Implications of population aging for education, technological progress and economic growth

Xiao Chen

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Economics

Australian School of Business

The University of New South Wales

June 2014

Abstract

This thesis is primarily concerned with the potential roles of individual aging (changes in people's productivity and innovation capacity from young to old) and population aging (increased proportion of old to young people in the population) in determining rates of innovation and technological progress. Since economic growth depends largely on innovation, this thesis will investigate whether and how economic growth is affected by individual aging or population aging.

To study individual aging, this thesis constructs a model of technology innovation and adoption and economic growth. Individuals' innovative and adoptive abilities change with age. This model shows that individual aging slows down technological progress of countries far from the world frontier more than it does for countries closer to the world frontier. This suggests that individual aging could be an explanation for the technology differences between developed and developing countries.

To examine the impact of population aging on growth, this thesis constructs models combining population aging, discrete time overlapping-generation, endogenous economic growth and education in general equilibrium. Under both autarky and international trade environments, the impacts of population aging on three major aspects are studied, and they are educational efforts, directed technical change and skill premia. Population aging tends to increase educational efforts and the rate of technological progress. The impacts of population aging upon directed technical change depend on the relative strength of price and market size effects. Moreover, population aging decreases skill premia under autarky, but increases skill premia under trade equilibrium.

Overall, this thesis highlights the important roles of individual and population aging in understanding innovation and economic growth.

Acknowledgments

I would first like to express my sincere gratitude to my supervisor, Professor Alan Woodland, for his guidance, support and encouragement during my PhD candidature. I am very thankful to Professor Woodland for his valuable suggestions and comments on this thesis, and for discussions during our regular meetings.

I also wish to thank the faculty and research students in the School of Economics for their part in creating a vibrant academic environment. This makes the task of a postgraduate student much easier and more enjoyable.

Moreover, I owe a lot to my parents and my wife, for their endless support and encouragement. No word is enough to express my gratitude to them.

Table of Contents

Chapter 1

Introduction

1.1 Background

Well established results from psychology and biology literature confirm that individual aging (the fact of becoming old) has remarkable effects on cognitive skills. The evolution of cognitive skills with age is quite complicated. There are two separate and independent components of human intelligence. The first is called fluid intelligence, which is important for innovation, and this component of human intelligence decreases monotonically with age. The second component is called crystalized intelligence, which increases with experience and is important for knowledge adoption. In contrast with fluid intelligence, crystalized intelligence increases with age which broadens experience.

In modern economic growth theory, in a global context, technological progress of each country consists of technological innovation with each country itself, as well as technology adoption from other more advanced countries. With individual aging, changes in cognitive skills at the individual level will be translated to the national level and individual aging matters a great deal for technology innovation as well as adoption.

Population aging has become one of the most important demographic phenomena facing many countries in the world. One feature of population aging is the shift in the distribution of a country's population structure toward older ages. It is well known that as the baby-boom generation retires, OECD countries' population will age rapidly over the next decades. Two main reasons for population aging are a reduction in fertility and a decline in mortality, resulting in relatively more old people per unit of young.

Numerous studies have analyzed the impacts of population aging upon macroeconomic performance, such as factor accumulation and output growth. Among all impacts of population aging, its implication for human capital investment via education deserves special attention. While studies using neoclassical models argue that population aging has negative effects on the economy, current economic literature utilizing endogenous growth models tends to agree with the proposition that population aging would encourage educational efforts, thus technological progress and economic growth.

Realizing the importance of population aging upon overall educational effort and human capital investment, how different research sectors are influenced by population aging is largely ignored in the literature. Population aging would change the relative amount of various input factors and different research sectors yield technologies augmenting different factors. Therefore, besides the impact of population aging upon overall technological progress, whether different research sectors are affected to the same extent deserves some serious economic analysis.

While the literature studies population aging and education under closed economy, the impacts of aging under international trade is also important. This thesis also analyzes the the relationship between population aging and education in a two-country world under general equilibrium. Moreover, skill premium has increased remarkably during the recent decades and the international trade model allows to investigate whether population aging can be an explanation for this phenomenon.

1.2 Objectives

The main theme of the thesis is to give a comprehensive analysis of aging, including both individual aging and population aging, and economic growth, in the context of endogenous economic growth. In particular, this thesis analyzes the the direct impacts of aging upon technological progress, which is the engine of modern economic growth. Besides the impacts of aging upon technological progress, this thesis also investigates whether technological progress has certain bias in different research sectors. Moreover, international trade models are constructed to study the impacts of population aging upon education and skill premium.

The objectives of the thesis include four aspects, to

1. model individual aging from a more realistic point of view, in accordance with biology and psychology literature, and analyze its impact on technological progress, in an international context with both technology innovation and adoption;

2. amend the classical overlapping-generation model, with survival uncertainty and human capital investment via costly education, and analyze how population aging affects educational effort and other macro economic performance, in a one-sector growth model;

3. extend the above model into a two-sector growth context, so as to analyze the impact of population on the relative bias of technological progress in the two sectors;

4. study the impacts of population aging in an international trade model, upon educational efforts, technological progress and skill premium.

In pursuit of the first objective, it is assumed (in the model of chapter 3) that each agent has two separate abilities related to technological progress, innovation ability and adoption ability. As an agent becomes old, his technology innovation ability could fall, and his technology adoption ability could rise. Technological progress is triggered by research and development at the individual firm level. Anticipating the changes in the employees' technology innovation and adoption abilities, firms decide whether to continue hiring the employees or simply dismiss them. Different countries have different relative incentives and this is reflected in the retaining rules of employees. This difference further translates into different impact of individual aging upon technological progress. This analysis also provides an insight to understand the cross-country differences in their technology levels.

In pursuit of the second objective, an overlapping-generation model is utilized in chapter 4. Each agent has some probability of surviving into the old age and population aging is via an increase in the survival rate. Each agent can choose to do education and become a scientist doing research and development (R&D) upon surviving into the old age, or he can simply stay unskilled. With population aging, the present value of further R&D payoff is higher and this induces more agents to do costly education. Therefore, with population aging, there is an increase in the overall educational effort. As will be shown in chapter 4, the model has the neoclassical growth nature, and the result shows that an endogenous growth model is not necessary to show the positive relation between population aging and educational effort, overall technological progress and economic growth.

Moreover, in pursuit of the third objective, in chapter 5 the above one-sector growth model is extended to a two-sector growth framework, and it features one research sector for each intermediate good sector. Scientists can freely choose to work in the research sector that can give them the higher expected payoff. With this model, chapter 5 investigates how population aging affects the relative numbers of scientists in these two research sectors, the relative phase of technological progress and thus the direction of technical change. The model features both the price and market size effects, as two key factors in the directed technical change literature and it is shown that the price and market size effects jointly determine the bias of technical change. Analysis of the connection between population aging and directed technical change is very rare in the current literature and my model is an early attempt in this aspect.

Finally, chapter 6 of the thesis works with a two-country two-good trade model with an overlapping-generation structure to study the impacts of aging upon education and skill premium under international trade. Survival into the second period is uncertain and population aging is via an increase in the survival probability. The model studies how population aging would affect educational effort and it is found that population aging encourages domestic educational effort, while it discourages educational effort of the other country. This can be labeled the education stealing effect. Moreover, this model studies how population aging affects the skill premium. In autarky, it is found that population aging decreases skill premium. In contrast, in the international trade equilibrium, population aging in any country will increase skill premium in both countries. This is consistent with the empirical finding on the skill premium change experienced by many countries in recent decades.

1.3 Outline

Chapter 2 provides an overview of the literature. Some classical literature on endogenous economic growth, focusing on innovation and technological progress is reviewed. Then a general review of individual aging, population aging, and related economic studies is provided. The main purpose of chapter 2 is to provide contexts and motivations for chapters 3 to 6 of the thesis.

Chapter 3 of the thesis focuses on the relation between individual aging and technological progress, via technology innovation and adoption in an international context.

Chapters 4 and 5 both focus on population aging. Chapter 4 studies the impacts of population aging upon overall educational effort, technological progress and other macroeconomic performance. Chapter 5 extends the main result of chapter 4 into a two-sector growth model, and more importantly, investigates how population aging affects the direction of technical change across different research sectors.

While chapters 4 and 5 focus on closed economy, chapter 6 investigates the impacts of population aging upon educational efforts in an international trade model, with two countries identical in every aspect except the degree of population aging. Moreover, chapter 6 studies the relationship between population aging and skill premium, both under autarky and international trade equilibrium.

In the end, chapter 7 concludes the thesis with an overview of the thesis' contributions. The limitations of the current analysis and suggestions for future research are identified.

Chapter 2

Literature review

2.1 Introduction

This part of the thesis serves as an overall literature review. One of the main focus is on aging and economic growth. First of all the literature on economic growth, innovation, and especially the endogenous growth literature, is reviewed. After reviewing the literature on economic growth, attention is drawn to the literature on individual aging, cognitive skills and economic growth. This topic is relatively unexplored in economics, since the literature is generally from biology and psychology. Individual aging gives problems as well as opportunities for technological progress and economic growth, and countries with different technology levels deal with individual aging quite differently. This not only presents an opportunity to investigate the impacts of individual aging upon economic growth, but it also uncovers a channel for explaining significant cross-country differences in technology levels and economic growth.

Next the literature on population aging is reviewed. Population aging is characterized by a shift in the world population distribution towards older ages, which is now an important phenomena for most countries. Besides reviewing the trends in fertility decline, mortality decline, which gave rise to population aging, more attention is given to population aging's effects upon economic performance. While there exists a significant relationship between population aging, factor accumulation and output growth, what matters more in the modern economic growth literature is the level of overall educational efforts. The literature review includes results concerning population aging and educational efforts, which is studied as one of the key questions in this thesis.

Recently, the directed technical change literature has drawn much attention. This literature gives a better understanding of the modern economic growth, especially when research and development in different sectors can not be treated the same. The current literature hardly investigates the importance of population aging upon the relative direction of technical change, which is very important in the real world. The brief literature on population aging and directed technical change is reviewed, with an attempt to tackle this question in chapter 5 of the thesis.

Moreover, the literature on recent changes in skill premium is reviewed. In recent decades, skill premium has steadily risen in both developed and developing countries. Most of literature focuses on technological change and international trade to explain this phenomenon, while some literature mentions the possible role played by demographic trends in raising skill premium. This literature review lays the background for chapter 7 of the thesis, which studies the relationship between population aging, international trade and skill premium.

2.2 Innovation and Growth

The considerably large cross-country income differences start the exploration into the sources of economic growth and development. One early landmark in this area is the so-called Solow model, first published in Solow (1956). Despite of its simplicity from today's point of view, it lays the foundation for subsequent research in this area. The Solow model features a neoclassical aggregate production and it constructs a simple one-good economy, with little reference to individual decisions. The Solow model emphasizes the role of factor accumulation and provides valuable insights on issues such as equilibrium, convergence, transitional dynamics.

One major limitation of the Solow model is the lack of optimization based on well-defined household preferences. In Solow model, households always save a constant fraction of their income. However, in reality, the consumption and saving decisions are made in a forward-looking manner. This initiated the research on neoclassical growth. In the standard neoclassical model, usually known as the Ramsey or Cass-Koopmans model, households have well-defined preferences, according to which they optimize, making consumption and saving plans. The neoclassical growth model opens the black box of savings and capital accumulation. Moreover, since preferences are explicitly specified, equilibrium and optimal growth can be compared. Roughly speaking, the neoclassical model can be thought as a combination of Solow model and household dynamic optimization.

Another major limitation of the Solow model, shared also by the neoclassical model, is the lack of treatment of the technological progress. In both the Solow and neoclassical models, economic growth is generated by technological progress. However, the progress of technology is totally exogenous. In most cases, a pure number, interpreted as the technological index, grows according to an exogenous rate and totally independent of the economic performance. Alternatively, as in the so-called AK model, growth is sustained by (linear) capital accumulation, or growth is taken place simply as a by-product of knowledge spillovers (a relatively formal treatment can be found in Acemoglu (2009) chapter 11).

Economic growth literature later on formalized the idea that technological progress is a consequence of purposeful investments by individuals (such as scientists) and firms (such as R&D sector in the IT companies). These models endogenize technological progress and discover how it is related with economic factors. Endogenous growth literature explains how economic activities can influence technological progress (via, for instance, changing the incentives to do research), which further affects economic growth. Below three aspects of endogenous growth literature are reviewed: (1) general concepts about technological change; (2) expanding variety models; and (3) Schumpeterian growth models. These have laid the foundation for directed technical change, as will be discussed in detail soon. Expanding variety type and Schumpeterian type models are two pillar stones for the modern endogenous growth literature.

2.2.1 General concepts of technological change

There is an important distinction between macro and micro innovations. Macro innovations refer to more profound and revolutionary discoveries of technologies that can be applied in general. Examples of macro innovations include steam power, electricity and computer, which change the entire world in different ways. Micro innovations, in contrast, focuses more on specific goods or services. Innovations that give new models of existing products (such as from iPhone 5 to iPhone 5S), reduce time costs (the introduction of assembly lines) can be seen as examples of micro innovations. Most of the innovations in economic models are micro innovations.

In regards to the incentive to conduct scientific research, some historians argue that the determinant of innovation is large exogenous and not for the pursuit of profit. In his seminal study, Schmookler (1966, page 206) writes that "invention is largely an economic activity which, like other economic activities, is pursued for gain." Much more recent empirical studies supporting this view include Newell, Jaffee and Stavins (1999), Popp(2002) and Acemoglu and Linn (2004).

Numerous studies investigate the value of innovation in partial equilibrium. The classic reference on the private and social values of innovation is Arrow (1962). Romer (1990) emphasizes the importance of monopoly power for innovation. Nevertheless, growth theory is mainly about general equilibrium. The most popular model in this type is developed by Dixit and Stiglitz (1977), which formalizes the key features of monopolistic competition, discussed early by Chamberlin (1933).

2.2.2 Expanding variety models

The simplest models of endogenous technological change feature a R&D sector where the varieties of inputs or outputs expand. The first type of models has expansion in input varieties. Purposeful research leads to the creation of new varieties of inputs, and a higher variety of inputs increases the division of labor across different production sectors, which further raises the productivity of the final good firms. Two classic references of the expanding input varieties model are Romer (1987, 1990).

The second type of expanding variety models, developed early by Grossman and Helpman (1991a, 1991b), focuses on the expansion of output varieties. In this model, research leads to the invention of new (final) goods for consumption. It is assumed that consumers have a love for varieties, who can enjoy higher welfare if a given amount of budget is spent on more varieties of final goods. This amounts to an increase in real income.

There are some critiques on the endogenous growth models, which mainly point at the scale effect in the sense that a larger population corresponds to a higher growth rate. As mentioned in Backus, Kehoe and Kehoe (1992), larger countries do not necessarily grow faster, such as the more populous countries in the postwar ear. Moreover, Jones (1995) finds that the fraction of resources (labor or final output) devoted to R&D appears to increase constantly, but there is no associated increase in the growth rate. Looking over a very long time period, from one million B.C. to 1990, Kremer (1993) argues that, based on the world population estimates, there must be an increase in economic growth over the past one million years.

2.2.3 Schumpeterian growth models

While the expanding variety models have certain attractive features, they have one common feature: relatively new and old varieties exist at the same time. In reality, most innovations increase the quality of *existing* product, resulting in the new products replacing old products. Models of expanding variety are not a good description of innovation dynamics because they do not capture the competitive aspect of innovations. This gives rise to the whole realm of Schumpeterian growth literature, in which the innovations are creatively destructive: new firms replace initial incumbents and new inputs or products replace old ones. The expanding variety models deal with product innovation, while the Schumpeterian type models deal more with process innovation. Innovations in input or output varieties are called horizontal innovations, compared to innovations raising the product quality, which are called vertical innovations.

The basic model of competitive innovation was first introduced by Aghion and Howitt (1998), and then further developed by Grossman and Helpman (1991a, 1991b). Aghion and Howitt (1998) gives an excellent survey. Schumpeterian growth literature highlights at least two important issues. First, since innovation is done with the presence of the old vintage, there are direct price competitions between old and new vintages, with different qualities or production costs. This competition can have significant effects on the growth process. Secondly, the competition between incumbents and entrants bring about the replacement and business stealing effects.[1] These two effects could cause excessive innovation, which does not exist in the expanding variety literature.

Schumpeterian growth models feature the replacement effect, in that the entrants have higher incentives to conduct radical innovations, and as a result, it is predicted that (nearly) all new innovations should be done by entrants. However, in reality the evidence shows that innovations can be done by both incumbents and entrants, contradicting the Schumpeterian growth models. One recent paper, Acemoglu and Cao (2010), assuming that the incumbents can conduct more radical innovations than entrants, gives a first attempt to resolve this conflict.

2.3 Individual aging, cognitive skills and economic growth

Individual aging, the fact of becoming old, is an event happening to everyone in the world. As a person becomes old, some of his characteristics, such as earning potential, energy level, problem solving skills, vary significantly compared to when he was young. Important as these changes in characteristics are, what matters more are the happenings on the biological level.

Despite of significant inter-personal physical differences, the impact of becoming old upon

[1]It is sometimes called the "Arrow's replacement effect", first pointed out in Arrow (1962), that the monopolist has lower incentives to undertake innovation than the firm in a competitive environment, since the innovation will replace its own already existing profits. Business stealing effect is closed related to the replacement effect. By replacing the incumbents, the entrant is stealing the business from the incumbent.

the mental skills is nearly the same for each single person and is the most prominent among all impacts of individual aging. According to psychological literature, *cognitive skills* refer to mental skills associated with acquiring knowledge and innovating.

Knowledge and innovation, as key components for research and development, play a vital role in economic development and growth. This section reviews some psychological and biological research on individual aging, and some related economic research about cognitive skills for economic growth. The combination of these two literature naturally yield interesting research questions, as will be addressed in chapter 3.

2.3.1 Individual aging and cognitive skills

The relation between cognitive skills and age has been studied intensively by psychologists and biologists and some results have been well established. Instead of treating cognitive skills as an one-dimensional variable lying on a continuum, cognitive psychologists and neuroscientists find that cognitive skills are determined by two types of intelligence that evolve quite distinctly over the life-cycle.

The earliest formal definition is given by Raymond Cattell (1971). The first type of intelligence is *fluid intelligence* (FI), 'the capacity to think logically and solve problems in novel situations, independent of acquired knowledge'. It is the ability to analyze novel problems, identify patterns and relationships that underpin these problems and the extrapolation of these using logic. It is necessary for all logical problem solving, especially scientific, mathematical and technical problem solving. Since fluid intelligence is relatively independent of acquired knowledge, it can rarely enhanced just by getting more experience.

The second type of intelligence is labeled *crystallized intelligence* (CI), 'the ability to use skills, knowledge, and experience'. Crystallized intelligence is the lifetime achievement of an individual, and it is demonstrated mainly through one's vocabulary and general knowledge. In contrast to fluid intelligence, crystallized intelligence improves with age since experiences tend to expand one's knowledge. Fluid and crystallized intelligence are governed by separate neural and mental systems. Consequently, instead of treating cognitive skill as a one-dimensional variable lying on a continuum, it should be viewed as consisting of two variables lying on two different and separate dimensions.

In terms of the impact of individual aging upon cognitive skills, it is found that, (contrary to a commonly held view), is not the same at all for fluid and crystallized intelligence. With individual aging, there is a predictable shift toward more knowledge about how things are done, coupled with a reduction in the speed with which new ideas are grasped (Salthouse, 1990).

Roughly speaking, fluid intelligence is independent of learning or experience and is more important for innovating, while crystallized intelligence is based upon facts and rooted in experiences and is more important for adopting existing knowledge. FI is decreasing in age while CI is increasing in age as more experience is accumulated.

Not only psychologists and biologists, economists have also noticed the two different components of cognitive skill and intelligence. One of the first economic papers distinguishing the difference between fluid and crystallized intelligence is Heckman (1995). In a review article, he criticizes an influential book by Herrnstein and Murray (1994), in that: "inequality is a multidimensional problem, but Herrnstein and Murray consider only a unidimensional version of it. They hold that a single factor of intelligence accounts for inequality in modern society". Heckman (1995) points out that "Indeed, the existence of multiple abilities bolsters the empirical case for heterogeneity in ability as an important fact of social and economic life". Very recently, Rohwedder and Willis (2010) emphasize the fact that the psychological theory of fluid and crystallized intelligence has a clear parallel with economic theories of investment in human capital (in which Ben-Porath (1967) is an early contribution), and how researchers in each discipline have been almost completely unaware of the related literature in the other discipline (Heckman (1995) is an early exception).

2.3.2 Connecting cognitive skills and economic growth

In the modern economic literature, the endogenous nature of economic growth, as well as R&D, technological progress and human capital accumulation via education, is emphasized to a large extent. The important role played by cognitive skills upon economic growth is mainly through the quality of education.

