2).

$$o_t = \beta_0 + \sum_{i=1}^{q} \beta_i o_{t-1} + I(o_t = o_t^d) \sum_{i=1}^{q} \gamma^d y_{t-1} + I(o_t = o_t^d) \sum_{i=1}^{q} \gamma_i^d y_{t-1} + \epsilon_t \tag{21.3}$$

[3] $o_t^{d-} I(o_t q_t - q_{t-1}) > 0) o_t$ where $o_t = p_{ot} - p_{ot-1}$ therefore, $o^d + o\, s = o_t$.

Table 21.3 Granger-casuality tests between macroeconomic variables and oil price driven by supply factor

Null hypothesis		1990:7 ∼ 2010:7, q=6		1991:2 ∼ 2010:7, q=12	
		F(6,240)	(p-value)	F(12, 234)	(p-value)
IP	H1: IP↛Dubai	0.66	(0.68)	0.75	(0.70)
	H2: Dubai↛IP	0.22	(0.97)	1.17	(0.31)
CPI	H3: CPI↛Dubai	1.56	(0.16)	2.38	(0.01)
	H4: Dubai↛CPI	1.29	(0.26)	0.83	(0.62)
IPI	H5: IPI↛Dubai	1.84	(0.09)	1.82	(0.05)
	H6: Dubai↛IPI	0.85	(0.53)	0.64	(0.81)
M1	H7: M1↛Dubai	0.50	(0.81)	1.51	(0.12)
	H8: Dubai↛M1	2.13	(0.05)	1.72	(0.06)

Table 21.4 Granger-casuality tests between macroeconomic variables and oil price driven by demand factor

Null hypothesis		1990:7 ∼ 2010:7, q=6		1991:2 ∼ 2010:7, q=12	
		F(6,240)	(p-value)	F(12, 234)	(p-value)
IP	H1: IP↛Dubai	2.76	(0.01)	2.24	(0.01)
	H2: Dubai↛IP	4.76	(0.00)	3.28	(0.00)
CPI	H3: CPI↛Dubai	1.33	(0.24)	2.33	(0.01)
	H4: Dubai↛CPI	0.79	(0.58)	0.52	(0.89)
IPI	H5: IPI↛Dubai	2.49	(0.02)	2.84	(0.00)
	H6: Dubai↛IPI	1.08	(0.38)	1.22	(0.27)
M1	H7: M1↛Dubai	0.40	(0.88)	0.53	(0.89)
	H8: Dubai↛M1	0.92	(0.48)	1.09	(0.37)

If the variable y Granger-caused the oil price fluctuation driven by supply factor, the null hypothesis $\gamma_i^s = 0,\ \forall i$ should be rejected in Eq. (21.3). If the variable y Granger-caused the oil price fluctuation driven by demand factor, the null hypothesis $\gamma_i^d = 0$ should be rejected. The estimation Eq. (21.3) is the estimation formula considering the method of Chow test because the variable y affects the oil price differently by the supply and demand factor. If we use o^s and o^d as a direct explanatory variables as in Eq. (21.2), we cannot get the intended test since there is a zero in o^s and o^d.

Therefore, it is reasonable to analyze the estimation Eq. (21.3) in order to test the Granger-causality from variable y to the oil price fluctuations by supply and demand factor. In Tables 21.3 and 21.4, the null hypotheses H1, H3, H5, and H7 were tested and the Granger causality test was performed by estimating the Eq. (21.3).

Table 21.3 shows the Granger causality test results between oil price fluctuations of supply factors and macroeconomic variables. As shown in Table 21.2, the effect of oil price fluctuations on macroeconomic variables was not found statistically significant except for the effect on the money supply. That is, H2, H4 and H6 were not rejected. The fact that there was no significant change in the supply of crude oil during the sample period could lead to this result. Even when the supply of crude oil was shrinking, it has little impact on production because it has not shrunk to such an extent as to hinder production. The effect of the industrial production index on oil price driven by supply factors was not statistically significant at the 10% significance level, but the remaining variables were significant. That is, H1 was not rejected, but H3 and H5 were rejected. In Fig. 21.3, we saw that inflation dropped oil prices. The rise in inflation can be interpreted as having the effect of lowering oil prices by expanding the supply of crude oil by causing OPEC to worry about the contraction caused by rising costs.

Oil price fluctuations due to OPEC production adjustments and industrial production are generally expected to be negatively correlated, but during the sample period the oil price changes due to the reduction or expansion of crude oil supply did not affect production. As can be seen from the positive relationship between production and oil prices in Fig. 21.2, it can be said that oil prices and production during the sample period were not caused mainly by oil price fluctuations driven by supply factors. And the Granger Causality results of oil prices also support that oil prices fluctuate due to demand factors rather than supply factors in this sample period.

Table 21.4 shows the results of the Granger causality test between oil price fluctuations driven by demand factors and macroeconomic variables. Oil price fluctuations driven by demand factors have no statistically significant effects on inflations but on industrial production. In other words, the null hypothesis (H4, H6) that the oil price fluctuation does not Granger-cause the inflations was not rejected. If oil prices fluctuate due to demand factors, the economic fluctuations, i.e., changes in production and prices, will precede fluctuations in oil prices. However, both leads and lags' relations of oil prices and productions are statistically significant.

The increase in demand for crude oil stemmed from the economic boom which is the increase in productions and price increases due to the economic boom. Therefore, industrial production and inflations has Granger-caused the fluctuations of oil prices driven by demand factor.