Since the late 1980's, much of the attention of macro-economists has focused on the determinants of long-term economic growth. Nelson and Phelps (1966), Romer (1990) and Rebelo (1991) and others have initiated the literature on endogenous growth models. In the simplest formulation, growth rates are affected by research and development and adoption behavior, and the stock of human capital plays a vital role in the process.

Human capital is thought to accumulate via education, and the raw years of schooling are used to measure the quantity of education. Hanushek and Kimko (2000) is one of the first papers emphasizing the quality of education. Rather than using the raw number of years of schooling, they construct new measures of quality based on student cognitive performance on various international tests of academic achievement in mathematics and science. Their measures of quality are directly related to labor-force skills and productivity of individuals. Using data for OECD, Asian, Latin American and African countries, ranging from 1960 to 1990, they find that labor-force quality differences are important for growth. They argue that math and science skill is a primary component of human capital relevant for the labor force. Such cognitive skill of a population is not well captured simply by measures of school quantities. The labor-force quality differences are related to schooling and that quality has a causal impact on growth.

Barro (2001), using scores on internationally comparable examinations (science, mathematics and reading), investigates the quality of education and economic performance for roughly 100 countries, over three ten-year periods: 1965-1975, 1975-1985, and 1985-1995. Barro finds that given the level of GDP, a higher initial stock of human capital signifies a higher ratio of human to physical capital. This higher ratio tends to generate higher growth through at least two channels. First, more human capital facilitates the absorption of superior technologies from leading countries. This channel is likely to be especially important for schooling at the secondary and higher levels. Second, human capital tends to be more difficult to adjust than physical capital. Therefore, a country that starts with a high ratio of human to physical capital tends to growth rapidly by adjusting upward the quantity of physical capital. Given the quality of education (represented by the test scores), the quantity of schooling (measured by average years of schooling) is still positively related to economic growth. However, the effect of quality of education is quantitatively much more important.

More recently, the study by Jamison, Jamison and Hanushek (2007) confirms the positive impacts of education quality upon economic growth. They find that tests in mathematics appear to be measuring an element of human capital that is important to growth in income per capita and that is not captured by quantity (years) of schooling on its own. Their study involves more countries and over more time periods with additional controls. More importantly, they begin to investigate the *mechanism* underlying the positive effect of school quality upon growth. They test whether educational quality may operate through the level of output (country fixed effects), through the rate of technical progress, or through the size of the increment to output caused by an increase in a country's average quantity of education. Using the data from 62 countries, with 10-year intervals from 1960 to 2000, they find that the strongest support for the idea that quality impacts economic output via changes in the rate of technical progress. Moreover, they find such positive impact depends importantly on the openness of the economy. Education quality improves productivity most significantly in countries open to trade, thus pointing out the importance of technological diffusion and adoption. This supports the idea that education improves productivity most significantly in an economic environment that is open to outside trade and influence.

In a survey article, Hanushek and Woessmann (2008) point out that cognitive skills are most important in deciding the quality of education and they argue that cognitive skills are THE key issue. They claim: "*schooling that does not improve cognitive skills, measured here by comparable international tests of mathematics, science, and reading, has limited impact on aggregate economic outcomes and on economic development*." They find that in the regression with economic growth as the dependent variable, including the cognitive skills measures makes the coefficient on years of schooling goes to zero. Conventional measures using years of school attainment or enrolment rates in schools, though readily observed and convenient to obtain, are very misleading in the policy debates. They also include factors other than cognitive skills (such as established property rights, open labor and product markets, participation in international markets), but they find cognitive skills independently affect economic outcomes. In the end they make a very positive conclusion on causality: "*While it is difficult to establish conclusively that this is a causal relationship, the robustness of the result lends considerable credence to such an interpretation. The relationship does not appear to result from particular data samples or model specifications. Nor can it be explained away by a set of plausible alternative hypotheses about other forces or mechanisms that might lie behind the relationship*."

2.3.3 Motivation for chapter 3

The impacts of individual aging upon crystallized and fluid intelligence are quite different, and these two types of intelligence play different roles for technological progress.

Though noticing the importance of cognitive skills on technological progress, most economic studies treat cognitive skills as constant throughout lifetime. Studies in psychology and neuroscience show that an agent's cognitive skills vary a lot throughout his lifetime (Hunt, 1995). Given that cognitive skills are important for technological progress, the changing pattern of cognitive skills should be considered in economic analysis. An early economic study that considers how individual aging affects abilities dates back to Clark and Spengler (1980). They argue that there is an inverse-U-shape relation between ability and age. While this relation is true for abilities such as productivity in manual work which requires energy and strength, it is not for cognitive skills. To deal with the current modeling limitations, this thesis (Chapter 3) will work with a relatively realistic assumption about individual aging and ability change, by having both components of human intelligence included.

Technological progress consists of technology adoption from more advanced countries and innovation within each country itself (Acemoglu, 2009; Acemoglu, Aghion and Zilibotti, 2006). For each country, the relative importance of adoption and innovation depends on its technology level. While technology innovation is more important for countries with high technology levels (advanced countries), technology adoption is more important for countries fallen behind. Moreover, the impacts of individual aging on technology innovation and adoption skills are distinct. Individual aging lowers innovation skills while raising adoption skills. Due to the different importance of innovation and adoption for various countries and the distinct impacts that individual aging have upon innovation and adoption abilities, a natural questions that arises is: whether the impact of individual aging upon technological progress differs for countries with different technology levels (advanced vs. lagged countries). In chapter 3 it is found that individual aging causes further *divergence* between advanced and lagged countries.

Individual aging lowers fluid intelligence which in turn slows down technology innovation. On the other hand, aging raises crystallized intelligence which in turn fasters technology adoption. Therefore, individual aging yields a *slower innovation vs. faster adoption* tradeoff for technological progress. As mentioned before, technology innovation is more important for advanced countries while adoption is more important for lagged countries. As a result, it would be expected that individual aging will cause technological convergence among countries. Hunt (1995) expresses a similar opinion: "(individual) aging increases the value of a workforce when the workplace is static, but it may decrease the value of the same workforce if the methods and technology of the workplace are changing".

Current aging literature focus mostly on *population aging* (Weil, 2006; Lee, 2003; Hock and Weil, 2006; Cutler et al., 1990). Individual aging, on the other hand, has not received enough attention.[2] Individual aging is a problem as important as, if not more than, population

[2]Population aging usually means a shift in the distribution of a country's population toward older ages

aging. Individual aging happens to everyone in every country, so it is much more widespread than population aging, which happens mostly to industrialized countries (United Nations, 2004; United Nations, 2007). Chapter 3 gives an early attempt to point out the importance of individual aging, as a value-added to the aging literature.

2.4 Population aging

Population aging has become one of the most important demographic phenomena facing many countries in the world. One feature of population aging is the shift in the distribution of a country's population toward older ages (Weil, 2006). It is well known that as the baby-boom generation retires, OECD populations will age rapidly over the next decades (Fougere and Merette, 1999). In most industrialized countries, birth rates have been below the replacement rate, and life expectancy has been increasing, leading to population aging (Zhang et al., 2003).

From a historical point of view, demographic transition starts with mortality declines, followed by fertility reduction. These two factors first lead to an interval of increased, and then decreased population growth, and finally, to population aging (Lee, 2003).[3]

This section reviews some stylized facts about population aging. In particular, the historical trends of mortality and fertility declines are reviewed. This helps to understand the past, present and future of the world'd demographic structure.

2.4.1 Mortality decline in recent decades

Besides a decline in fertility rate, there has been also a reduction in the mortality rate worldwide since decades ago. Since 1950, advances in the treatment and prevention of communicable diseases coupled with improved nutrition and sanitation and with effective methods to control the vectors of disease have brought about significant declines in mortality around the world (Patton, 2011). Inequalities in access to food, safe drinking water, sanitation,medical care

(Weil, 2006); or an increase in the dependency ratio (ratio of dependants to the working-age population). Individual aging, at a superficial level, means the mere fact that an agent becomes old. Population aging focuses on changes of the aggregate age distribution and structure in a society, while individual aging focuses on changes of each individual's abilities when he becomes old. Population aging happens at a national level, while individual aging occurs at a personal level.

[3] A lower fertility could reinforce the decline in mortality. As Lee (2003) suggests, if low fertility is associated with increased human capital investments per child, then these might lead to longer life for those children eventually.

and other basic human needs give rise to wide disparities in mortality levels across countries and regions. However, the reduction of mortality, particularly child and maternal mortality, is part of the internationally agreed development.

Over the past sixty years there have been a tremendous gains in survival, with life expectancy for the world's population increasing from 48 years in 1950-1955 to 68 years in 2005-2010 according to estimates presented in United Nations (2011(b)).[4]

The pace of improvement in life expectancy has varied vastly across countries. By 2005-2010, life expectancy at birth in the more developed regions (excluding Eastern Europe[5]) has passed 80 years, while the average length of life was 6 years shorter in Latin America and the Caribbean,[6] 11 years shorter in Asia,[7] 15 years shorter in developing Oceania[8] and nearly 25 years shorter in Africa (Figure I in United Nations (2012) page xi includes more details).[9] In addition to United Nations (2012), Patton (2011) gives an excellent cross-country comparisons about youth mortality and the corresponding causes.

The increases in survival among both children and adults account for the improvement in life expectancy. However, the relative contribution of different age groups changes in the entire demographic transition. In populations with low life expectancy at birth in the early stages of their demographic transitions, more importantly is the survival improvements among children, rather than among adults. As life expectancy at birth increases, the marginal improvement contributed by progress in older-age survival is more dominant. Figure II.2 in United Nations (2012) page 4 and figure II.3 in United Nations (2012) page 5 contain more details, decomposing the changes in life expectancy at birth between 1950-1955 and 2005-2010

[4]United Nations (2011(a)) describes in great details, the data requirements to estimate child and adult mortality, together with the major large-scale demographic household survey programs, basic child and adult mortality indicators that can be derived from the surveys and different approaches to calculate estimates of these indicators.

[5]Life expectancy at birth increased in Eastern Europe over the period 1950-1955 to 1965-1970 from 64 years to 70 years, but then stagnated and even declined somewhat at various periods over the next several decades. By 2005-2010, life expectancy at birth in Eastern Europe was 70 years, no higher than the level estimated in 1965-1970. Figure II in United nations (2012) page 3 has more details.

[6]Latin America and the Caribbean saw an increase in life expectancy from 51 years in 1950-1955 to 73 years in 2005-2010.

[7]Asia saw a 26-year increase in life expectancy at birth, from 43 years in 1950-1955 to 69 years in 2005-2010.

[8]Developing Oceania saw a 25-year increase in life expectancy by 2005-2010.

[9]In Africa the average length of life increased by 17 years over the last half century, from 38 years in 1950-1955 to 55 years in 2005-2010.

according to the contributions of mortality decline in various age groups.

Increases in life expectancy with the demographic transition is a common pattern, known as the epidemiologic transition. First posited by Omran (1971), this is characterized by initial declines in the death rates due to communicable diseases in the early states of the transitions, followed subsequently by reductions in mortality because of non-communicable diseases in the advanced stages of the transitions (United Nations, 2012).

In regards to how and where mortality is headed during the coming decades. On the optimistic side, Oeppen and Vaupel (2002) suggests that it could reach 97.5 years by mid-century and 109 years by 2100. Less optimistic projections by Lee and Carter (1992) and Tuljapurkar, Li and Boe (2000) suggest more modest gains for the high-income nations of the world, with average life expectancy approaching 90 years by the end of the twenty-first century.

2.4.2 Fertility decline in recent decades

One of the key characteristics of demographic change all over the world is a decline in the fertility rate during the recent decades. According to United Nations (2011(b)), at the world-wide level, the total fertility rate was 4.45 during 1970-75, 3.04 during 1990-95 and only 2.52 during 2005-10. This global trend is similar in all areas. In Europe, total fertility rate (the average number of children a woman would bear if fertility rates remained unchanged during her lifetime) was 2.17 during 1970-75, 1.57 during 1990-95 and 1.53 during 2005-10, very close to the replacement rate. In Northern America, it was 2.05 during 1970-75 and slightly lower at 2.03 during 2005-10. Actually the total fertility rate has always been low in Europe and North America, compared to other areas like Asian, Arab region and so the decline in fertility is more profound and significant outside Europe and North America.

In the Arab region, in the early 1950s, the total fertility rate was 7.2 births per woman. Virtually all countries were characterized by a total fertility rate above 6.5 (18 countries) and a majority above 7.0 (15 countries). Five decades later in 2005-10, the average fertility has declined by more than one-half (56 percent) to 3.1 births per woman. At present the estimated total fertility rate falls below 2.5 births per woman. Much potential for substantial further decline in most countries in the region has been argued for (Casterline, 2011).

In India, the second populous country in the world, from the early 1950s, the 6.0 percent total fertility rate declined to about 2.7 percent in 2007. Moreover, it is projected that it will continue to fall to 2.1 percent in 2034 (Haub, 2011). Dividing 1961-2011 into ten-year time intervals, India's decade growth in population has been above 20 percent, ranging from 21.5 to 24.8 percent (for more information, see Table 1 in Haub (2011)).

The fertility prospect in East Asia is even more pessimistic than India. In the first decade of the twenty-first century, some East Asian countries have undercut the European countries characterized by "lowest-low" fertility of below 1.3 percent. China's total fertility rate has been estimated to range from 1.33 to 1.44 between 2001 and 2005. It has been below replacement level for almost two decades, since about 1990. Other East Asian countries, such as Thailand, Vietnam, Myanmar and Indonesia are all low-income countries where fertility has fallen to close or even below replacement level (Jones, 2011).

While the increase in longevity is due to advances to medical technology, which mostly is beyond the choice, how many children to bear is under control. There are some economic studies trying to investigate the economic reasons behind the fertility decline. Most economic theories argue that couples wish to have a certain number of surviving children. Under this view, when parents recognize an exogenous increase in child survival, fertility should decline. However, the mechanism could be very complicated. For example, as argued in Nerlove (1974), parents might want to invest more time and resource rearing a smaller number of children and this is the initial reason for improvement in children survival. In this process, economic factors influence parents' decisions.

One notable study by Galor and Weil (1996) points out that increases in capital per worker raise women's relative wages, since capital is more complementary to women's labor input than to men's. Rearing children is time intensive and usually women have primary responsibility for child-rearing. Therefore, a rise in the relative wages of women makes children more costly relative to other activities. This is the reason leading to a decline in the fertility rate. At the same time, rising incomes shift consumption demand toward non-agricultural goods and services, and education is particularly important for these sectors. A rise in return to education naturally leads to increased education investments. The overall effects are: children become more expensive, especially for educated parents, so parents with higher incomes choose to have fewer children while devoting more resources to each child.

2.5 Population aging and economic performance

Realizing the importance of demographic transition, numerous studies have investigated the macroeconomic implications of population aging. This section reviews the literature on population aging's impacts on economic performance, with particular focus on the effects of population aging upon saving, factor accumulation and economic growth. First, the literature where endogenous technological progress is absent is reviewed. Then the literature where endogenous technological progress is accounted for in economic growth is reviewed. Because in the endogenous growth literature, the role of human capital is prominent (see, for example, Lucas (1988)), the review will focus especially on the impacts of population aging upon human capital investment, via purposeful education.

2.5.1 Population aging, factor accumulation and economic growth

Different types of evidence have been produced to shed light on the relationship between demographic structure, saving and capital accumulation. In the early literature, there are some cross-country aggregate studies dealing with saving behavior and demographics (see, for example, Graham (1987), Koskela and Viren (1989) and Masson et al. (1996)). Usually they use a saving regression with various factors including demographics, interest rates, income, wealth, inflation and others. The results show that populations with relatively more young people (younger than 30 years of age) or old people (older than 65 years of age) tend to have lower savings rates.

Besides using aggregate data, some studies use micro data, with the availability household level information. From Miles (1999), there is a survey in six of the major economies (Canada, Germany, Italy, Japan, United Kingdom, United States), where the saving rates of different age groups are summarized. It is found that: (1) the scale of any effect from demography upon saving behavior is small, and the age-saving relation appears to be flat; (2) aging could even slightly *increase* saving rates in the long run. This is inconsistent with results from most cross-country studies using aggregate data.

While studies using aggregate data and micro data are useful sources, calibrated general equilibrium overlapping-generation models are more widely used in this area. First introduced by Auerbach and Kotlikoff (1987), followed by studies such as Auerbach et al. (1989), Miles (1999) and Hviding and Merette (1998), the effect of aging upon macroeconomic performance is studied intensively. Their results suggest that population aging will lead to a reduction in national saving rates and in real output per capita over the next decades. These papers all use large-scale simulations, which makes the mechanism complicated and different conflicting forces difficult to see.

2.5.2 Population aging and endogenous technological progress

Studies in the previous subsection have a common feature: they are silent on how economic activities affect the technological progress. Technological progress is either absent from their models (such as Auerbach and Kotlikoff (1987) and Hviding and Merette (1999)), or simply exogenously (with respect to economic factors) given and evolving (such as Miles (1999) and Auerbach et al. (1989)). This treatment of technology ignores the relationship between demography and technological change. Technological progress is considered to be a more important factor for economic growth, than merely (physical) capital accumulation. Actually, since last century, the connection between demography and technological change has drawn major attention. One early notable paper by Cutler et al. (1990) points out that, "We have only scratched the surface in assessing the macroeconomic implications of demographic change. Among the main priorities for future research...any effects of demography on the rate of technical change are likely to dwarf its other consequences." Therefore, the missing of modeling technological progress could yield limitations on implications of the models.

Actually, though omitting technological progress in their simulated models, Hviding and Merette (1998, p.30) admits that a purely neoclassical production function fails to include any positive spillover effects from increased human capital investment. They also postulate that aging populations may stimulate human capital investment.

While early simulation studies on population aging agree on the negative impact upon real output per capita, it has been observed that, related with economic growth are a decline in mortality, an increase in life expectancy, and a rise in human capital investment (Ludwig and Vogel, 2010). In their seminal work, Barro and Sala-I-Martin (1995) find that a 13 year increase in life expectancy is estimated to raise the annual growth rate by 1.4 percentage points. Since a decline in mortality rate and an increase in longevity are key features of population aging, the empirical observation suggests a positive relation between population aging and economic growth. This contradicts the conclusion from early studies with exogenous technological progress. Some other early evidence similar to Barro and Sala-I-Martin (1995) can be found in Ram and Schultz (1979) and Eckstein et al. (1999).

Using a large panel of countries, Feyrer (2007) examines the impacts of demographic structure on productivity. Instead of focusing only on dependency ratio, he examines different age groups and the corresponding productivity. He finds that movements into the 40-year-old group are associated with higher productivity. In contrast, cohorts aged 50-59 and 60-plus are associated with lower productivity. Moreover, cohorts aged 15-39 are associated with significantly lower productivity as well. In a theoretical model with standard 2-period OLG setup where there is no difference between young and working age, Feyrer (2007)'s result implied a negative relation between aging and productivity. This paper also conducts cross-country comparisons and suggests differences in demographic structure is another reason for income differences among countries.

Kögel (2005) also studies the impacts of demographic structure on productivity and he finds youth dependency is negatively correlated with productivity and he agrees with Feyrer (2007) in attributing international income differences to differences in demographics. Beaudry et al. (2005) examines the same question, from the viewpoint of technological transition from 1970 to mid-1990s. He finds the rate of labor growth is a key factor determining a country's adjustment to the new technology and he concludes that demographic difference accounts for much of the cross-country difference in economic performance, as opposed to other factors emphasized in the literature.

The modern economic growth theory, developed by Romer (1986) and Lucas (1988), emphasize factors such as learning by doing, spillover effects, especially human capital investment and R&D as key components in economic growth. Lucas (1988) has shown that human capital, accumulated through purposeful schooling and other forms of investment, serves as the engine of economic growth. Therefore, to resolve the above conflict, introducing endogenous technological progress seems a natural remedy.

One early study by Fougere and Merette (1999) analyzes an endogenous growth model with both physical and human capital accumulation. They use a computable OLG model for seven industrialized countries and model population aging via a lower birth rate. In their model human capital is non-rival while physical capital is rival, hence from the viewpoint of the whole economy, human capital accumulation is better than physical capital accumulation in booming economic growth. When aging occurs (lower birth rate), physical capital is relatively abundant compared to labor, and this decreases the return on physical capital while raising wage rate (since labor is now more scarce). Since the payoff of accumulating human capital (via early schooling) is future wage rate, a higher wage rate makes schooling more attractive. As a result, agents invest less in physical capital and more in human capital. This would stimulate economic growth and reduce any negative impacts that population aging could have upon output per capita.

After Fougere and Merette (1999), many more recent studies have focused on the connection between population aging, economic growth, education investment and human capital accumulation. These studies all make use of the overlapping-generation model, either in discrete time (such as Ludwig and Vogel (2010)), or the continuous time (such as de la Croix and Licandro (1999), Hu (1999), Kalemli-Ozcan et al. (2000), Boucekkine et al. (2002), Echevarria and Iza (2006), and Heijdra and Romp (2009)). These studies highlight the importance of labor force quality in economic growth progress. They find that an increase in longevity, as a key feature of population aging, encourages human capital investment and this contributes to economic growth. The general intuition behind their results is: with increased longevity, the payoff of accumulating human capital (usually via schooling during the young age) is realized for a longer period and this enhances the discounted present value of doing education, which raises educational efforts.

de la Croix and Licandro (1999) model population aging as a decrease in the birth rate. As people live longer, they will do more education when young. This is because education raises working income, and longer working time due to an increase in longevity makes doing education more attractive. Assuming education corresponds to accumulating human capital, increased longevity results in accumulating more human capital, hence the effect of longer life expectancy on human capital accumulation is unambiguously positive.

In contrast to the above models, where production only requires one input (labor) so the general equilibrium effect of demographic shock onto relative factor prices is ignored, Ludwig and Vogel (2010) propose a model where both physical capital and effective labor are inputs into production process. They study the impacts on physical and human capital accumulation, following both a lower birth rate and a higher survival rate. In their model, agents have some uncertainty of surviving to their second old stage. They find a lower birth rate increases the accumulation of physical and human capital. However, the impact of a higher survival rate upon (physical and human) capital accumulation is ambiguous. A higher survival rate decreases the effective discount rate, which gives higher incentive to accumulate capital when young. At the same time, in their model agents spend an exogenous fraction of time working when old, so a higher survival rate also means a higher income in the old stage, and this decreases the incentive to save in the young stage. This is a key difference against standard OLG model where agents have no income at all when old, and this is the reason why they get ambiguous impacts of a higher survival rate upon capital accumulation.

2.5.3 Motivation for chapter 4

Early studies without endogenizing technological progress, especially simulations of large-scale overlapping-generation models, find a positive impact of population aging upon physical capital accumulation and hence economic growth. However, a simple and theoretical model can potentially yield the same insights with better and clearer intuition. To this end, chapter 4 of the thesis constructs a simple overlapping-generation model and analyzes how population aging affects physical capital accumulation, saving behavior and hence economic growth and welfare.

Despite of papers in the previous subsection having technological progress endogenous, compared with the early models with exogenous technology, many of the above models assume that the final production sector uses only one factor, either human capital or effective labor (taking skill level into account), or they with only small open economies to pin down the interest rate.[10] In this way, the general equilibrium feedback effect of population aging upon relative prices (wage rate and interest rage) is absent by construction.[11] Models such as Hu (1999), Kalemli-Ozcan et al. (2000) and Ludwig and Vogel (2009) overcome the partial equilibrium problem in a tractable way, including both physical capital and effective labor in the final production. Though in general equilibrium, the above models assume that all agents are identical and will all accumulate some human capital. There is no isolation between unskilled and skilled labor. In reality, both unskilled labor (such as manual workers) and

[10]For tractability, the literature has eliminated the endogeneity of interest rate. For example, Boucekkine et al. (2002) achieve the constancy of the interest rate by assuming that the felicity function is linear, i.e. that the inter-temporal substitution elasticity is infinite. Heijdra and Romp (2009), closely related to Boucekkine et al. (2002), assume that the economy is of a small size and have access to well-functioning markets including the world capital market thus the interest rate is exogenously given and constant. The latter has the advantage that they can postulate a concave felicity function, giving rise to well-defined consumption profiles.

[11]On the other hand, early studies with exogenous technological progress are in general equilibrium, with both wage rate and interest rage endogenized.

skilled labor (such as scientists) exist simultaneously, doing different kinds of jobs (such as manual work and academic research respectively).

Chapter 4 of the thesis also investigates how population affects education affect, in a general equilibrium model. In contrast with the current literature, the model features both unskilled and skilled labor at the same time, and what is more, it is shown that this is a possible outcome even if all new born agents are identical. Unskilled labor provide labor service and physical capital, while skilled labor provide research machine, and they all enter into the final production function. In this manner, the model has more general equilibrium feedback than the current literature. Population aging affects the consumption and saving decisions of unskilled labor, and the formation and size of skilled labor, and this in turn affects wage rate, interest rate, and price of research machine simultaneously.

Early literature with exogenous technological progress finds a negative impact of population aging upon physical capital accumulation, output and economic growth, while recent literature emphasizing the endogeneity of technological change finds the opposite. This seems to suggest that having endogenous technological progress is necessary for population to boom economic growth. This concern is implicitly mentioned in Fougere and Merette (1999).

However reasonable it seems, there is no paper showing that having endogenous technological progress is necessary for population aging to have positive effect on macroeconomic performance and economic growth. To this end, the third goal of chapter 4 is to analyze: in a model with essentially neoclassical growth progress, can population aging have positive effect on physical capital accumulation, output and economic growth?

As emphasized earlier on, the model features human capital accumulation and endogenous technological progress. To have an essentially neo-classical growth feature, an insight introduced in Acemoglu (2009, chapter 13) is borrowed. The marginal cost of research machine is specified in a special way such that the resulting model is neoclassical in essence (more details in chapter 4). The current literature focuses on the quality of labor force. With population aging, there is an increase in the investment of human capital, which in turn booms economic growth. However, the neoclassical channel is somewhat ignored in the literature. This is the reason that I intentionally remove the positive effect of higher machine quality on the output growth. The output growth and capital accumulation are essentially neoclassical in nature, where the endogenous technical progress plays no role. Though having removed this effect, I still find that population aging has positive impacts on output growth and this is all due to physical capital accumulation. In this aspect, Chapter 5 uncovers a new channel, having been ignored till now, through which population aging can boom economic growth.

2.6 Directed technical change

Previous models of economic growth share one feature: technologies of different sectors evolve in the same fashion. For instance, in Romer (1986), Aghion and Howitt (1992) and Jones (1995), there is only one sector and technology level of this sector represents technology of the whole economy. In Grossman and Helpman (1991), even though there is a continuum of sectors and a continuum of types of technologies, in equilibrium all sectors are the same and the technology of any one sector again represents the overall technology. As pointed out by Acemoglu (1998), there is a significant bias of technical change in various sectors (for example, skilled labor intensive versus unskilled labor intensive sectors), with more R&D investment conducted towards certain sectors. This gives rise to the directed technical change (DTC hereafter) literature. This section reviews some early related literature, usually known as the *induced innovation* and then the modern treatment of DTC, initiated mainly through the works of Acemoglu, is introduced.

2.6.1 Early view: induced innovation

The modern DTC literature is largely influenced by and closely related to the earlier literature on induced innovation. This literature dates back to the influential work by Hicks (1932). In his book, *The Theory of Wages*, page 124-125, he wrote: "*A change in the relative prices of the factors of production is itself a spur to invention, and to invention of a particular kind– directed to economizing the use of a factor which has become relatively expensive*." Later on, Kennedy (1964) introduced the concept of "innovation possibilities frontier". He argued that, instead of the shape of given neoclassical production function, it is the form of this innovation possibilities frontier that determines the factor distribution of income. In his famous work, *American and British Technology in the Nineteenth Century: The Search for Labor-Saving Inventions*, Habakkuk (1962) argued that labor scarcity and labor-saving inventions were central to technological progress.

As can been seen, the induced innovation literature focuses largely on the effect of relative price on the direction of technical change, with more research effort put into the technology that augments the relatively scarce factor. While this gives valuable insight, it overlooks another important factor, the size of the market into which the technology is applied. This idea is well captured by a famous quote from Matthew Boulton, James Watt's business partner, who wrote to Watt: "It is not worth my while to manufacture your engine for three countries only, but I find it very well worth my while to make it for all the world" (Scherer (1984), page 13).

The problem with emphasizing too much more about the relative price than the market size it similar to that in a partial equilibrium. Relative prices are not primitive, but determined and affected by other economic factors. If only the relative prices are emphasized, various other important factors might be missed out in analysis. The directed technical change literature takes both of the price and market size effects into consideration.

2.6.2 Modern view: directed technical change (DTC)

Modern literature on DTC is initiated by Acemoglu (1998, 2002, 2003, 2007, 2009), among others. The start of this literature is inspired by a phenomena in labor economics. As shown in figure 15.1 in Acemoglu (2009), in the U.S., the relative supply of skills (defined as the number of college equivalent workers divided by non-college equivalents) has increased steadily since the later 1960s. During the same time period, there has been no tendency for the returns to college to fall. On the other hand, the wage premium has risen, which suggests an upward-sloping demand curve for skills, contradicting the traditional economic theory. The standard explanation is that new technologies over the postwar period have been skill biased, whose advancements favor more skilled labor relative to the unskilled.

This idea was firstly formalized in Acemoglu (1998). Two conflicting effects are pointed out that jointly determine the direction or bias of technological progress. The first is the price effect, where the incentives to develop certain technologies are stronger if the goods produced using these technologies have higher prices. At the same time, there is a second factor, the market size effect, where it is more profitable to develop technologies if the goods produced using these technologies are sold in a larger market. R&D is driven by profitability, and hence the price and market size effects together determine the direction of technical change. These two effects are conflicting. Since prices are endogenized in general equilibrium, relatively more scarce factors command higher price, and the price effect drives more R&D to this sector. However, at the same time, since the factor is scarce, the market size effect will drive R&D

favoring technologies in other sectors. There is always a relative tension in the process. For the skill market in the U.S. since the late 1960s, apparently the market size effect has dominated the price effect.

2.6.3 Motivation for chapter 5

Population aging can have vital effects on educational effort, human capital investment and overall technological progress. Technical change is considered to be the engine of economic growth, so any effect that population aging has upon technical change matters for economic growth. Many studies investigate the relation between demographic structure and economic growth. However, there is little research investigating the effect of population aging upon the direction of technical change, and this gives much room for future research.

To my best knowledge, the first paper (and the only theoretical paper till now) dealing with population aging and directed technical change is Irmen (2009). Irmen devised a new neoclassical growth model with endogenous capital- and labor-saving technological change. The channels via which population aging occurs are (1) a decline in the birth rate and (2) an increase in the survival rate. There are one final good, two intermediate goods (labor and capital intensive) and two factors (labor and capital), where two intermediate firms (labor and capital intensive, respectively) conduct labor- and capital- saving technologies respectively. A lower birth rate or a higher survival rate leads to relatively more capital than labor, which induces changes in relative scarcity of factors, and in turns relative prices. As labor becomes more expensive compared to capital, more innovation is directed to labor-saving technology, away from capital-saving technology. Assuming both labor- and capital-saving technology are conducted in equilibrium, this model implies that the long-run capital-labor ratio is unchanged with population aging. This further implies that the growth rate of aggregate labor-saving technology in the long run is unaffected by aging. This model defines economic growth as the growth in aggregate labor-saving technology, hence a conclusion is long-run economic growth is unaffected by aging. However, the long-run growth rate of the capital-saving technology is lower due to aging (simply because some innovation is directed away from capital-saving innovation) and if economic growth is defined alternatively as some the growth rate of some combination of both labor- and capital-saving technology, his model would yield a negative relation between the long-run economic growth rate and population aging.

Irmen's model is very different from the standard DTC literature. Standard DTC literature features both the price and market size effects. However, Irmen introduces production capacities in the two intermediate sectors and this essentially rules out the market size effect by construction. It is exactly due to this reason that his model implications are consistent to what is implied by solely the price effect. It is valuable to include the market size effect and get a general equilibrium effect of population aging upon the direction of technical change. This not only makes the model richer but also has a better resemblance with the current DTC literature. This will be tackled in chapter 5 of this thesis.

2.7 Population aging, technical change, international trade and skill premium

In recent decades, there has been a remarkable increase in skill premium, in near all developed and many developing countries. Autor, Katz and Kearney (2008) plots the US full-time weekly wage distribution from 1963 to 2003 and they find the gap in pay between college and high school educated workers have risen consistently since 1979. Similar rises in skill premium since 1970s are found in the UK and other OECD countries as well (Acemoglu and Autor, 2010; Gosling et al., 2000; Atkinson, 2007). In the developing countries, there is less systematic evidence for the evolution of the wage distribution, but wage premium has also risen in many developing countries (Parro, 2013; Burstein et al., 2013, Anderson, 2005; Goldberg and pavcnik, 2007). In fact, the rise in skill premium is not just a recent event. Van Zanden finds that skill premium rose consistently in England and Florence/Milan during the period of 1300-1800.

The rise in skill premium represents a form of wage inequality, which is important for economic welfare. Many studies try to explain the rise in skill premium, mainly from two points of view: (1) skill-biased technical change and (2) international trade. These studies are shortly reviewed below.

2.7.1 Skill-biased technical change

In recent decades, there is a co-occurrence between the rise in skill premium and remarkable technological progress. This fact has lead to many studies exploring the effect of technological progress upon skill premium.

In the past 30 years there has been significant technological progress, especially in information and communication technologies. Following the technological progress, there is a worldwide decline in the relative price of capital goods.[12] Therefore, many studies support the idea that technology is embodied in capital goods (Parro, 2013).

Declined relative price of capital goods will increase the relative demand of skilled labor via a mechanism called capital skill complementarity. In an early study of Griliches (1969), it is argued that compared to unskilled labor, skilled labor are more complementary to capital goods, the so-called capital skill complementarity. After Griliches (1969), many studies provide empirical evidence consistent with the hypothesis of capital skill complementarity. For example, Autor, Kazt and Krueger (1998) find a positive relation between computerization and the employment of skilled labor. More recently, Koren and Csillag (2011) find that imported machinery and relative wages of skilled labor are positively correlated, using firm-level data for Hungary. Following modern technological progress, capital goods are relatively cheaper, and the production sectors have used more of capital goods. Because of capital skill complementarity, the relative demand for skilled labor will increase. Following the same logic, more use of capital goods will decrease the relative demand of unskilled labor, because capital goods substitute for unskilled labor more than it does for skilled labor. An increase in the relative demand for skilled labor, together with a decline in the relative demand for unskilled labor, lead to a higher skill premium.[13]

Though the capital skill complementarity can explain the rise in skill premium, it is treating the direction of technical change (skilled biased in this case) exogenous. As mentioned in the directed technical change literature by Acemoglu, the direction of technical change is endogenous, caused by a combination of price and market size effects. The fact that the market size effect dominates the price effect explains why skill premium has risen since 1970s, while the relative supply of skilled labor also rose during the same period.

In a more general sense, capital skill complementarity and directed technical change can be unified. According to Acemoglu (2003), skilled-biased technical change can be defined as “any change in technology that increases the aggregate demand for skills.” Therefore, the capital skill complementarity argument can be treated as a form of skill-biased technical change.

[12] Parro (2013) Figure 1 plots the change in the relative price of capital goods for many countries during 1990-2007. The relative price of capital goods has declined in all countries in his sample.

[13] Formal models featuring capital-skill complementarity include Jovanovic (1998) and more recently, Krusell et al. (2000).

2.7.2 International trade

Besides skill biased technical change, many papers study the rise in skill premium from an international trade point of view. One trigger lies in the fact that many trade models focus on the effects of international trade upon wage inequality.

In the classic international trade model, the two-country two-good two-factor (skilled and unskilled labor) Heckscher-Ohlin model, the Stolper-Samuelson effect predicts that, international trade should increase skill premium in skill-abundant countries while decrease that in skill-scarce countries. According to this theorem, since developing countries are relatively unskilled labor abundant, international trade should lead to a decrease in skill premium in developing countries. This is in sharp conflict with the observed data, where skill premium has risen in both developed and developing countries. Davis and Mishra (2006) point out recently that "[i]t is time to declare Stolper-Samuelson dead."

Another more serious problem with the trade explanation on skill premium increase is related with the relative price of skill-intensive goods. According to Acemoglu (2003), any trade explanation essentially means that trade should increase the relative price of skill-intensive goods, and this in turn will raise the derived demand for skills. However, this argument is inconsistent with empirical evidence, where the relative price of skill-intensive goods have stayed constant or declined.

Due to above two arguments, it is believed that international trade can not explain the rise in skill premium. However, even though international trade per se does not give satisfactory explanations for the rise in skill premium, it could raise skill premium by inducing skill-biased technical change. A key mechanism for skill biased technical change is the rise the relative price of skill intensive goods. Since international trade can increase the relative price of skill-intensive goods, it could be an important factor in skill premium increase. Acemoglu (2003) proposes a formal model studying the relationship between international trade, skill-biased technical change and skill premium. He argues that the two explanations for the increase in the demand for skills, trade and technology, may be related.

Some studies incorporate capital skill complementarity into trade and explain the rise in skill premium. Under the assumption that technical change has the feature of capital skill complementarity, international trade will affect skill premium via its impacts upon equipment accumulation and capital goods prices. If technical change has increased skill premium via decreasing the costs of capital goods, a reduction in trade costs should also increase skill premium because trade liberalization decreases capital goods costs. Following a decline in trade costs, capital goods are cheaper, and since capital goods substitute more for unskilled labor, skill premium will rise. Parro (2013) calls this *skill-biased trade*. Parro (2013) combines skill-biased trade with skill-biased technical change and constructs a multi-country multi-sector general equilibrium model of international trade, and quantifies the contribution of each force to the increase in skill premium. Both trade and technical change are skill-biased, and both could potentially induce an increase in the skill premium. He finds that without capital-skill complementarity (so that only the standard Stolper-Samuelson effect is operating), the magnitude of the effect of trade on the skill premium is very close to zero . With capital skill complementarity, the impact of skill-biased trade upon skill premium is much larger than Stolper-Samuelson effect, and it is of a similar magnitude to the effect of skilled-biased technical change.

Papers closely related to Parro (2013) include Burstein et al. (2013) and Burstein and Vogel (2010). Burstein et al. (2013) construct a model evaluating the impact of trade upon skill premium via the impact of trade upon the accumulation of capital equipments, and they study the relationship between observable changes in import shares by sector to changes in real wages of skilled and unskilled workers. Their model gives a mapping between the extent of capital skill complementarity and the strength of trade's effect and they quantify the importance of this effect for a large set of countries. Burstein and Vogel (2010) construct a model of international trade and multinational production to examine the impact of globalization on skill premium in skill-abundant and skill-scarce countries. They find that, if technology is skill biased, as trade costs decline, the relative demand for skill increases because labor shifts within sectors towards the most productive producers, which have the highest skill intensity. As a result, trade liberalization increases the relative demand for skill, analogous to the effect of skill-biased technological change.

2.7.3 Motivation for chapter 6

The technological progress literature explaining the rise in skill premium relies heavily on the assumption of skill biased technical change. It requires that the technical change must be biased towards the skilled labor in order to explain why skill premium has increased. Moreover, in the international trade argument, skilled biased technical change is necessary in order to explain the rise in skill premium. The standard two-country two-good two-factor trade model (Heckscher-Ohlin model) yields results in sharp contrast with empirical findings. Only after being combined with skilled biased technical change, can standard trade models explain the rise in skill premium.

While skill biased technical change could be important for the rise in skill premium, whether it is a necessary component is still an open question. Chapter 6 aims to solve the problem: could any factor other than skill biased technical change explain the rise in skill premium? Moreover, while chapters 4 to 5 of the thesis all focus on close economy for tractability, it is natural to study the impacts of population aging in an international trade model.

Studying population aging as a possible reason for the skill premium increase is triggered by two studies. First, Acemoglu and Autor (2011) point out that skill biased technical change is caused by an increase in the relative supply of skilled labor which is possibly due to demographic trends. However, Acemoglu and Autor (2011) does not explicitly model how population aging can increase the relative supply of skilled labor. Second, the study by Van Zanden finds that skill premium rose consistently in England and Florence/Milan during the period of 1300-1800, and this period is associated with population growth. These two studies both imply population aging could be an important explanation for skill premium increase and a related study of this issue can contribute to the literature on population aging, international trade and skill premium. This problem is investigated in chapter 6 of the thesis.

Chapter 3

Individual aging and technological progress

3.1 Introduction

Individual aging is the fact of becoming old. As a person becomes old, some of his characteristics, such as earning potential, energy, problem solving skills, vary significantly compared to when he was young. Individual aging has one of its most prominent effects on *cognitive skills*–mental skills associated with acquiring knowledge and innovating. Since the ability to acquire knowledge and innovate is essential for Research and Development (R&D) and technological progress, changes in cognitive skills due to individual aging might have impacts on technological progress. The first question this chapter addresses is that, since individual aging affects agents' cognitive skills, how a country's technological progress is affected by individual aging.

Technological progress consists of technology adoption from more advanced countries and innovation within each country itself (Acemoglu, 2009; Acemoglu, Aghion and Zilibotti, 2006). For each country, the relative importance of adoption and innovation depends on its technology level. While technology innovation is more important for countries with high technology levels (advanced countries), technology adoption is more important for countries fallen behind.

Moreover, the impacts of individual aging on technology innovation and adoption skills are distinct. Individual aging lowers innovation skills while raising adoption skills. Due to the different importance of innovation and adoption for various countries and the distinct impacts that individual aging has upon innovation and adoption abilities, the second question this chapter addresses is: whether the impact of individual aging upon technological progress differs for countries with different technology levels (advanced vs. lagging countries). This chapter shows that individual aging causes further *divergence* between advanced and lagging countries, in terms of technological progress.