For the industrial production and inflations, the supply factor fluctuations of oil prices have no effect on them, but the demand factor fluctuations of oil price is statistically significantly affected by them. Oil price fluctuations driven by supply factors are expected to have an impact on macroeconomic variables, but OPEC crude oil production fluctuations did not provide enough negative impact on the industrial production during the sample period. On the other hand, industrial production fluctuations influenced and were influenced by the oil prices fluctuations driven by demand factors. The rise in oil prices due to demand factors was due to the increase in demand for crude oil which means the increase in industrial production, and this makes positive correlations between oil prices and industrial productions. On the other hand, the rise in oil prices in the supply side is resulted from the decrease in OPEC crude oil production, which may have a negative impact on industrial production due to the reduction of energy supply. However, OPEC's production reductions during the sample period were not large enough to affect the industrial productions. The result showed that the oil price fluctuations were affected by the inflations. It is reasonable to assume that OPEC countries' production decisions were reacted to the global inflations in order not to influence the demand of crude oil.

21.3 Conclusion

In this paper, to analyze the relationship between macroeconomic variables and oil prices in response to changes in supply and demand factors in the crude oil market, we conducted empirical analysis using data from January 1990 to July 2010. Although the rise in oil prices in 2000s exceeded the level of oil shocks on a nominal basis, industrial production contraction due to the oil price hike could not be found in Korea. In particular, the effect of oil prices on inflations was not statistically significant during the sample period.

Using crude oil production data, we could analyze oil price fluctuations by causes. During the sample period, either demand or supply driven oil price fluctuations did not Granger-cause inflations. In particular, oil price fluctuations driven by supply factors had no effect on industrial production, only oil price fluctuations driven by demand factors affected industrial production. Unlike the two oil shocks, oil prices mainly affected by crude oil demand so the industrial production and oil prices moved in the same direction.

During the sample period, the effect of oil price fluctuations on inflations was minimal, unlike in the oil shocks. Even the oil prices fluctuations driven by supply factor did not Granger-cause inflation. The emergence of emerging economies, and the development of energy-saving technologies have all helped offset the negative effects of rising oil prices. In this paper, we use oil production data to classify the reason of oil price fluctuations into demand factor and supply factor, and show the effect of those through the empirical analysis.

References

Adelman, M.A. 1986. Scarity and world oil prices. *Review of Economics and Statistics* 68 (3): 387–397.

Bomberger, W.A., and G.E. Makinen. 1993. Inflation and relative price variability: Park's study reexamined. *Journal of Money, Credit, and Banking* 25 (4): 854–861.

Borenstein, S., and A. Shepard. 2002. Sticky prices, inventories, and market power in wholesale gasoline markets. *RAND Journal of Economics* 33 (1): 116–136.

Bruce, A. 1980. Oil supply projection techniques: Past developments and future requirements. *Managerial and Decision Economics* 1 (2): 61–66.

Eastwood, R.K. 1992. Macroeconomic impacts of energy shocks. *Oxford Economic Papers* 44: 403–425.

Gisser, M., and T.H. Goodwin. 1986. Crude oil and macroeconomy: Tests of some popular notions. *Journal of Money, Credit, and Banking* 18 (1): 95–103.

Griffin, J.M. 1985. OPEC behavior: A test of alternative hypotheses. *American Economic Review* 75 (5): 954–963.

Hamilton, J.D. 1983. Oil and the macroeconomy since World War II. *Journal of Political Economy* 92 (5): 228–248.

Hillman, A.L., and N.V. Long. 1985. Monopolistic recycling of oil revenue and intertemporal bias in oil depletion and trade. *Quarterly Journal of Economics* 100 (3): 597–624.

Hooker, M.A. 2002. Are oil shocks inflationary? Asymmetric and nonlinear specifications versus changes in regime. *Journal of Money, Credit, and Banking* 34 (2): 540–561.

Hotelling, H. 1931. The economics of exhaustible resources. *Journal of Political Economy* 39 (2): 137–175.

Huh, In. 2006. Analysis on oil price and economic cycle. Samsung Research Institute of Finance. (In Korean).

Kilian, L. 2009. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review* 99 (3): 1053–1069.

Kim, Young-Duk. 2002. A Study on the Spreading Path of Oil Price Fluctuation Impact. Korea Energy Economic Institute (In Korean).

Krautkraemer, J.A. 1998. Nonrenewable resource scarcity. *Journal of Economic Literature* 36: 2065–2107.

Newell, R.G., A.B. Jaffe, and R.N. Stavins. 1999. The induced innovation hypothesis and energy-saving technological change. *Quarterly Journal of Economics* 941–975.

Popp, D. 2002. Induced innovation and energy prices. *American Economic Review* 92 (1): 160–180.

Weitzman, M.L. 1999. Pricing the limits to growth from minerals depletion. *Quarterly Journal of Economics* 114 (2): 691–706.

Correction to: Contemporary Issues in Applied Economics

Moriki Hosoe, Biung-Ghi Ju, Akira Yakita and Kiseok Hong

Correction to:
M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*, https://doi.org/10.1007/978-981-13-7036-6

The original version of the book was inadvertently published with incorrect information in Chapters 7 and 11. The following corrections have been made:

In Chapter 7, the third author's name was changed from "Akira Kawasaki" to "Akio Kawasaki".

In Chapter 11, the acknowledgement text that was missing has now been included.

The correction book has been updated with the changes.

The updated version of these chapters can be found at
https://doi.org/10.1007/978-981-13-7036-6_7
https://doi.org/10.1007/978-981-13-7036-6_11
https://doi.org/10.1007/978-981-13-7036-6

M. Hosoe et al. (eds.), *Contemporary Issues in Applied Economics*,
https://doi.org/10.1007/978-981-13-7036-6_22