Recently, economists have noticed the importance of cognitive skills in economic growth, via the impacts on the quality of education.[1] However, current economic literature either assumes that cognitive skills stay constant throughout lifetime (such as Jamison, Jamison and Hanushek (2007), Hanushek and Woessmann (2008), and Hanushek and Kimko (2000)),

[1] Chapter 2, section 2.3 gives a literature review on cognitive skills and economic growth.

or there is an inver-U-shape relation between ability ana age (Clack and Spengler, 1980). These two assumptions are incorrect in terms of how cognitive skills evolve with age.

Cognitive skills consist of two components, fluid intelligence (FI) and crystallized intelligence (CI). Psychology and biology literature finds that fluid intelligence is more important for innovating, while crystallized intelligence is more important for adopting existing knowledge. Moreover, fluid intelligence is decreasing in age while crystallized intelligence is increasing in age. This implies that with individual aging, there is a predictable shift toward more knowledge about how things are done, coupled with a reduction in the speed with which new ideas are grasped (Salthouse, 1990).

This chapter models the impacts of individual aging on two aspects of cognitive skills (FI and CI), in a manner consistent with psychological evidence. It is assumed that as an agent becomes old, there is some probability that his innovation ability (which is linked to FI) will fall while his adoption ability (which is linked to CI) will rise. This contributes to the literature in that (1) changes in cognitive skills (due to individual aging) have significant impacts on technological progress and so assuming constant cognitive skills (the case for most of current literature) is problematic; (2) the impacts of individual aging on cognitive skills are modelled in a more realistic manner, rather than assuming a simple (yet incorrect in this situation) inverse-U-relationship.

Individual aging lowers fluid intelligence which in turn slows down technology innovation. On the other hand, aging raises crystallized intelligence which in turn enhances technology adoption. Therefore, individual aging yields a *slower innovation vs. faster adoption* tradeoff for technological progress. As mentioned before, technology innovation is more important for advanced countries while adoption is more important for lagging countries. As a result, individual aging could cause technological convergence among countries. Hunt (1995, page 18) argues that "*(individual) aging increases the value of a workforce when the workplace is static, but it may decrease the value of the same workforce if the methods and technology of the workplace are changing*". Since most modern developed countries are experiencing fast technological progress via innovation while technological progress in developing countries is rather slow, Hunt's argument would imply that aging slows down technological progress of developed countries and does the opposite to developing countries, and this should cause a convergence in economic growth.

However, evidence in the economic growth literature shows a divergence among countries, especially in the last 50 years (Acemoglu (2009), chapter 1 provides an overview). Since technological progress is considered to be the engine of economic growth, economic divergence implies a technological divergence.

From the previous two paragraphs, there is a conflict between Hunt's argument and empirical evidence of economic growth. The first contribution of this chapter is to solve the above conflict by pointing out an important issue: retaining rules of old agents. Advanced countries put more value on innovation than adoption. Knowing individual aging lowers agents' innovation skills, firms of advanced countries will dismiss old agents and hire young agents, who have higher innovation abilities on average. Countries fallen behind, on the other hand, tend to retain old agents, even though old agents on average have lower innovation abilities. The difference in retaining rules of old agents explains why individual aging could cause countries to diverge in technological levels.

The second contribution of this chapter is to point out the above reason explaining divergences among countries. While current literature focuses on factors such as the quality of schooling (Barro, 2001; Jamison, Jamison and Hanushek, 2007), institutions (Hall and Jones, 1999; Acemoglu, Johnson and Robinson, 2001) or demographic structure (Feyrer, 2007) in explaining the divergence among countries, the effects of individual aging are neglected. By emphasizing the impacts of individual aging on cognitive skills, this chapter uncovers another channel for growth divergence.

The rest of the chapter is organized as follows. Section 3.2 describes the basic model structure, including the problems facing firms and agents. This chapter focuses on technology evolution for each country and this is solved in two steps in sections 3.3 and 3.4. Section 3.3 solves the optimization problems of firms and agents during each period, given the technological level of that period, and this is named *static equilibrium*. Using the choices of agents and firms in each period, section 3.4 solves for how technology evolves through time, and this is called *dynamic equilibrium*. Sections 3.3 and 3.4 together establish the evolution of technological progress of each country. Section 3.5 investigates the impacts of individual aging, in terms of how individual aging could affect technological progress of each country and the main analytical results are established. Following the analytical results in section 3.5, section 3.6 provides a numerical study of this chapter. In section 3.6 several data sets will be chosen randomly and they are used to test if the analytical results in section 3.5 are correct. In the end, section 3.7 concludes the chapter with a summary of the main results and some future research ideas.

3.2 Model

3.2.1 Basic structure

This chapter only deals with individual aging and there is no population aging. In this chapter *aging* always refers to *individual aging*. This chapter's model is based on Acemoglu, Aghion and Zilibotti (hereafter AAZ) (2006), but extends it to allow for individual aging. In a world economy with many countries, technological progress of each country consists of technology adoption from the world frontier and innovation within each country itself. The speed of technological progress is dependent on agents' abilities (high or low).

Since this model is closely linked to AAZ, two main differences between this model and AAZ are highlighted here. Details of the differences are produced in sections 3.2 and 3.4.

1. **Adoption skill:** In AAZ, agents differ in their technology innovation ability, but all agents have the same adoption adoption. This chapter allows heterogeneities in both adoption and innovation skills (full details in section 3.2).

2. **Age-dependent abilities:** In AAZ, each agent's ability is unchanged with age. In contrast, this chapter models a probabilistic formation of changes in abilities with age, based on evidence from psychology (full details in section 3.4).

These two differences make this model more realistic, and new insights obtained from this will be developed in section 3.5. The first difference is motivated by evidence from neuroscience and psychology, suggesting that cognitive skills change with age. Moreover, while individual aging affects both innovation and adoption abilities, these two effects are not in the same direction. While individual aging decreases an agent's innovation ability, it increases the agent's adoption ability and this presents a trade-off. Explicitly modelling both adoption and innovation abilities allows to see the trade-off and this is why adoption ability heterogeneity is introduced. Section 3.4 discusses how this model is a general version of the model in AAZ (2006). Section 3.4 how illustrates how putting special restrictions on this model results in the model of AAZ as a special case. Section 3.5 shows taking both adoption and innovation abilities into account could yield ambiguity about the impacts of aging on technological progress.

This chapter works with an economy where everyone is risk-neutral and lives for two periods, young and old. There are two types of agents, principals and workers. Principals own factories but have no skills, while workers have skills but no ownership of factories. One principal must hire one technology adopter and one technology innovator to operate his factory, and the words factory and firm will be used interchangeably. The term *manager* means either a technology adopter or innovator. All workers not hired as managers serve as manual labor in final good production.

In each generation, the mass of principals (also called *capitalists* in AAZ) is 1/2, while that of workers is $(N+2)/2$.[2] Workers are equally productive if they serve as manual labor (in the final good sector), but they differ in their adoption and innovation abilities. For simplicity, this chapter assumes there are two levels of adoption (innovation) abilities: high and low. Each new-born young worker has probability $\lambda^A \in (0,1)$ to have high adoption ability, and $1-\lambda^A$ to have low adoption ability ($\lambda^I \in (0,1)$ is defined likewise for innovation).[3]

There is a unique final good that is produced using labor and a continuum of intermediate goods. The final good is also the unique input to each intermediate good. Final good is the numéraire with price fixed to one.

Production function for the final output is

$$y_t = \frac{1}{\alpha} N^{1-\alpha} \int_0^1 A_t(v)^{1-\alpha} x_t(v)^{\alpha} dv\,, \qquad \alpha \in (0,1), \tag{3.1}$$

where y_t is the output level in period t, N is the mass of labor employed in the production, $A_t(v)$ is productivity of intermediate good v (or technology of firm v), and $x_t(v)$ is the amount of intermediate good v used in production. Each intermediate good is produced by only one firm (thus the mass of firms equals the mass of intermediate goods). v is the index for intermediate good. v also indexes the firm producing intermediate good v.

Production function (3.1) is homogeneous of degree one in labor (N) and intermediate goods ($x_t(v)$'s). The final producer is assumed to take all output and input prices as given,

[2] For each generation during each period, the mass (size) of principals is $1/2$. Summing up two generations at each time, the mass of principals, $1/2+1/2=1$, equals the mass of adopters, equals the mass of innovators. The mass of manual labor is $(N+2)-1-1=N$ which will be employed in the final good sector. As a result, labor market clears at every time.

[3] The probabilities λ^A and λ^I are assumed to be independent of agent's ex post job positions. In other words, being hired as an innovator does not increase this worker's innovation ability and the same applies to adoption ability. Relaxing this assumption is mentioned in the conclusion.

and this implies the profit of the final production sector is zero. Therefore, the ownership of the final production sector is irrelevant.

In this chapter one unit of each intermediate good is produced using one unit of final output, hence the marginal (also average) cost of each intermediate firm is 1 (recall the final output with price set to one is the numéraire). There is a fringe of additional firms who can imitate the existing firms and produce the same intermediate good, with the same productivity $A_t(v)$. Exogenous reasons, such as technological gaps or government regulations, require imitators to charge a price of at least χ. Since the marginal cost of each existing intermediate firm is one (recall each intermediate good requires one unit of final good to produce, and the price of final good is set to unitary), it is assumed that $\chi > 1$ so the cost of the fringe of additional firms producing any intermediate good is higher than existing firms. Here, χ can be seen as the degree of competition in the intermediate good industries, between existing firms and potential entrants.[4] A higher χ corresponds to more monopoly power of existing firms and more protection for them against potential entrants. The assumption $\chi > 1$ implies existing firms will prevent entry by charging the price

$$p_t(v) = \chi. \tag{3.2}$$

The final good sector is assumed to be perfectly competitive and take all input and output prices as given. The final good producer will maximize its profit, with respect to the inputs, $x_t(v)$'s and N. The problem of the final producer is expressed by the formula below

$$\max_{N,x_t(v)} Profit = \frac{1}{\alpha}N^{1-\alpha}\int_0^1 A_t(v)^{1-\alpha}x_t(v)^{\alpha}dv - w_tN - \int_0^1 p_t(v)x_tdv,$$

For the producer's problem, taking first order conditions with respect to $x_t(v)$ gives the price of intermediate good v expressed as

$$p_t(v) = \frac{\partial y_t}{\partial x_t(v)} = \left(\frac{A_t(v)N}{x_t(v)}\right)^{1-\alpha}. \tag{3.3}$$

Combining equation (3.2) and (3.3) yields $x_t(v) = A_t(v)N\chi^{-\frac{1}{1-\alpha}}$, and gives the profit of intermediate firm v as

$$\pi_t(v) = (p_t(v) - 1)\,x_t(v) = \delta A_t(v)N\,, \qquad \delta \equiv (\chi - 1)\chi^{-\frac{1}{1-\alpha}}. \tag{3.4}$$

[4]Here the existence of competition is only potential, since firms other than the existing firms are not in the market yet, hence this model specification is quite different from monopolistic competition. What is more, since pricing rule of each intermediate firm is set to χ which is exogenous, this model does not need to explicitly consider potential impacts on prices from competitions among intermediate firms, and this chapter can abstract from (monopolistic) competition among intermediate firms.

It is natural to have the profit of each intermediate firm being positive and increasing in the price they charge, which is the case if an only if $\chi < \frac{1}{\alpha}$. Therefore, this chapter imposes the reasonable assumption that $1 < \chi < \frac{1}{\alpha}$. A higher level of χ or δ corresponds to more monopoly power of intermediate firm v and thus higher profit, given $A_t(v)$ and N.

The competitive wage rate is solved by taking first order condition in the producer's problem with respect to N, and it is expressed by

$$w_t = \frac{\partial y_t}{\partial N} = \frac{1-\alpha}{\alpha} \chi^{-\frac{\alpha}{1-\alpha}} A_t, \tag{3.5}$$

where $A_t \equiv \int_0^1 A_t(v)dv$ is a technology index. In this model the mass of intermediate firms is one, hence $A_t \equiv \int_0^1 A_t(v)dv$ incorporates the technology of all intermediate firms in the country and A_t represents a country's overall technological level.

3.2.2 Technology evolution

Technological progress consists of technology adoption and innovation. The world technology frontier at time t, $\bar{A}_t$, is assumed to grow at an exogenous rate g.[5] It follows that

$$\bar{A}_t = \bar{A}_0(1+g)^t.$$

It is assumed that no country has technology that is more superior than the world frontier, thus $A_t \leq \bar{A}_t$. The measure of a country's distance to the world technology frontier is defined as

$$a_t \equiv A_t/\bar{A}_t,$$

where $a_t \leq 1$ due to the assumption $A_t \leq \bar{A}_t$. A higher a_t means a country is closer to the world technology frontier.

The technology evolution process in this chapter is based upon AAZ but extends it in a significant way to allow for individual aging. In AAZ (2006)'s model, technology evolves because intermediate firms conduct research projects. In each firm, there is both technology adoption and innovation. One firm hires one manager and the manager's ability can be high or low, which affects the productivity of the firm. The firm can choose the size of the project, larger or small, and this also affects the productivity.

In AAZ (2006) equation (8), technology evolves according to

$$A_t(v) = s_t(v)\left[\eta \bar{A}_{t-1} + \gamma_t(v) A_{t-1}\right].$$

[5]Endogenizing g does not qualitatively change the results. The appendix shows how to endogenize g.

$\eta\bar{A}_{t-1}$ represents technology adoption from the world frontier in the previous period, and $\gamma_t(v)A_{t-1}$ represents technology innovation of the firm. $\gamma_t(v)$ denotes the manager's ability and it is equal to 1 if the manager has high ability, or $\gamma < 1$ if the manager has low ability. $s_t(v)$ notes the project size and it is equal to 1 if the project size is larger, or $\sigma < 1$ if the project size is small. Other things equal, technological progress is faster if the manager has high ability and if the project size if large.

This chapter generalizes the technology evolution in AAZ in several ways. Firstly, in AAZ model there are two innovation abilities levels, high and low. However, AAZ assumes that all agents have the same adoption ability, represented by η. This model extends AAZ in assuming that adoption ability could be high or low for different agents. AAZ focuses on innovation ability, and hence they put more weight on innovation and assumes away differences in adoption abilities. In contrast, this chapter focuses on differences in both innovation and adoption abilities, and this creates more symmetry in innovation and adoption (they are equally important in this chapter). This model differences makes it more suitable to study individual aging since aging affects both adoption and innovation ability, and more importantly, the impacts are opposite (good for adoption, while bad for innovation). Using this approach, the tradeoff between these two effects can be analysed, which gives a better insight into individual aging. The advantage of having ability differences in both innovation and adoption is that, because individual aging has opposite effects on adoption and innovation, only focusing on one effect of aging (either on innovation or adoption) will yield results with limitations. This will be clear from the analysis in section 3.5.[6]

Secondly, in AAZ model one manager is hired for one intermediate firm, doing both technology adoption and innovation. Because there is only one manager, there is only one project size $s_t(v)$, affecting the (marginal) contribution of both adoption and innovation. In their model $s_t(v)$ is determined only in the innovation sector, hence decision in the innovation sector indirectly affects the adoption sector. As stated in the introduction, fluid and crystalized intelligence are independent and so innovation and adoption abilities should also be independent and this requires an independence between innovation and adoption sector. In this model, to capture the combined but *independent* effects from adoption and innovation, the the projects sizes of adoption and innovation will be separated. To ensure decisions in innovation

[6]Since the focus of AAZ (2006) is not on investigating the impacts of aging, they do not need to consider this, but this chapter should take this into account.

and adoption sectors are made independently, it is assumed that two separate agents work in adoption (adopter) and innovation (innovator) sectors. In this model, adoption manager and innovation manager are doing different and independent jobs, and the retaining rule of the adoption manager is independent from that of the innovation manager (more details of the retaining rules are developed in section 3.3).

With the above extensions, equation (8) in AAZ (2006) is modified to the formula below

$$A_t(v) = s_t^A(v)\left[\theta^A + \gamma_t^A(v)\right]\bar{A}_{t-1} + \left[\theta^I + s_t^I(v)\gamma_t^I(v)\right]A_{t-1}. \tag{3.6}$$

where $s_t^A(v)$ is the project size of adoption and $\gamma_t^A(v)$ is the skill level of the adopter, with $s_t^I(v)$ and $\gamma_t^I(v)$ as counterparts in the innovation sector. $s_t^A(v) = 1$ if the adoption project size is large and $s_t^A(v) = \sigma^A \in (0, 1)$ if the adoption project size is small. $\gamma_t^A(v) = \gamma^A > 0$ if the adopter has high adoption ability and 0 if low. Innovation sector has a similar story, with σ^I and γ^I replacing σ^A and γ^A. $\theta^A > 0$ and $\theta^I > 0$ are included in (3.6) to avoid trivial results. In equation (3.6), $s_t^A(v)\left[\theta^A + \gamma_t^A(v)\right]\bar{A}_{t-1}$ denotes technology adoption from the world frontier, and $\left[\theta^I + s_t^I(v)\gamma_t^I(v)\right]A_{t-1}$ denotes technology innovation.[7] In this model, $\gamma_t^A(v) = 0$, so an adopter with low adoption ability can contribute nothing to adoption; the same logic applies to a low ability innovator. Ceteris paribus, a larger adoption/innovation project is better for a firm, in terms of technological progress.[8]

Using the technology evolution for each intermediate firm, namely equation (3.6), the technology evolution for the whole country can be solved. One country's overall technology consists of technology of all intermediate firms, so integrating (3.6) at t for all v gives the expression for a country's overall technology level during t (namely A_t), and at $t-1$ for all v gives A_{t-1}. Dividing A_t by A_{t-1} yields the equation for technology growth rate as

$$\frac{A_t}{A_{t-1}} = \frac{\int_0^1 A_t(v)dv}{A_{t-1}} = \underbrace{\frac{1}{a_{t-1}}\int_0^1 s_t^A(v)\left(\theta^A + \gamma_t^A(v)\right)dv}_{(Adoption)} + \underbrace{\int_0^1 \left(\theta^I + s_t^I(v)\gamma_t^I(v)\right)dv}_{(Innovation)}. \tag{3.7}$$

Equation (3.7) shows that a country's overall technology growth rate consists of two parts, technology adoption and innovation. From (3.7), the larger a_{t-1} is, the more important is innovation in (3.7). Therefore, innovation is more important for countries close to the frontier, while adoption is more important for countries fallen behind.

[7] This equation includes the case where the most efficient firm adopts from other less efficient firms. This is a limitation, due to the simplicity of the modeling strategy.

[8] The terminology, technological progress, is used as an inclusive term for technology adoption and innovation.

Following AAZ, the investment costs associated with projects, conditional on the project size s, are assumed to be the following

$$k_t^A(s) = \begin{cases} \phi^A \kappa^A \bar{A}_{t-1}, & \text{if} \quad s = \sigma^A \quad \text{where} \quad \phi^A \in (0,1), \\ \kappa^A \bar{A}_{t-1}, & \text{if} \quad s = 1, \end{cases}$$

and

$$k_t^I(s) = \begin{cases} \phi^I \kappa^I \bar{A}_{t-1}, & \text{if} \quad s = \sigma^I \quad \text{where} \quad \phi^I \in (0,1), \\ \kappa^I \bar{A}_{t-1}, & \text{if} \quad s = 1, \end{cases}$$

where κ^A, κ^I, ϕ^A and ϕ^I are exogenous constant. Since $\phi^A \in (0,1)$ and $\phi^I \in (0,1)$, a large project incurs more investment cost. The key point about the project costs is that larger project costs more than a smaller project. As long as this feature is kept, changes to the functional forms in the project costs will not qualitatively change the main results.

Investment costs can be financed either from contributions by managers, or from principals themselves. The principals can borrow from intermediaries to finance the project. Young managers do not have any wealth to finance the project, while old managers can use their retained earning from the previous period. In this way, the old managers' ability to finance the project (fully or partially) could provide incentives for the principal to retain them, rather than hire a young agent. Thus, retained earnings can serve as a shield protecting old managers against the young.

3.2.3 Individual aging and physical abilities

In contrast to AAZ (2006), this chapter allows agents' abilities to vary with age (they assume agents' abilities are unchanged after agents become old). This will change their results significantly about retaining rules of old agents as discussed in section 3.4. As an agent becomes old, his principal, when deciding whether to retain him or not, will take into account of the probabilistic changes in his abilities. The impacts of individual aging on ability, then on retaining rules, will affect technological progress.

As mentioned in the introduction, when a person becomes old, his innovation ability may decline while his adoption ability may increase. Individual aging is modelled in the following manner: at the end of each agent's first period, those with low adoption ability has a probability $p^A \in [0,1]$ to have high adoption ability when they become old; and agents with high innovation ability has a probability $p^I \in [0,1]$ to have low innovation ability when old.

For any particular agent, whether or not his ability really changes is unknown to his principal when deciding the retaining rules for this agent.[9]

In terms of the assumptions about individual aging and ability change in the previous paragraph, two things need to be pointed out here. First, it is true that besides technology adoption and innovation abilities, individual aging also affect agents' ability to do manual work, but this is assumed away in this model (young and old workers are equally productive if they serve as manual labor). This is because this model focuses on the effect of aging on technological progress and this effect has it dominance over any other effects of aging. Second, if $p^A = p^I = 0$, the abilities of all agents do not change with age and this is the case in AAZ (2006). Section 3.4 provides details how restricting the model to a special case yields their results and more importantly, what new results can be obtained and their importance.

For tractability, this chapter assumes that if a worker has low innovation ability when young, he still has low innovation ability when old, namely, there is not a *lower-than-low* innovation ability. While it is more realistic to add another innovation ability level lower than low ability, this is unnecessary in this model. Equation (3.6) assumes $\gamma_t^I = 0$ for a low ability innovator, so a low ability innovator already contributes nothing to technology innovation (but still his principal need to hire him in order to operate the factory). Assuming hiring an innovator does not dis-contribute to technological progress (γ_t^I cannot be negative), it can be said that a low innovation ability is already "too low to decline". Similar but a more restrictive logic justifies why a high ability adopter's adoption ability has no further increase when old.[10]

In this chapter, there is no relationship between the magnitudes of p^A and p^I. In other words, aging's effects on innovation and adoption abilities are independent. Up to now, researches in biology and neuroscience do not have any evidence about relations between changes in crystalized intelligence (CI) and fluid intelligence (FI). In other words, for any randomly chosen person, a decrease in FI does not necessarily lead to an increase in CI, and vice versa.

[9]If the principal knows whether the agent's ability changes *before* the principals determines the retaining rule, this model will end up with trivial results. Here, probabilistic aging presents a risk to principals when deciding whether to retain their previous managers.

[10]The case for adopter requires stronger assumptions. For simplicity, this chapter assumes that there is no adoption ability higher than 'high'.

3.2.4 Contracts between principals and agents

At the beginning of each period, each principal writes one-period contracts to hire two managers, one for an adoption (he is called adopter) and one for innovation (innovator).[11] On the contract the principal specifies three things, which are (i) payment to the adopter and innovator (denoted S_t^A and S_t^I respectively), (ii) project size (denoted s_t^A and s_t^I) and (iii) how much the adopter and innovator need to contribute to the projects (denoted $\widehat{RE}_t^A \leq RE_t^A$ and $\widehat{RE}_t^I \leq RE_t^I$ respectively where RE_t^A is the total retained earnings that an old adopter has from previous period). Here the arguments of all variables are omitted to simplify the exposition. Full details are in the appendix 3.8.2.

It is assumed that young workers' abilities are unknown, to both themselves and the principals. If a young worker is hired as an adopter, at the end of the first period, his adoption skill will be revealed to himself, and the principal hiring him. However it is assumed that his abilities are *not revealed to other agents.* A young principal can only hire young workers, while an old principal can choose from young and old workers. For each young principal, at the end of the first period, after his worker's skill level is revealed, the principal needs to decide whether he will retain the old manager, or hire someone else.[12]

At the end of each period, after the profit is revealed, an adopter(innovator) can take away a fraction equal to $\mu^A(\mu^I)$ of the firm's profit without being prosecuted by anyone. This moral hazard gives a minimum payments to agents, as will become clear soon. Then incentive compatibility for the managers requires that payments to them must be at least a fraction of the ex post profits, namely,

$$S_t^A \geq \mu^A \pi_t,$$
$$S_t^I \geq \mu^I \pi_t,$$

where π_t is the ex post profit of a firm. π_t is calculated from equations (3.4) and (3.6).

The above incentive compatibility rules out long-term contracts where the decision to hire an old agent depends on whether he has stolen when he was young. In this chapter there is

[11] For simplicity, all contract offers are assumed one-period.

[12] In this model setup, the old agents' ability to finance research projects is important for the main results. Consider a young innovator who turns out to be low skill, then without financing issue, his principal will dismiss him for sure since a new hired innovator has a weakly higher ability, given none of them helps the principal to finance the project. However, if the old innovator helps finance the project from his savings, the principal might have incentives to keep him. The ability to finance projects plays a key role in this model.

no commitment for such long-term contracts, and an old agent will be hired, whenever the principal finds it profitable to do so, even if this agent is "dishonest".

The relation between principals and workers are described, starting from a young principal. A young principal offers contracts, and hires two young workers as the technology adopter and innovator respectively. Then at the end of the first period, the principal knows the agents' skills and pays the agents according to the contract offer. After this, at the beginning of the second period, the principal offers new contracts. Now the principal needs to decide whether to retain his previous workers (whose skills are known now), or hire the young.[13]

At first glance, it seems obvious that if a young adopter/innovator turns out to be low skill, his principal should dismiss him and hire a young agent next period. However, young workers are born with no wealth, so they can contribute nothing to finance the project. On the other hand, since an old manager has some retained earning from previous period, he is able to contribute his wealth to finance the project, and this could give incentives to the principal to retain him, even if he is low skill. This shows financing is essential in determining the incentive to retain old agents.

Till now the description of the model is finished. The following two sections solve the model in two steps. In section 3.3, during each period t, each intermediate firm takes the country-wide technology as given and makes firm-level decisions. It is called *static equilibrium*. Section 3.4, conditional on firms' decisions about project sizes and whom to hire as its managers,[14] solves how technology evolves, which is called *dynamic equilibrium* or *endogenized law of motion*.

3.3 Static equilibrium

This section and the next section 3.4 solve the model to find the equilibrium. This section solves the model for each single period. In each period, the technology level is taken as given and intermediate firms decide the finance requirements, payments to managers, project sizes

[13] It is assumed that an old principal cannot hire people who served as manual labor when young, the appendix shows that no old principal will hire people who served as managers for other principals when young. This left the only possibility that if an old principal decides to dismiss his managers, he will only hire from the new-born young.

[14] A firm's decision includes more than project size and its manager, but only these two are directly relevant for technology progress.

and which managers to hire. This is called static equilibrium, as in AAZ (2006). After solving the model for each single period, the next section 3.4 analyzes how the technology evolves over time.

Taking the state of technology at time t as given, a principal maximizes the net value of his intermediate firm and he solves the problem expressed as

$$Max\mathbb{E}_t V_t,$$

$$V_t \equiv \pi_t - S_t^A - S_t^I - \max\left\{k_t^A - \widehat{RE}_t^A, 0\right\} - \max\left\{k_t^I - \widehat{RE}_t^I, 0\right\},$$

where V_t is the net value of an intermediate firm. π_t is the firm's profit, S_t^A and S_t^I are payments to the adopter and innovator. k_t^A is the cost of the project size in adoption and $\widehat{RE}_t^A$ is how much the adopter can afford to finance. If $k_t^A < \widehat{RE}_t^A$, the adoption project can be fully funded by the adopter and the principal needs to pay nothing. If $k_t^A > \widehat{RE}_t^A$, the adopter can not fully fund the adoption project and the principal needs to pay the difference. This shows the amount $\max\left\{k_t^A - \widehat{RE}_t^A, 0\right\}$ is paid by the principal for the adoption project. Similar logic explains that $\max\left\{k_t^I - \widehat{RE}_t^I, 0\right\}$ is paid by the principal for the innovation project. The expectation operator applies because if a principal hires young managers, there is uncertainty concerning the managers' skill levels.[15]

In this model it is assumed that a young principal can only hire young managers, while an old principal has the option to retain his current (old) managers or hire the young. Moreover, the principal must decide project sizes, payments to the managers and how much the managers need contribute to the projects. As in the standard agent-principal model, the principal's decision must satisfy incentive compatibility and participation constraint of the managers.

From the viewpoints of the principals, the static equilibrium is solved in the following steps,

1. conditional on the managers' skill level or expected skill level, the principal decides payment to the managers and how much he wants the managers to contribute whatever projects he chooses, namely, S_t^A, S_t^I, $\widehat{RE}_t^A$ and $\widehat{RE}_t^I$;

2. conditional on the managers' skill or expected skill, S_t^A, S_t^I, $\widehat{RE}_t^A$ and $\widehat{RE}_t^I$, the principal decides project sizes s_t^A and s_t^I;

[15]The arguments of variables in expressing V_t so on for the arguments of variables below where no confusion could arise. Details are in appendix 3.8.2.

3. conditional on S_t^A, S_t^I, $\widehat{RE}_t^A$, $\widehat{RE}_t^I$, s_t^A and s_t^I, an old principal decides whether to retain his current managers or hire new ones. Note that a young principal needs to solve the first two steps, since he can only hire young managers by assumption.

For the rest of this chapter, for expositional convenience, the following notations are used,

YA	for young adopters;
YI	for young innovators;
OLA	for old adopters with low adoption ability;
OHA	for old adopters with high adoption ability;
OLI	for old innovators with low innovation ability;
OHI	for old innovators with high innovation ability.

With technical proofs in the appendix 3.8.1 and 3.8.2, the decisions of principals are presented in the four lemmas below.

Lemma 1 *If the mass of manual labor N is large enough (but still finite), participation constraints of all agents (adopters and innovators) are slack.*

Lemma 2 *In terms of finance requirements from the managers, namely, how much the managers need to contribute to the projects in their respective sector, this model finds that*

1. *because young agents are assumed to be born with no initial wealth, they pay nothing to finance the projects;*

2. *all old managers (OHA, OLA, OHI and OLI) will be asked to finance all their retained earnings from previous period, as long as the maximum amount they can contribute does not exceed the investment cost required in their respective sector.*

Lemma 3 *In terms of **payments** to the managers, this model finds that*

1. *all adopters (YA, OHA, OLA) are paid proportional (μ^A) to the ex post profit in their firm.*

2. *all innovators (YI, OHI, OLI) are paid proportional (μ^I) to the ex post profit in their firm.*

The insights from lemmas 1-3 are that, as long as N is large enough (but still finite), participation constraints of all kinds of managers will be slack, even if they are asked to pay all their previous retained earnings, and even if they are paid the minimum amount consistent with incentive compatibility. This greatly simplifies the analysis below, since knowing managers can be treated so "harshly", the principals will just pay the managers a fixed proportion of ex post profits, and ask them to contribute all their retained earnings from the previous period.

A principal decides the project sizes for managers, conditional on they are working for him and this is presented in the lemma below.

Lemma 4 *In terms of project sizes, the model gives the following results, where*

1. *if a principal hires young agents, he chooses a small project, if δ (or χ) is small; or a large project if δ (or χ) is large;*

2. *if a principal hires old agents, as long as N is large (but still finite), he chooses large projects.*

In this chapter the focus is on the case where young managers run small projects only. The other case gives the same qualitative results (footnote 16 provides more details on this). From lemma 4, a young manager (adopter or innovator) will run a large project if and only if δ is large enough. The profit of each intermediate firm (equation (3.4)) is increasing in δ, and δ is increasing in χ, a measure of how competitive the market is. Intuitively, since young managers can not finance the projects, a principal may not be willing to take the burden for the entire high investment cost (a large project incurs a higher cost). However, if the principal can exploit high profits (a higher δ) then he will be willing to pay all the investment cost and choose a large project for the young manager.

Moreover, old managers will run large projects if N is large enough. This is because that, a large project is always better for a firm, *if* there is no associated investment. As long as N is large enough, old managers can get enough retained earnings from the previous period and fully finance large projects. This gives incentives to the principal to choose large projects since now he can simply ignore investment costs.

Conditional on payments to agents, financing from agents, and project sizes, an old principal decides whether to retain his current managers or hire the young. In this chapter it is assumed that δ is small hence young agents run small projects (footnote 16 provides more details on this). In this model, retaining rules of old managers play essential roles in technological progress. With proofs in appendix 3.8.3, retaining rules of various kinds of managers are presented below.

Proposition 1 *Let*

$$p^A(L) \equiv \sigma^A \lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A} - \frac{\phi^A \kappa^A}{\gamma^A(1-\mu^A-\mu^I)\delta N} < 1.$$

Retaining rules for the (old) technology adopters are

1. $R_t^A(H) = 1$ *(retained);*
2. $R_t^A(L) = \begin{cases} 1 \ (\textit{retained}) & \textit{if} \quad p^A > p^A(L), \\ 0 \ (\textit{dismissed}) & \textit{if} \quad p^A < p^A(L). \end{cases}$

Intuitively, since individual aging never decreases adoption ability, an old manager with high adoption ability will always be retained. For an old adopter with low adoption ability, individual aging could raise his adoption ability with some probability p^A. If p^A is high enough so that he has a high chance to have high adoption ability, he will be retained. If p^A is low, then he will be dismissed. Proposition 1 can be illustrated using the following Figures 3.1 and 3.2.

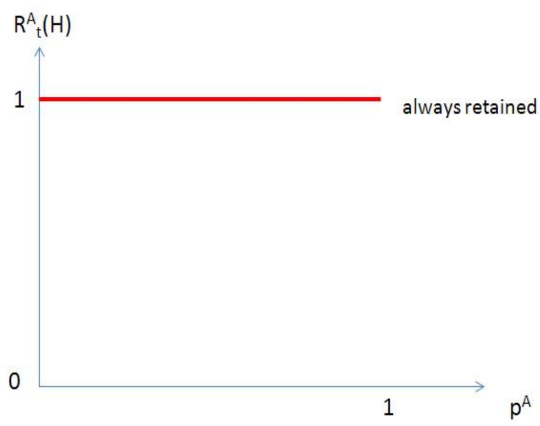


Figure 3.1. Retaining rules for old high skill adopters

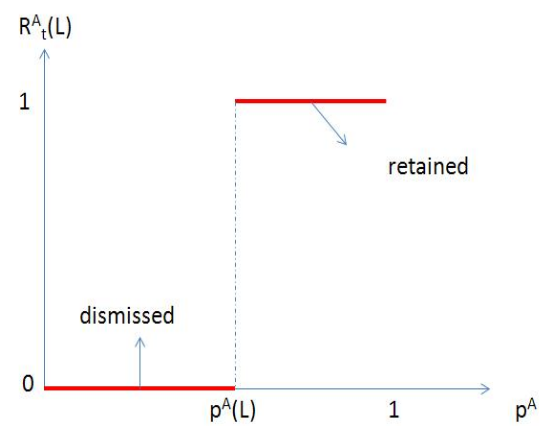


Figure 3.2. Retaining rules for old low skill adopters

In Figures 3.1 and 3.2, retaining rules are represented on the vertical axis, with 0 denoting dismissed and 1 denoting retained. As can be seen from Figure 3.1, an old high skill adopter is always retained, and this is not affected by individual aging, since aging only (weakly) increases an old manager's adoption skill. Figure 3.2 shows that an old, low skill adopter will be retained if p^A is larger than a certain value, namely he has a high enough probability of having high ability when old.

Finishing the retaining rules for adopters, the following proposition presents the retaining rules for innovators, with proofs left in appendix 3.8.3.

Proposition 2 *Let*

$$a(L,I) \equiv \frac{\phi^I \kappa^I}{(1-\mu^A-\mu^I)\delta N \sigma^I \lambda^I \gamma^I};$$
$$a(H,I) \equiv \frac{\phi^I \kappa^I}{(1-\mu^A-\mu^I)\delta N \left(\sigma^I \lambda^I \gamma^I - (1-p^I)\gamma^I\right)}.$$

Retaining rules for the (old) technology innovators are

1. $R_t^I(L) = \begin{cases} 0 \ (dismissed) & if \ \ a_{t-1} > a(L,I), \\ 1 \ (retained) & if \ \ a_{t-1} < a(L,I). \end{cases}$

2. $R_t^I(H) = \begin{cases} 1 \ (retained) \ \ if \ \ p^I < 1-\sigma^I\lambda^I \ \ OR \ \ a_{t-1} < a(H,I), \\ 0 \ (dismissed) \ \ if \ \ p^I > 1-\sigma^I\lambda^I \ \ AND \ \ a_{t-1} > a(H,I). \end{cases}$

The results in Proposition 2 are presented using the following two Figures 3.3 to 3.4, followed by discussions.

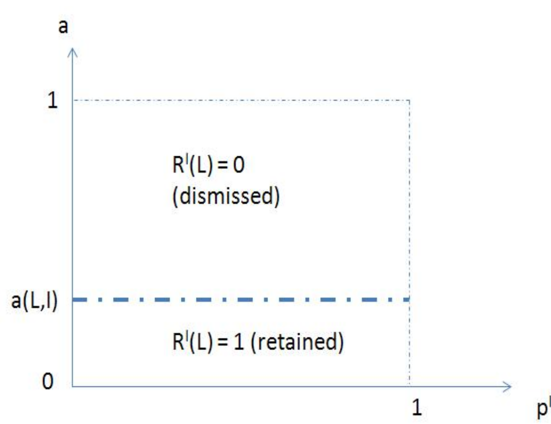


Figure 3.3. Retaining rules for old low skill innovators

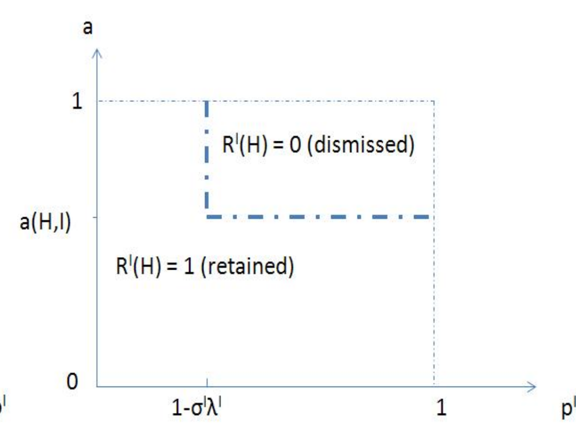


Figure 3.4. Retaining rules for old high skill innovators

In Figures 3.3 and 3.4, depending on values of the distance to frontier a and p^I, the whole space is partitioned into areas of *dismissed* and *retained*. As can be seen from Figure 3.3, old low-skill innovators are retained only in countries with $a < a(L,I)$, namely developing countries. The intuition is: if a country is close to the world technology frontier, innovation is relatively more important, hence a manager with low innovation skill is more likely to be replaced by a young innovator. Note the retaining rule of OLI does not depend on individual aging (p^I does not enter the expression of $a(L, I)$), since individual aging (weakly) lowers innovation skill and an OLI already has the lowest innovation skill level by assumption ($\gamma_t^I = 0$ for OLI).

In Proposition 2, $a(L, I)$ is decreasing in $\lambda^I\gamma^I$. $\lambda^I\gamma^I$ measures the average innovation ability of YI. A higher $\lambda^I\gamma^I$ makes YI more attractive than OLI, and so OLI are more likely to be dismissed, hence a lower $a(L, I)$ (the threshold to dismiss OLI).

Different from $R_t^I(L)$ (which only depends on a), Figure 3.4 shows that the retaining rule for OHI ($R_t^I(H)$) depends on both a and p^I. This is because individual aging only lowers the innovation skill of an OHI (an OLI's innovation skill is the lowest by model construction). Note $R_t^I(H) = 0$ if (and only if) *both* $p^I > 1 - \sigma^I\lambda^I$ *and* $a_{t-1} > a(H, I)$ hold, which means if $a_{t-1} < a(H, I)$, even if p^I is very close to unitary (nearly all OHI become low skilled), an OHI is still retained. Similarly if p^I is not so large, OHI will be retained even if a_{t-1} is very close to one (the country values innovation very much), and this is just because their innovation skill is still 'acceptable'(p^I not so small). Overall, OHI are dismissed if and only if the country is close to frontier (hence innovation skill is important), and OHI has a high probability to become low skilled (hence they can not do innovation well enough, compared to young agents).

Note if $a_{t-1} \in (a(L, I), a(H, I))$, an OLI will be dismissed while an OHI will be retained. And $a(H, I) \geq a(L, I)$ with the equality hold iff $p^I = 1$, which means as long as $p^I < 1$, even a very high probability of becoming low skilled will not result in OHI being dismissed. Intuitively, this is because a higher innovation skill level (even in an expected term) is always preferred, regardless of the country's distance to the world frontier.[16]

Section 3.3 solves the model for each single period, taking the technology level for each period as given. The choices of intermediate firms (project sizes and retaining rules in particular) determine how the technology evolves from one period to the next. This is analyzed in the following section 3.4.

3.4 Endogenized law of motion

While the above section solves for the static equilibrium of each period, this section investigates how technology evolves through time. The evolution of technology depends on project sizes and retaining rules of intermediate firms. Following AAZ (2006), the process of solving technology evolution through time is labeled *dynamic equilibrium*.

3.4.1 Dynamic equilibrium

Conditional on firm-level choices about project sizes (lemma 4) and retaining rules concerning different managers (proposition 1 and 2), technology evolution can be solved using equation (3.6). What directly matters for technological progress are the abilities of adopters and innovators and the projects sizes. Since project sizes depend on whether the manager is old (large projects) or young (small projects), this chapter analyses different cases based on retaining rules for managers.

Retaining rules depend on p^A, p^I and a_{t-1}, as shown in Proposition 1 and 2. According to Proposition 1, OHA are always retained. In contrast, OLA could be retained or dismissed. In the analysis below, the case where OLA are dismissed ($R_t^A(L) = 0$) is labeled case **I** and the case where OLA are retained ($R_t^A(L) = 1$) is labeled case **II**.

[16]It is shown in the appendix 3.8.1 that for a sufficiently large N, all old managers can fully finance large projects, in their respective sectors. Therefore, N is assumed to be large enough and analysis is conducted under the condition that all old managers, if retained, run large projects. Alternatively, without imposing a condition on N, both OHI and OLI must be assumed to able to *fully* finance a large project size. The purpose here is to magnify and focus on the differences in financing abilities between old and young agents, and hence this model abstracts from old agents' differences in financing abilities. If this restriction is relaxed (assuming N is small and OLI cannot fully finance large projects), $a(H, I) > a(L, I)$ even at $p^I = 1$, because now a higher ability to finance a large project size gives the principal even more incentive to retain OHI than OLI, besides OHI's higher (expected) innovation potential.

According to Proposition 2, both OLI and OHI can be either dismissed or retained. In the analysis below, the case where both OLI and OHI are retained ($R_t^I(H) = R_t^I(L) = 1$) is labeled case **i**, the case where OLI are dismissed and while OHI are retained ($R_t^I(H) = 1, R_t^I(L) = 0$) is labeled case **ii**, and the case where both OLI and OHI are dismissed ($R_t^I(H) = R_t^I(L) = 0$) is labeled case **iii**. According to Proposition 2, OHI are always more likely to be retained than OLI, hence it is impossible that OLI are retained and OHI are dismissed. Namely, $R_t^I(H) = 0, R_t^I(L) = 1$ if impossible.

In the analysis below, all possible combinations of the above cases **I-II** and **i-iii** are analyzed. For exposition convenience, the term **II(iii)** is used to denote the case where OLA are retained ($R_t^A(L) = 1$) and both OLI and OHI are dismissed ($R_t^I(H) = R_t^I(L) = 0$).

For an intuitive interpretation, the magnitude of p^A (the probability that a worker with low adoption ability will have high adoption ability when becoming old) can be related with the *severity* of individual aging. In this manner, the more severe is individual aging, the more likely that adoption ability will rise. In this chapter the severity of individual aging is assumed identical for all workers in the same country (but inter-country differences are allowed). Under this interpretation, case **I** corresponds to a country where individual aging is not severe, and case **II** happens in a country where aging is severe enough so that adopters are more likely to be retained when old (since there is a high probability their adoption ability will rise when old). Similarly, cases **i-iii** are ranked according to the severity of aging: the more severely individual aging happens, the more likely workers' innovation abilities fall, and the more likely they will be dismissed when old.

3.4.2 Equations of technological progress

Under each case of subsection 3.4.1, equation (3.6) allows to solve for a country's distance to the world frontier (a_t) as a linear function of a_{t-1} (calculations in the appendix), in the form of

$$2(1+g)a_t = \text{intercept } + \text{slope} \cdot a_{t-1}. \tag{3.8}$$

Appendix 3.8.4 shows how to derive the intercept and slope in (3.8) under various retaining rules, and the results are presented here directly. In terms of the intercept in (3.8), this chapter has

I. OLA **are dismissed.**

$$\text{intercept } \mathbf{I} \equiv \left[\left(\lambda^A + (1-\lambda^A)\sigma^A + \sigma^A\right)\theta^A + \left(1 + (1-\lambda^A)\sigma^A + \sigma^A\right)\lambda^A\gamma^A\right]. \tag{3.9}$$

II. OLA **are retained.**

$$\text{intercept } \mathbf{II} \equiv \left[\left(1+\sigma^A\right)\theta^A + \left(\sigma^A\lambda^A + \lambda^A + (1-\lambda^A)p^A\right)\gamma^A\right]. \tag{3.10}$$

In terms of the slope in (3.8), this chapter has

i. OLI **and** OHI **are both retained.**

$$\text{slope } \mathbf{i} \equiv \left[2\theta^I + (\sigma^I + 1 - p^I)\lambda^I\gamma^I\right]. \tag{3.11}$$

ii. OLI **are dismissed but** OHI **are retained.**

$$\text{slope } \mathbf{ii} \equiv \left[2\theta^I + \left(\sigma^I + 1 - p^I + (1-\lambda^I)\sigma^I\right)\lambda^I\gamma^I\right]. \tag{3.12}$$

iii. OLI **and** OHI **are both dismissed.**

$$\text{slope } \mathbf{iii} \equiv \left[2\theta^I + 2\sigma^I\lambda^I\gamma^I\right]. \tag{3.13}$$

For a specific example, in case **II(iii)**, the technological progress equation is

$$2(1+g)a_t = \underbrace{\left[\left(1+\sigma^A\right)\theta^A + \left(\sigma^A\lambda^A + \lambda^A + (1-\lambda^A)p^A\right)\gamma^A\right]}_{\text{Intercept } \mathbf{II}} + \underbrace{\left[2\theta^I + 2\sigma^I\lambda^I\gamma^I\right]}_{\text{Slope } \mathbf{iii}} a_{t-1}.$$

The distance to frontier a is a proxy of a country's technological level, and by analyzing how individual aging affects the evolution of a, namely the intercept and slope in (3.8), aging's impacts upon technological progress can be analysed. This analysis is done in section 3.5 below. Before that, the following subsection gives a detailed comparison with AAZ (2006).

3.4.3 Comparison with Acemoglu et al. (2006)

Having solved the retaining rules (in section 3) and the evolution of technologies, this chapter's model can be compared with AAZ in full detail. First, the retaining rules (proposition 1 and 2) are compared with AAZ (2006). The retaining rule for OLI is very similar to AAZ in that $R_t^I(L)$ depends only on a_{t-1}, with $R_t^I(L) = 1$ if a_{t-1} is not so large. The retaining rule for OHI in this chapter differs from AAZ. In AAZ, OHI are always retained while in this model they could be dismissed, if p^I is large enough. AAZ assume the abilities of agents are unchanged with age, while this chapter allows OHI to have some probability of becoming low skill when old. Showing $R_t^I(L)$ is influenced by p^I introduces another source of employment rigidity. As will be shown in section 3.5, if $p^I > 1 - \sigma^I \lambda^I$, slope in (3.8) is larger if OHI are dismissed and so technological progress is faster after aging. However, countries with $a_{t-1} < a(H, I)$ will still retain OHI, and this is an obstacle for technological progress.

The model in this chapter is richer than AAZ in the sense that it allows for heterogeneity in the adopter's ability levels, and this shows an *anti-rigidity*. From proposition 1, if p^A exceeds some threshold, OLA will be retained. However, retaining OLA does not necessarily lead to a larger intercept. In this case, the inclination to retain OLA is an anti-rigidity, since the employment is "too flexible". However, this is also an obstacle to technological progress.[17]

This chapter's model nests AAZ as a special case. To see this, equation (3.6) needs to be put on some restriction. AAZ only have an innovation sector, without the adoption sector, and principals only hire innovators in their model. What is more, the project size chosen for the innovator also affects the adoption. Due to these reasons, if the following conditions are imposed, namely $\gamma^A = 0$, $s_t^A(v) = s_t^I(v)$, $\theta^A = \eta$, $\theta^I = 0$, $p^A = p^I = 0$, and $\sigma^A = \sigma^I = \sigma$, the model in this chapter essentially becomes the model of AAZ (σ and η are introduced here simply to keep consistency with their model notations).

This chapter has two types of managers (adopters and innovators) while they have only one (innovators). To relate this model with AAZ, the following conditions are needed, where

1. within each firm, either OHA match with OHI (simultaneously high), or OLA match with OLI (simultaneously low);

2. within each firm, the adopter and innovator are retained simultaneously or dismissed simultaneously.

The above conditions essentially ensure two managers within any firm can be treated as just one manager. Since two managers within each firm are required to be retained or dismissed at the same time, only case **II(i)**, where all old managers are retained, and case

[17]The term *rigidity* means the flexibility of employment rules and it has no unambiguous implications on the speed of technological progress.

I(ii), where all high (adoption and/or innovation) ability managers are retained and all low (adoption/innovation) ability managers are dismissed, are relevant.

In case **II(i)**, where all old managers are retained, the equation for technological progress becomes theirs (the case where $R_t = 1$ in their equation (23)). In case **I(ii)**, where old low ability managers are dismissed but OHI are retained (note in this model OHA are always retained, hence again the same retaining rules hold for adopters and innovators with the same level of ability), and consequently the equation for technology progress becomes theirs (the case where $R_t = 0$ in their equation (23)). In sum, this chapter's model nests that of AAZ (2006) as a special case. In addition some new results are obtained with extra insights, which are absent in their model. These are discussed in the next section.

3.5 The impacts of individual aging upon technological progress

Equation (3.8) shows that the distance to frontier this period a_t, has a linear relationship with a_{t-1}. In this chapter the distance to frontier a_t is the proxy of a country's technological level. Equation (3.8) shows how technology evolves through time. Individual aging's impacts upon technological progress can be fully analyzed through how aging affects equation (3.8), its intercept and slope. If aging increases (decreases) the intercept of (3.8) for a particular country (but slope is unchanged), then this country experiences a sudden increase (decrease) in its technological level (while its technology is increasing at the same speed). If aging increases (decreases) the slope of (3.8) for a particular country, then this country's technology will increase at a higher (lower) speed.

Individual aging's impacts upon adoption ability is represented by p^A while its impacts upon innovation ability is represented by p^I. Equations (3.9) to (3.13) imply that p^A only affects the intercept and p^I only affects the slope. Therefore, aging's impacts upon technological progress, via its impacts upon adoption ability, can be fully analyzed from the intercept. Similarly, aging's impacts upon technological progress, via its impacts upon innovation ability, can be fully analyzed from the slope. The analysis below investigates how individual aging affects the intercept and slope of equation (3.8), separately. These two results can then be combined together to discuss individual aging's overall impacts upon technological progress.[18]

[18] In this model, agents' adoption and innovation abilities are assumed to exogenously change with age. Evidence from psychology (Carttell, 1971; Hunt, 1995) shows while individual aging affects each person's ability, the magnitude of such changes depend on various reasons. Two economically relevant reasons are

3.5.1 Aging's impacts upon technological progress via adoption

In this model, when agents become old, those with high adoption ability when young will remain high adoption ability. Those with low adoption ability when young have probability p^A to become high adoption ability. An increase in p^A can be thought of as an increase in severity of aging, in terms of raising adoption ability.

In the analysis below, the term *severity* is used for the magnitude of p^A (adoption) or p^I (innovation). The term *more severe* means p^A increases or p^I increases. Whether it is referring to adoption (p^A) or innovation (p^I) will be specified by the context. The arguments from the next paragraph till Proposition 3 show how Proposition 3 is derived. The intuition and economic logic behind these will be developed in the two paragraphs after Figures 3.5-3.8.

To see how changes in p^A affects the intercept of (3.8), first note that from Proposition 1, OLA are retained if and only if $p^A > p^A(L) \equiv \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A} - \frac{\phi^A\kappa^A}{\gamma^A(1-\mu^A-\mu^I)\delta N}$. Moreover, intercept **II** $>$ intercept **I** $\Longleftrightarrow p^A \geq \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, namely, aging increases the intercept if it is severe enough in terms of technology adoption. It is easy to see that $p^A(L) < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, so retaining OLA does not necessarily lead to a higher intercept.[19]

The analysis below investigates the case if individual aging becomes more severe, say p^A increases from p_1^A to p_2^A ($p_1^A < p_2^A$) and how this change affects the intercept. If $p_1^A < p_2^A < p^A(L)$, OLA are dismissed both before and after aging becomes more severe. The intercept I are using is intercept **I** (equation (3.9)) both before and after aging becomes more severe. In this case a more severe aging does not affect the intercept, hence no impacts upon technological progress.

If $p_1^A < p^A(L) < p_2^A < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, OLA are dismissed before aging becomes severe but retained after aging becomes severe, so intercept **I** (equation (3.9)) is used before aging becomes severe and intercept **II** (equation (3.10)) is used after aging becomes severe. Note that intercept **II** $<$ intercept **I** $\Longleftrightarrow p^A < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, and so after aging becomes more severe, the intercept becomes smaller, implying a decrease in the level of technology.

---

educational level and the nature of their jobs during adulthood. A person with high educational level and doing innovative jobs (such as academic research) shows much less decrease in his innovation ability after becoming old. In this model, these factors are assumed away, hence differences in p^A and p^I across countries can be seen as due to country-level differences partly from government policy (such as education subsidy). This is left for future research.

[19] This is because OLA could be retained due to their ability to finance project costs, not necessarily due to their expected higher adoption ability after aging.

If $p_1^A < p^A(L) < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A} < p_2^A$, OLA are dismissed before aging becomes severe but retained after aging becomes severe, so intercept **I** applies before aging becomes severe and intercept **II** applies after aging becomes severe. Intercept **II** > intercept **I** $\Longleftrightarrow$ $p^A > \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, and so after aging becomes more severe, the intercept increases, implying an increase in the level of technology.

If $p^A(L) < p_1^A < p_2^A$, OLA are retained both before and after aging becomes more severe and intercept **II** applies both before and after aging becomes more severe. Intercept **II** is increasing in p^A, so after aging becomes more severe, the intercept is larger, implying an increase in the level of technology.

The above results are summarized in Proposition 3 and Figures 3.5 to 3.8 below

Proposition 3 *In a particular country, assume that aging does not affect innovation ability (p^I is fixed) then individual aging does not affect the slope of equation (3.8), hence no impacts upon the speed of technological progress. Now suppose aging becomes more severe in terms of raising adoption ability, which is represented by an increase in p^A, from p_1^A to p_2^A with $p_1^A < p_2^A$. In terms of how this affects technological progress, it follows that (using notation $p^A(L)$ from Proposition 1),*

1. *if $p_1^A < p_2^A < p^A(L)$, a more severe aging has no impacts upon the intercept of (3.8) and no impacts on the level of technology;*

2. *if $p_1^A < p^A(L) < p_2^A < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, a more severe aging causes a decrease in the intercept of (3.8) and a decrease in the level of technology;*

3. *if $p_1^A < p^A(L) < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A} < p_2^A$, a more severe aging causes an increase in the intercept of (3.8) and an increase in the level of technology;*

4. *if $p^A(L) < p_1^A < p_2^A$, a more severe aging causes an increase in the intercept of (3.8) and an increase in the level of technology.*

The four scenarios in Proposition 3 are illustrated using the Figures 3.5 to 3.8 below.

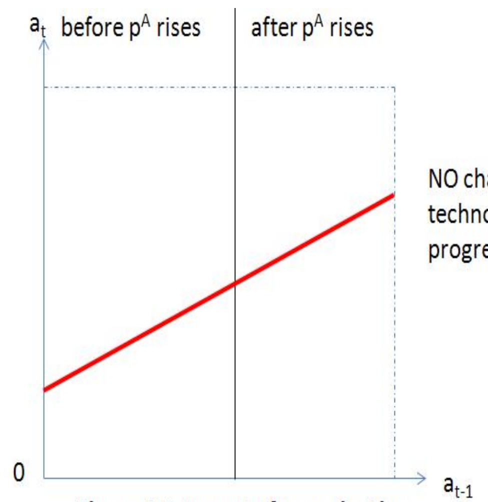


Figure 3.5. Impacts from adoption scenario 1: $p^A_1 < p^A_2 < p^A(L)$

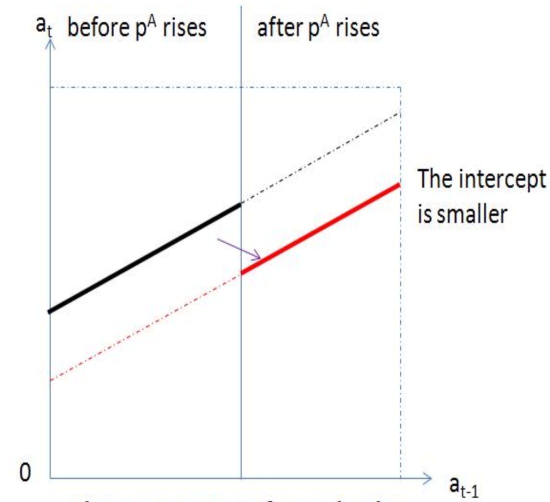


Figure 3.6. Impacts from adoption scenario 2: $p^A_1 < p^A(L) < p^A_2 < \sigma^A\lambda^A - (1-\sigma^A)\theta^A/\gamma^A$

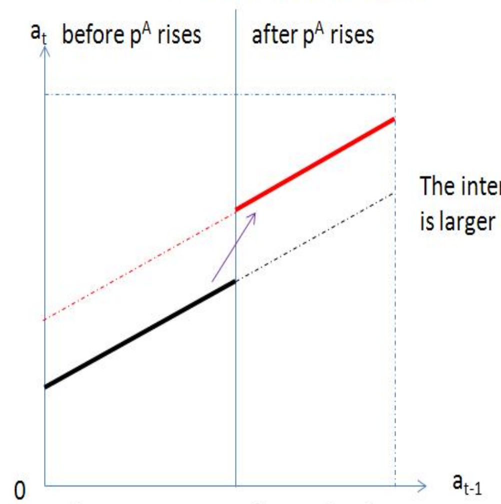


Figure 3.7. Impacts from adoption scenario 3: $p^A_1 < p^A(L) < \sigma^A\lambda^A - (1-\sigma^A)\theta^A/\gamma^A < p^A_2$

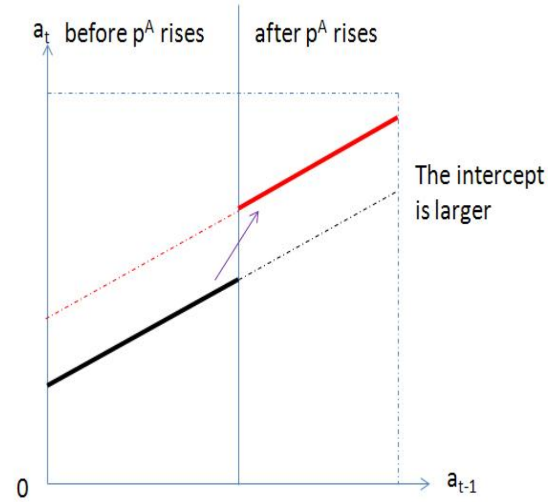


Figure 3.8. Impacts from adoption scenario 4: $p^A(L) < p^A_1 < p^A_2$

In Figures 3.5 to 3.8, technological progress represented by equation (3.8) is plotted. In each figure, the whole space is partitioned into left and right parts. On the left of graph is the technological progress before p^A rises. On the right is the technological progress after p^A rises. Only the solid lines are effective and the dashed lines (extending solid lines) are drawn to emphasize the two solid lines are parallel (slope does not change after p^A rises). Scenario 1 is plotted in Figure 3.5, where if aging is not severe, namely $p_1^A < p_2^A < p^A(L)$, it has no impacts upon technological progress. In scenarios 3 and 4, where aging is severe enough namely $p_1^A < p^A(L) < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A} < p_2^A$ or $p^A(L) < p_1^A < p_2^A$, technological progress after aging evolves on a path with a higher intercept (with the same slope) and this means an increase in the technological level. Therefore, if aging is severe enough, it has positive impacts upon technological progress.The intuition is: since becoming old will potentially raise agents' adoption abilities, a very severe individual aging will raise average adoption ability in the economy enough so as to yield positive impacts upon technological progress.

Following the above intuition, since aging never decreases adoption skills, it might seem that a more severe individual aging will never have negative impacts upon technological progress. (Under scenarios 1, 3 and 4, aging has either no impacts or positive impacts on technological progress.) However, under scenario 2, which is plotted in Figure 3.6, the technological progress experiences a decrease in this level after aging becomes more severe. When $p_1^A < p^A(L) < p_2^A < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, a more severe individual aging does have negative impacts upon technological progress. To understand this, recall that Lemma 2 shows that old agents can fully finance the project sizes, and this serves as a shield protecting them from being dismissed. When OLA are retained, it could be due to their project-financing ability, not necessarily higher adoption ability and if OLA are retained only for this reason, the average adoption ability in the economy is lower than if OLA are replaced by young agents.[20] If $p_1^A < p^A(L) < p_2^A < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, OLA do not have a high probability to have high adoption abilities, and they are retained only due to their project-financing ability. Therefore, in this case, a more severe individual aging has negative impacts upon technological progress.

3.5.2 Aging's impacts upon technological progress via innovation

In this model, when agents become old, those will low innovation ability when young will remain low innovation ability. Those with high innovation ability when young have probability p^I to become low innovation ability. An increase in p^I can be thought of as an increase in the severity of aging, in terms of lowering innovation ability.

The analysis is conducted in two steps. First, in any particular country, if individual aging becomes more severe in terms of lowering innovation ability, how is technological progress affected. Next, whether a particular individual aging, represented by a fixed p^I, has different impacts in countries with different distances to the frontier.

[20]If an OLA is retained, there is probability p^A that he will have high adoption ability when old. If a young agent is hired, with probability λ^A will he has high adoption ability. Hence if $p^A < \lambda^A$, then retaining an OLA will result in lower adoption ability on average.

In any particular country, individual aging becomes more severe.

Suppose p^I increases from p_1^I to p_2^I with $p_1^I < p_2^I$. Changes in p^I will affect the slope of (3.8). In contrast to subsection 3.5.1, the impacts of aging upon the slope depend not only on changes of p^I, but also on a, the country's distance to frontier. Here the expressions of $a(L, I)$ and $a(H, I)$ from Proposition 2 are used.

If $a < a(L, I)$ (poor countries), then from Proposition 2, OHI and OLI are both retained, regardless of changes in p^I, and slope **i** (equation (3.11)) applies. Slope **i** is decreasing in p^I, hence after a more severe aging, the slope of equation (3.8) decreases. A more severe individual aging slows down the rate of technological progress.

If $a(L, I) < a < a(H, I)$ (intermediate countries), then from Proposition 2, OLI are dismissed and OHI are retained, regardless of changes in p^I, and slope **ii** (equation (3.12)) applies. Slope **ii** is decreasing in p^I, hence after a more severe aging, the slope of equation (3.8) decreases. A more severe individual aging causes a slow down in technological progress.

If $a > a(H, I)$ (advanced countries), then from Proposition 2, OLI are dismissed regardless of changes in p^I, but the retaining rules of OHI depend on p^I and there are three possible situations:

1. If $p_1^I < p_2^I < 1 - \sigma^I \lambda^I$, OHI are retained both before and after aging becomes more severe and slope **ii** applies. Slope **ii** is decreasing in p^I, hence after a more severe aging, the slope of equation (3.8) decreases. A more severe individual aging slows down the rate of technological progress.

2. If $p_1^I < 1 - \sigma^I \lambda^I < p_2$, OHI are retained before aging becomes severe but dismissed after aging becomes more severe and slope **ii** applies before aging becomes severe but slope **iii** (equation (3.13)) applies after aging becomes more severe. Slope **iii**> slope **ii** when $p > 1 - \sigma^I \lambda^I$, so in this case the slope is larger after aging becomes more severe. A more severe individual aging increases the rate of technological progress.

3. If $1 - \sigma^I \lambda^I < p_1^I < p_2^I$, OHI are dismissed both before and after aging becomes more severe and slope **iii** applies. Slope **iii** is not affected by changes in p^I, so a more severe individual aging does not affect technological progress.

A particular individual aging happens to countries with different a.

The analysis here assumes the severity of individual aging is fixed at p^I, but the same severity of aging happens to countries with different a and studies if the impacts of aging upon slope of (3.8) vary across countries. The expressions of $a(L, I)$ and $a(H, I)$ from Proposition 2 are used here.

If $p^I < 1 - \sigma^I \lambda^I$,

1. In countries with $a < a(L, I)$, OLI and OHI are both retained, and slope **i** applies;

2. In countries with $a(L, I) < a < a(H, I)$, OLI are dismissed while OHI are retained, and slope **ii** applies;

3. In countries with $a > a(H, I)$, OLI are dismissed while OHI are retained, and slope **ii** applies.

It is simple to check that slope **i**<slope **ii**. Therefore, if $p^I < 1-\sigma^I \lambda^I$, the slope of technological progress is slower in countries with $a < a(L, I)$, than those with $a > a(L, I)$, but no difference among countries with $a(L, I) < a < a(H, I)$ and those with $a > a(H, I)$.

If $p^I > 1 - \sigma^I \lambda^I$,

1. In countries with $a < a(L, I)$, OLI and OHI are both retained, and slope **i** applies;

2. In countries with $a(L, I) < a < a(H, I)$, OLI are dismissed while OHI are retained, and slope **ii** applies;

3. In countries with $a > a(H, I)$, OLI and OHI are both dismissed, and slope **iii** applies.

It is simple to check that slope **i**<slope **ii**. Therefore, if $p^I > 1-\sigma^I \lambda^I$, the slope of technological progress is smaller in countries with $a < a(L, I)$, than those with $a(L, I) < a < a(H, I)$. Moreover, slope **ii**<slope **iii** if $p^I > 1 - \sigma^I \lambda^I$, hence the slope of technological progress is smaller in countries with $a(L, I) < a < a(H, I)$, than those with $a > a(H, I)$.

The findings in subsection 3.5.2 are summarized in the two propositions below:

Proposition 4 *Assuming that aging does not affect adoption ability (p^A is fixed), individual aging does not affect the intercept of equation (3.8), hence no impacts upon the level of technology. Given a particular country, suppose individual aging becomes more severe in terms of lowering innovation ability, namely p^I increases from p^I_1 to p^I_2 with $p^I_1 < p^I_2$, then*

In poor or intermediate countries, namely $a < a(H, I)$**,** *A more severe individual aging decreases the slope in (3.8) hence slows down technological progress;*

In advanced countries, namely $a > a(H, I)$**,** *there are three possibilities:*

1. *If* $p_1^I < p_2^I < 1 - \sigma^I \lambda^I$*, a more severe individual aging decreases the slope in (3.8) hence slows down technological progress;*
2. *If* $p_1^I < 1 - \sigma^I \lambda^I < p_2^I$*, a more severe individual aging increases the slope in (3.8) hence speeds up technological progress;*
3. *If* $1 - \sigma^I \lambda^I < p_1^I < p_2^I$*, a more severe individual aging does not affect the slope in (3.8) hence has no impacts upon technological progress.*

Proposition 5 *Assuming that aging does not affect adoption ability (*p^A *is fixed), and* p^I *is the same for all countries. The countries differ in their distances to the world technology frontier. Aging's impacts upon technological progress in countries with different distances to frontier are*

1. *If* $p^I < 1 - \sigma^I \lambda^I$*, the slope of (3.8) is smaller hence the speed of technological progress is slower in countries with* $a < a(L, I)$*, than those with* $a > a(L, I)$*, but no difference among those with* $a(L, I) < a < a(H, I)$ *and those with* $a > a(H, I)$*;*
2. *If* $p^I > 1 - \sigma^I \lambda^I$*, the slope of (3.8) is smaller hence the speed of technological progress is slower in countries with* $a < a(L, I)$*, than those with* $a(L, I) < a < a(H, I)$*, than those with* $a > a(H, I)$*.*

Equation (3.8) is plotted in Figure 3.9 below to illustrate Proposition 4 (In any particular country, aging becomes more severe).

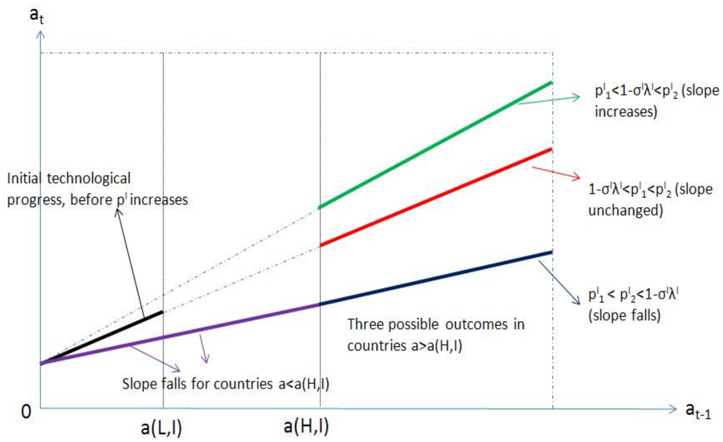


Figure 3.9. The impacts of an increase in p^I in different countries

The black bold solid line denotes the technological progress of all countries before p^I rises. All countries are assumed to have the same technological progress (same intercept and same slope) before p^I, just for illustration simplicity. All other solid lines denote the technological progress after p^I rises. The purple solid line is for countries with $a, a(H < I)$ (poor and intermediate countries), while the other three lines (green, red and blue) are for countries with $a > a(H, I)$. The blue line is for the case when $p_1^I < p_2^I < 1 - \sigma^I \lambda^I$, while the red is for the case $1 - \sigma^I \lambda^I < p_1^I < p_2^I$ and the green is for the case $p_1^I < 1 - \sigma^I \lambda^I < p_2^I$. As shown in the graph, after p^I rises, the speed of technological progress (slope) decreases in countries with $a, a(H, I)$, and countries with $a > a(H, I)$ if $p_1^I < p_2^I < 1 - \sigma^I \lambda^I$. The slope is unchanged in countries with $a > a(H, I)$ if $1 - \sigma^I \lambda^I < p_1^I < p_2^I$. The slope is larger in countries with $a.a(H, I)$ if $p_1^I < 1 - \sigma^I \lambda^I < p_2^I$.

In Figure 3.9, the purple line and the blue line has the same slope. This is for drawing convenience. After p^I, the slope for countries with $a < a(H, I)$ (purple line) and the slope for countries with $a > a(H, I)$ (blue line) do not necessarily equal, but this is an irrelevant point. The key point is that these two slopes are both less than the slope before p^I rises (the black line).

Proposition 4 implies that in any particular country, a mild change in the severity of aging ($p_1^I < p_2^I < 1 - \sigma^I \lambda^I$) will slow down technological progress, while a big change ($p_1^I < 1 - \sigma^I \lambda^I < p_2^I$) will speed up technological progress. Intuitively, individual aging potentially lowers innovation ability, hence dismissing OHI will result in an economy with higher average innovation ability.[21] However, OHI can finance projects so this protect them from being

[21] Average innovation ability of young agents is $\lambda^I \gamma^I$ while that of OHI is $(1 - p^I)\gamma^I$. If $p > 1 - \lambda^I$ young

dismissed. Only if individual aging becomes very severe, will a country adopt the retaining rule that is better in terms of technological progress. Therefore, a big change of aging severity (p^I changes from below to above $1-\sigma^I\lambda^I$) will speed up technological progress than a mild change of aging severity (p^I does not change from below to above $1-\sigma^I\lambda^I$). In the case that $1-\sigma^I\lambda^I < p_1^I < p_2^I$, the retaining rule is already optimal even before aging becomes more severe, hence a more severe aging will no more affect technological progress.

In Figures 3.10 and 11 below, equation (3.8) is plotted to illustrate Proposition 5 (a particular individual aging shock happens to different countries).

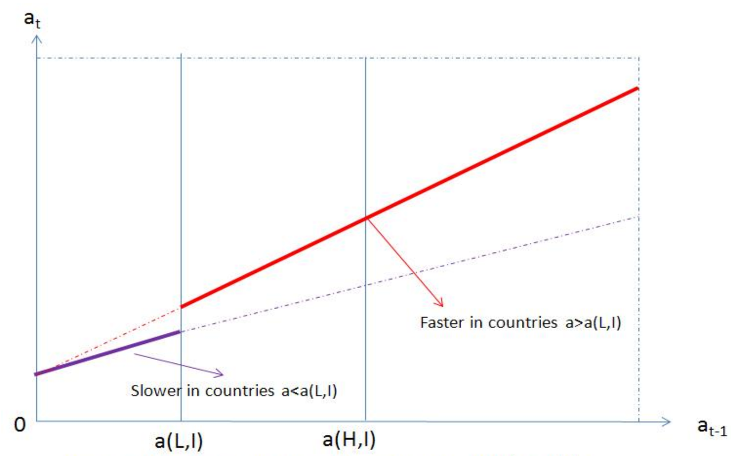


Figure 3.10. Impacts of a mild aging shock ($p^I<1-\sigma^I\lambda^I$) upon different countries

agents have higher innovation ability on average. However, $p^I > 1-\sigma^I\lambda^I > 1-\lambda^I$ is needed to have a higher slope, which means $p > 1-\lambda^I$, namely having agents with higher innovation ability on average, is not sufficient. This is because young agents run small projects, so they must have even higher innovation ability to compensate for this, which requires p^I to be higher than $1-\lambda^I$. If $\sigma^I = 1$, namely there are no differences in project sizes, this is not a problem.

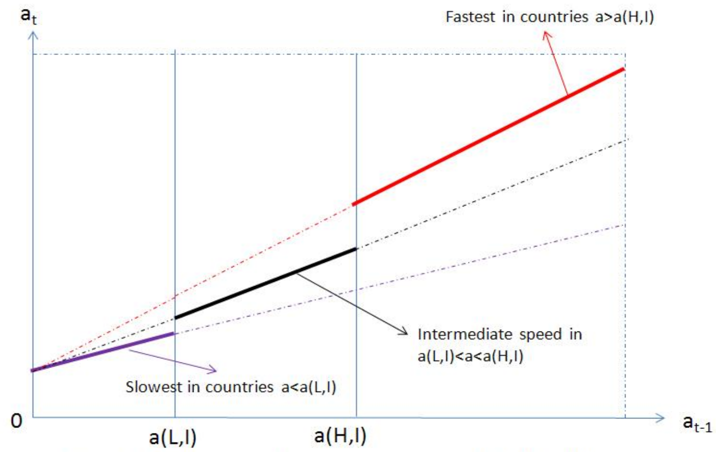


Figure 3.11. Impacts of a severe aging shock ($p^I > 1-\sigma^I\lambda^I$) upon different countries

In Figures 3.10 and 3.11, all countries are assumed to have the same intercept and slope before aging shock and that common is omitted for illustration simplicity. Only the technological progress after the aging shock is drawn and the graphs show that different countries have different slope (speed of technological progress) after the aging shock.

According to Figure 3.10, a mild aging shock ($p^I < 1 - \sigma^I \lambda^I$) will speed up technological progress of countries with $a > a(L, I)$ (intermediate and advanced countries). Figure 3.11 shows that a severe aging shock ($p^I > 1 - \sigma^I \lambda^I$) will also do so, and it even benefits advanced countries ($a > a(H, I)$) more than intermediate countries ($a(L, I) < a < a(H, I)$). Intuitively, individual aging may lower innovation ability, and so a country should dismiss old innovators in order to have faster technological progress. Advanced countries value innovation the most, and poor countries value innovation the least, so more advanced countries will adopt better retaining rules, and this is why the same individual aging shock will benefit advanced countries while hurt poor countries, and this causes a divergence among advanced and poor countries.

3.5.3 Overall impacts of individual aging

Individual aging might increase agents' adoption ability (better adoption) while decrease innovation ability (worse innovation), and so the net impacts of individual aging could be ambiguous in general. The net impacts of individual aging can be formally summarized by listing all possible combinations of all scenarios in Propositions 3, 4 and 5. However this is very tedious and not informative. Below one possible scenario is mentioned. This scenario is interesting in showing the tradeoff between faster-adoption and slower-innovation and pointing out the case when lagging countries would catch up with advanced countries because of individual aging.

According to Propositions 4 and 5, if individual aging decreases innovation ability, advanced countries ($a > a(H, L)$) tend to adopt better retaining rules of old innovators (namely dismissing old low skilled innovators and hiring young innovators), than poor countries ($a < a(L, I)$). This will speed up the technological progress of advanced countries. Individual aging, besides lowering innovation ability, also increases adoption ability, and since adoption is relatively more important than innovation for poor countries, poor countries might benefit enough from better adoption so as to offset any negative impacts from lowered innovation ability. In fact, the benefit from better adoption can be so large that individual aging could result in the initial poor country ($a < a(L, I)$) catching up with the initial advanced country ($a > a(H, I)$). This catching-up is more likely to happen if the severity of individual aging is different across countries, as illustrated by the hypothetical example below.

Consider a poor country (country 1) that is very far away from the world frontier, in particular, $a < a(L, I)$. Suppose poor country 1 is hit by an individual aging shock severe in both adoption and innovation. In particular, suppose $p^A > \sigma^A \lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$ and $p^I > 1 - \sigma^I \lambda^I$ in country 1. At the same time, suppose there is an advanced country 2, with $a > a(H, I)$, and is hit by a relatively mild individual aging. In particular, suppose $p^A < p^A(L)$ and $p^I < 1 - \sigma^I \lambda^I$ in country 2. According to Proposition 3, country 1's technological level will experience an increase from its agents' higher adoption ability while country 2 could experience a decrease (scenario 2 of Proposition 3). If country 1's technological level increases enough so as to exceed $a(H, I)$, then from Proposition 4, the speed of technological progress is faster in country 1 than country 2. In other words, the initial poor country 1 now has a higher level of technology and its technological progress is faster than the initial advanced country 2. Country 1 catches up with country 2, due to individual aging.[22]

Individual aging lowers fluid intelligence which in turn slows down technology innovation. On the other hand, aging raises crystallized intelligence which in turn fasters technology adoption. Due to this reason, Hunt (1995, page 18) argues that"(individual) aging increases the value of a workforce when the workplace is static, but it may decrease the value of

[22]In this model there is one unique final good (the consumption good) which is set as the numéraire with price set to one, hence a higher output of the final output is translated into a higher welfare. A higher level of technology corresponds to higher welfare. This can be seen by putting $x_t(v) = A_t(v) N \chi^{-\frac{1}{1-\alpha}}$ into equation (3.1) which yields $y_t = \frac{1}{\alpha} N \chi^{-\frac{1}{1-\alpha}} A_t$. Hence a higher technological progress means higher welfare.

the same workforce if the methods and technology of the workplace are changing". Since most modern developed countries are experiencing fast technological progress via innovation while technological progress in developing countries is rather slow, Hunt's argument would imply aging slows down technological progress of developed countries and do the opposite to developing countries, and this should cause a convergence in economic growth.

Hunt's argument is supported by the hypothetical example above, but not in general. Since aging's impacts on innovation ability is independent of its on adoption ability, there could be a case where aging has a very small probability of raising adoption ability (say $p^A < p^A(L)$) but a high probability of lowering innovation ability (say $p^I > 1 - \sigma^I \lambda^I$). In this case, Proposition 3 applies and the technological level will diverge between advanced and poor countries even more after aging. In fact, this divergence among countries is observed in economic literature, especially in the last 50 years (Acemoglu, 2009). In fact, this is more possible and realistic than the hypothetical example above. The example (which gives convergence rather than divergence) relies on the assumption that aging is more severe in poor countries compared to rich countries. However, in reality, due to more advanced medication system, people tend to live longer in advanced countries, so aging is more severe in advanced countries and this implies technological divergence should be expected, consistent with the empirical observation.

3.6 Numerical examples

This section studies the model numerically, in order to check the validity and robustness of the analytical results. This section first gives a detailed description of the model and then numerical analysis is conducted.

3.6.1 Model description

Functions

There are two functions in the model and they are:

1. The production function for the final good (equation (3.1)):

$$y_t = \frac{1}{\alpha} N^{1-\alpha} \int_0^1 A_t(v)^{1-\alpha} x_t(v)^{\alpha} dv \,, \qquad \alpha \in (0,1),$$

   where

- Parameters are: α and N,
- Endogenous variables are: y_t, $x_t(v)$ and $A_t(v)$.

2. Technology evolution (equation (3.6))

$$A_t(v) = s_t^A(v)\left[\theta^A + \gamma_t^A(v)\right]\bar{A}_{t-1} + \left[\theta^I + s_t^I(v)\gamma_t^I(v)\right]A_{t-1},$$

where

- Parameters are: θ^A and θ^I,
- Exogenous variables are: $\gamma_t^A(v)$, $\bar{A}_{t-1}$, and $\gamma_t^I(v)$,
- Endogenous variables are: $s_t^A(v)$, $s_t^I(v)$, and A_{t-1}.

Parameters

Below are the parameters in the model:

1. N: size of manual labor each period,
2. $\alpha \in (0,1)$: the power coefficient in the final production function,
3. $\lambda^A \in (0,1)$: the probability that a newborn young agent has high adoption ability;
4. $\lambda^I \in (0,1)$: the probability that a newborn young agent has high innovation ability;
5. χ: the limit price charged by all intermediate firms during each period,
6. $\delta \equiv (\chi - 1)\chi^{-\frac{1}{1-\alpha}}$: just a notation,
7. g: the increasing rate of the world technology frontier,
8. $\bar{A}_0$: the initial level for the world technology frontier,
9. $\sigma^A \in (0,1)$: the value for a small adoption project size,
10. $\sigma^I \in (0,1)$: the value for a small adoption project size
11. θ^A: constant parameter in equation (3.6),
12. θ^I: constant parameter in equation (3.6),
13. $\phi^A \in (0,1)$: constant parameter in the adoption project investment cost,

14. $\phi^I \in (0,1)$: constant parameter in the innovation project investment cost,

15. κ^A: constant parameter in the adoption project investment cost,

16. κ^I: constant parameter in the innovation project investment cost,

17. μ^A: the fraction of profit that can be taken by the adopter,

18. μ^I: the fraction of profit that can be taken by the innovator,

19. $p^A \in [0,1]$: the probability that low adoption ability will become high when old,

20. $p^I \in [0,1]$: the probability that high innovation ability will become low when old.

Exogenous variables

Below are exogenous variables in the model:

1. $\bar{A}_t$: world technology frontier level during time t,

2. $\gamma^A(v)$: ability of adopter, $\gamma^A(v) = \gamma^A > 0$ if high or 0 if low adoption ability,

3. $\gamma^I(v)$: ability of innovator, $\gamma^I(v) = \gamma^I > 0$ if high or 0 if low innovation ability.

Endogenous variables

Below are endogenous variables in the model:

1. y_t: the final output level during time t,

2. $A_t(v)$: the technology level of sector v during time t,

3. $x_t(v)$: the amount of intermediate good v used in final production during time t,

4. $p_t(v)$: the price of intermediate good v during t,

5. $\pi_t(v)$: the profit of intermediate firm v during time t;

6. w_t: the wage rate for manual labor during time t,

7. A_t: the average technology level during time t,

8. a_t: the distance to the world technology frontier during time t,

9. $s_t^A(v)$: the project size in the adoption sector of intermediate firm v,

10. $s_t^I(v)$: the project size in the innovation sector of intermediate firm v,

11. $k_t^A(s)$: the adoption investment cost for project with size s,

12. $k_t^I(s)$: the innovation investment cost for project with size s,

13. S_t^A: the payment to the adopter during time t,

14. S_t^I: the payment to the innovator during time t,

15. RE_t^A: total retained earnings of an old adopter accumulated from previous period,

16. RE_t^I: total retained earnings of an old innovator accumulated from previous period,

17. $\widehat{RE}_t^A$: the contribution of an old adopter to finance the adoption project,

18. $\widehat{RE}_t^I$: the contribution of an old innovator to finance the innovation project,

19. intercept **I**, expressed by equation (3.9),

20. intercept **II**, expressed by equation (3.10),

21. slope **i**, expressed by equation (3.11),

22. slope **ii**, expressed by equation (3.12),

23. slope **iii**, expressed by equation (3.13).

The above gives full description of the model. Numerical analysis is conducted below to investigate whether the analytical results are correct.

3.6.2 Numerical implications

The analytical results are formalized in Propositions 3, 4 and 5. Equations (3.8) to (3.13) fully capture the technology evolution and they are the main focus in this section.

Since the intercept and slope of equation (3.8) fully determine the technological progress of a country, they are studied to investigate the impacts of aging in Propositions 3, 4 and 5 (a higher intercept means higher technological level while a higher slope means faster technological progress speed). Analytical results are already obtained in Propositions 3, 4 and 5, to numerically test the analytical results, different values are given to the exogenous parameters, and intercepts (3.9) to (3.10) and slopes (3.11) to (3.13) are compared.

In equations (3.8) to (3.13), the parameters and endogenous variables are

16 parameters λ^A, λ^I, σ^A, σ^I, θ^A, θ^I, γ^A, γ^I, ϕ^A, κ^A, μ^A, μ^I, δ, N, p^A and p^I;

5 endogenous variables intercept **I**, intercept **II**, slope **i**, slope **ii** and slope **iii**.

The numbers of parameters and endogenous variables are less than the total numbers of parameters and endogenous variables in 3.6.1. This is because some of the parameters and endogenous variables in 3.6.1 are useful in the middle steps in solving the model and they are not directly related to the main results in Propositions 3,4 and 5. Therefore, they are not listed here.

In appendix 3.8.7, the numbers for parameters and endogenous variables are listed in three tables. In table 1, six different sets of data are chosen for the exogenous parameters. In table 2, the corresponding values for the three expressions in Proposition 3 are calculated. They do not directly matter for Propositions 4 and 5, but they are useful in choosing the values for p^A and p^I. In table 3, the values for p^A and p^I are chosen and the corresponding values for intercept **I**, intercept **II**, slope **i**, slope **ii**, and slope **iii** are calculated.

In getting the results in Propositions 3, 4 and 5, intercepts and slopes need to be compared in different situations. Full details are provided only for using the data in case 1, with less details for cases 2 and 3. Similar conclusions hold for cases 4 to 6.

The analysis below first tests analytical results in Proposition 3, which studies the impacts from adoption abilities change. Then focus is given to Propositions 4 and 5, which study the impacts from innovation abilities change.

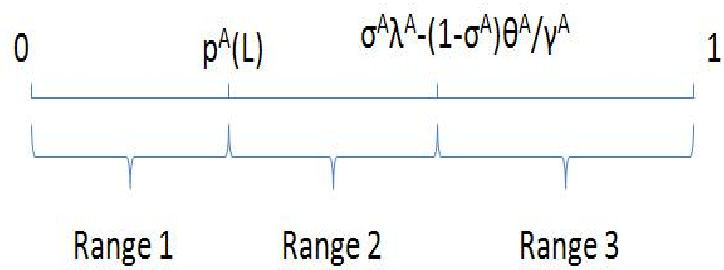
3.6.3 Numerical simulations

Numerical simulations for Proposition 3

Proposition 3 gives the impacts of individual aging from the adoption side. With individual aging, agents' adoption ability might rise and this could benefit technology adoption in a country. This is captured by equation (3.8). The intercept in equation (3.8) shows the how changes in technology adoption affect technological progress. Therefore, to test Proposition 3, numerical simulations should be done on the intercept in equation (3.8), namely investigating how the intercept in equation (3.8) changes with individual aging.

In Proposition 3, it is assumed that p^I stays unchanged and p^A increases. Denote the old and new values of p^A by p_1^A and p_2^A and suppose p^A rises from p_1^A to p_2^A ($p_1^A < p_2^A$). There are four scenarios in Proposition 3, depending on the retaining rules for old adopters. The retaining rules for old adopters are in Proposition 1, and the retaining rules depend on the relation between p^A and $p^A(L)$. Moreover, scenarios 2 and 3 in Proposition 3 depend on the relation between p^A and $\sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$. In order to test Proposition 3, for each choice of other parameters, different values of p^A are chosen to study how the intercept changes when p^A increases. Values of p^A are chosen in the three ranges drawn below.

In each case, one value for p^A is chosen in Range 1, one value in Range 2, and two values in Range 3.



In case 1, $p^A(L) = 0.01776$ and $\sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A} = 0.02$. Four possible values are chosen for p^A: 0.01, 0.0185, 0.025, 0.03. These four values for p^A, together with the parameters in case 1 are used to test the four scenarios in Proposition 3, as below.

According to scenario 1 of Proposition 3, if $p_1^A < p_2^A < p^A(L)$, a more severe aging has no impacts upon the intercept of (3.8) and no impacts on the level of technology. Testing of scenario 1 does not need any calculation because after p^A rises, the same intercept **I** applies, hence no change in the intercept, and no impacts upon technological progress.

According to scenario 2 of Proposition 3, if $p_1^A < p^A(L) < p_2^A < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$, a more severe aging causes a decrease in the intercept of (3.8) and a decrease in the level of technology. To test scenario 2 of proposition 3, p^A increases from 0.01 (in Range 1) to 0.0185 (Range 2). The corresponding intercept changes from Intercept **I**=1.93 to Intercept **II**=1.92925<1.93. The intercept decreases after aging becomes more severe and this is consistent with scenario 2 in Proposition 3.

According to scenario 3 of Proposition 3, if $p_1^A < p^A(L) < \sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A} < p_2^A$, a more severe aging causes an increase in the intercept of (3.8) and an increase in the level of technology. To test scenario 3 of proposition 3, p^A increases from 0.01 (in Range 1) to 0.0185 (Range 2). The corresponding intercept changes from Intercept **I**=1.93 to Intercept **II**=1.92925<1.93. The intercept decreases after aging becomes more severe and this is consistent with scenario 3 in Proposition 3.

According to scenario 4 of Proposition 3, if $p^A(L) < p_1^A < p_2^A$, a more severe aging causes an increase in the intercept of (3.8) and an increase in the level of technology. To test scenario 4 of proposition 3, p^A increases from 0.025 (in Range 3) to 0.03 (Range 3). The corresponding intercept changes from Intercept **II**=1.9325 to Intercept **III**=1.935>1.9325. The intercept increases after aging becomes more severe and this is consistent with scenario 4 in Proposition 3.

In sum, analytical results in Proposition 3 are supported by the data in case 1.

In the data in case 2, the following values for p^A are chosen: 0.2, 0.4, 0.5 and 0.6. It will be shown that the four scenarios in Proposition 3 are supported by the data in case.

1. Testing scenario 1: no calculations are needed, since the same intercept applies after aging becomes more severe.

2. Testing scenario 2: p^A increases from 0.2 to 0.4, and the corresponding intercepts change from intercept **I** = 3.9772 to intercept **II** = 3.9510, hence a decrease in the intercept, consistent with Proposition 3.

3. Testing scenario 3: p^A increases from 0.4 to 0.5, and the corresponding intercepts change from intercept **I** = 3.9772 to intercept **II** = 3.9910, hence an increase in the intercept, consistent with Proposition 3.

4. Testing scenario 4: p^A increases from 0.5 to 0.6, and the corresponding intercepts change from intercept **II** = 3.9910 to intercept **II** = 4.0310, hence an increase in the intercept, consistent with Proposition 3.

In sum, analytical results in Proposition 3 are supported by the data in case 2.

In the data in case 3, the following values for p^A are chosen: 0.7, 0.88, 0.91 and 0.93. The analysis below shows the four scenarios in Proposition 3 are supported by the data.

1. Testing scenario 1: no calculations are needed, since the same intercept applies after aging becomes more severe.

2. Testing scenario 2: p^A increases from 0.7 to 0.88, and the corresponding intercepts change from intercept I = 3.67215 to intercept II = 3.67087, hence a decrease in the intercept, consistent with Proposition 3.

3. Testing scenario 3: p^A increases from 0.88 to 0.91, and the corresponding intercepts change from intercept I = 3.67215 to intercept II = 3.67546, hence an increase in the intercept, consistent with Proposition 3.

4. Testing scenario 4: p^A increases from 0.91 to 0.93, and the corresponding intercepts change from intercept II = 3.67546 to intercept II = 3.67852, hence an increase in the intercept, consistent with Proposition 3.

Using similar logic, it can checked that Proposition 3 is supported by the data cases 4 to 6.

In sum, Proposition 3 is supported by the data in all six cases. The analysis below conducts numerical simulations for Propositions 4 and 5.

Numerical simulations for Proposition 4

Proposition 4 gives the impacts of individual aging from the innovation side. If population aging becomes more severe so that agents' innovation ability decreases, this would affect technology innovation. This is captured by equation (3.8). The slope in equation (3.8) shows the how changes in technology innovation affect technological progress. Therefore, to test Proposition 4, numerical simulations should be done on the slope in equation (3.8), namely investigating how the slope in equation (3.8) changes with individual aging.

In Proposition 4, it is assumed that p^A stays unchanged and p^I increases. The impacts of a higher p^I upon technological progress depend on the country's distance to the world technology frontier. To test this, denote the old and new values of p^I and p_1^I to p_2^I suppose p^I rises from p_1^I to p_2^I ($p_1^I < p_2^I$).

According to Proposition 4, in poor and intermediate countries, a higher p^I would decrease the slope of (3.8) so individual aging slows down technological progress. In advanced countries, the impacts of individual aging depends on the relation between p_1^I, p_2^I and $1 - \sigma^I \lambda^I$. To test Proposition 4, for each choice of data set, $1 - \sigma^I \lambda^I$ needs to be computed and values of p^I are chosen around $1 - \sigma^I \lambda^I$.

In the data in case 1, $1-\sigma^I\lambda^I=0.7$, and four values of p^I are chosen: 0.5, 0.6, 0.8, 0.85 (two less than 0.7 and two above).

In countries with $a<a(L,I)$, slope **i** applies both before and after p^I increases. In the data, when p^I goes up from 0.5 to 0.85, the corresponding slope **i** decreases from 1.95, to 1.90, to 1.8, to 1.775. Therefore, the speed of technological progress slows down after aging becomes more severe. This is consistent with Proposition 4.

In countries with $a(L,I)<a<a(H,I)$, slope **ii** applies both before and after p^I increases. In the data, when p^I goes up from 0.5 to 0.85, the corresponding slope **ii** decreases from 2.1, to 2.05, to 1.95, to 1.925. Therefore, the speed of technological progress slows down after aging becomes more severe. This is consistent with Proposition 4.

In countries with $a>a(H,I)$, there are three scenarios to check. When p^I increases from 0.5 to 0.6 (both less than $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **ii**=2.1 to slope **ii**=2.05, hence the slope is smaller, consistent with Proposition 4. When p^I increases from 0.6 to 0.8 (from less than to above $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **ii**=1.95 to slope **iii**=2, hence the slope is larger, consistent with Proposition 4. When p^I increases from 0.8 to 0.85 (both larger than $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **iii**=2 to slope **iii**=2, hence no change in the slope, consistent with Proposition 4.

In sum, the findings in Proposition 4 are supported by the data in case 1.

In the data in case 2, $1-\sigma^I\lambda^I=0.415$, and four values are chosen for p^I: 0.2, 0.3, 0.5, 0.6.

In countries with $a<a(L,I)$, slope **i** applies both before and after p^I increases. In the data, when p^I goes up from 0.2 to 0.6, the corresponding slope **i** decreases from 5.515, to 5.245, to 4.705, to 4.435. Therefore, the speed of technological progress slows down after aging becomes more severe. This is consistent with Proposition 4.

In countries with $a(L,I)<a<a(H,I)$, slope **ii** applies both before and after p^I increases. In the data, when p^I goes up from 0.2 to 0.6, the corresponding slope **ii** decreases from 1.5976, to 1.5829, to 1.5682, to 1.56085. Therefore, the speed of technological progress slows down after aging becomes more severe. This is consistent with Proposition 4.

In countries with $a>a(H,I)$, there are three scenarios to check. When p^I increases from 0.2 to 0.3 (both less than $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **ii**=5.6905 to slope **ii**=5.4205, hence the slope is smaller, consistent with Proposition 4. When p^I increases from 0.3 to 0.5 (from less than to above $1-\sigma^I\lambda^I$), the corresponding slope changes from slope

ii=4.8805 to slope **iii**=5.11, hence the slope is larger, consistent with Proposition 4. When p^I increases from 0.5 to 0.6 (both larger than $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **iii**=5.11 to slope **iii**=5.11, hence no change in the slope, consistent with Proposition 4.

In sum, the findings in Proposition 4 are supported by the data in case 2.

In the data in case 3, $1-\sigma^I\lambda^I = 0.85$, and four values are chosen for p^I: 0.7, 0.8, 0.9, 0.95.

In countries with $a < a(L,I)$, slope **i** applies both before and after p^I increases. When p^I goes up from 0.7 to 0.95, the corresponding slope **i** decreases from 1.5976, to 1.5829, to 1.5682, to 1.56085. Therefore, the speed of technological progress slows down after aging becomes more severe. This is consistent with Proposition 4.

In countries with $a(L,I) < a < a(H,I)$, slope **ii** applies both before and after p^I increases. When p^I goes up from 0.7 to 0.95, the corresponding slope **ii** decreases from 1.64905, to 1.63436, to 1.61965, to 1.6123. Therefore, the speed of technological progress slows down after aging becomes more severe. This is consistent with Proposition 4.

In countries with $a > a(H,I)$, there are three scenarios to check. When p^I increases from 0.7 to 0.8 (both less than $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **ii**=1.64905 to slope **ii**=1.63435, hence the slope is smaller, consistent with Proposition 4. When p^I increases from 0.8 to 0.9 (from less than to above $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **ii**=1.61965 to slope **iii**=1.627, hence the slope is larger, consistent with Proposition 4. When p^I increases from 0.9 to 0.95 (both larger than $1-\sigma^I\lambda^I$), the corresponding slope changes from slope **iii**=1.627 to slope **iii**=1.627, hence no change in the slope, consistent with Proposition 4.

In sum, the findings in Proposition 4 are supported by the data in case 3.

Using similar methods, it can be easily checked that Proposition 4 is supported by the data in cases 4 to 6.

In sum, the findings in Proposition 4 are supported by all six data sets. The analysis below gives the numerical simulations for Proposition 5.

Numerical simulations for Proposition 5

In Proposition 5, it is assumed that p^I is the same for all countries, and aging's impacts in different countries are compared. To test Proposition 5, this subsection assumes that a given individual aging happens to different countries, and investigates if the resulting impacts upon technological progress differ across countries. According to Proposition 5, the impacts of aging depend on the relation between p^I and $1 - \sigma^I \lambda^I$. To test it, for the data in each case, $1 - \sigma^I \lambda^I$ is computed and only two values are chosen for p^I, one above and one under $1 - \sigma^I \lambda^I$.

In the data in case 1, $1 - \sigma^I \lambda^I = 0.7$ and 0.5 and 0.85 are chosen for p^I. When $p^I = 0.5$ (scenario 1 of Proposition 5), in countries with $a < a(L, I)$, slope **i**=1.95 applies. In countries with $a > a(L, I)$, slope **ii**=2.1 applies and this is larger than 1.95. Therefore the speed of technological progress is faster in countries with $a > a(L, I)$ than those with $a < a(L, I)$ and this is consistent with scenario 1 of Proposition 5. When $p^I = 0.85$ (scenario 2 of Proposition 5), in countries with $a < a(L, I)$, the slope **i**=1.775 applies. In countries with $a(L, I) < a < a(H, I)$, slope **ii**=1.925 applies. In countries with $a > a(H, I)$, slope **iii**=2 applies. Therefore the speed of technological progress is faster in countries with $a > a(H, I)$, than those with $a(L, I) < a < a(H, I)$, than those with $a < a(L, I)$ and this is consistent with scenario 2 of Proposition 5.

In the data in case 2, $1 - \sigma^I \lambda^I = 0.415$ and 0.3 and 0.5 are chosen for p^I. When $p^I = 0.3$ (scenario 1 of Proposition 5), in countries with $a < a(L, I)$, slope **i**=5.245 applies. In countries with $a > a(L, I)$, slope **ii**=5.4205 applies and this is larger than 5.245. Therefore the speed of technological progress is faster in countries with $a > a(L, I)$ than those with $a < a(L, I)$ and this is consistent with scenario 1 of Proposition 5. When $p^I = 0.5$ (scenario 2 of Proposition 5), in countries with $a < a(L, I)$, slope **i**=4.705 applies. In countries with $a(L, I) < a < a(H, I)$, slope **ii**=4.8805 applies. In countries with $a > a(H, I)$, slope **iii**=5.11 applies. Therefore the speed of technological progress is faster in countries with $a > a(H, I)$, than those with $a(L, I) < a < a(H, I)$, than those with $a < a(L, I)$ and this is consistent with scenario 2 of Proposition 5.

In the data in case 3, $1 - \sigma^I \lambda^I = 0.85$ and 0.7 and 0.9 are chosen for p^I. When $p^I = 0.7$ (scenario 1 of Proposition 5), in countries with $a < a(L, I)$, slope **i**=1.5976 applies. In countries with $a > a(L, I)$, slope **ii**=1.64905 applies and this is larger than 1.5976. Therefore the speed of technological progress is faster in countries with $a > a(L, I)$ than those with $a < a(L, I)$ and this is consistent with scenario 1 of Proposition 5. When $p^I = 0.9$ (scenario 2 of Proposition 5), in countries with $a < a(L, I)$, slope **i**=1.5682 applies. In countries with $a(L, I) < a < a(H, I)$, slope **ii**=1.61965 applies. In countries with $a > a(H, I)$, slope **iii**=1.627 applies. Therefore the speed of technological progress is faster in countries with $a > a(H, I)$, than those with $a(L, I) < a < a(H, I)$, than those with $a < a(L, I)$ and this is consistent with scenario 2 of Proposition 5.

Using similar logic, it can easily checked that the findings of Proposition 5 are supported by data in cases 4 to 6.

In sum, the findings in Proposition 5 are supported in the data in all of the six cases.

3.7 Conclusion

Individual aging changes agents' abilities. As an agent becomes old, he is better at learning current knowledge, but worse at innovating new things. This chapter examines the impacts of individual aging on technological progress, and shows how the impacts of individual aging interact with a country's pre-aging technology level.

There are several interesting and intuitive results. Since individual aging lowers agents' innovation ability, this tends to slow down technological progress. However, expecting this negative impact, countries close to the world frontier will dismiss old innovators and hire young agents. On the other hand, countries faraway from the world frontier still retain old agents. Because the advanced countries hire young agents, who have higher average innovation ability, the average innovation ability is higher in advanced countries, which makes them grow faster. Individual aging magnifies the technology differences between advanced countries and lagging countries, causing a divergence. In sum, compared to the lagging countries, advanced countries will grow relatively faster because of individual aging.

While individual aging lowers innovation ability, it raises adoption ability. Technology adoption is also an important part of technological progress, especially for countries faraway from the world frontier. After taking individual aging's adoption-ability-raising effect into account, the net impact of individual aging on technological progress is ambiguous. Therefore, individual aging could even quicken technological progress of countries falling behind the world frontier, and so is beneficial to *all* countries (but still relatively more to advanced countries).

In this model technological progress depends crucially on the ability of managers (adopters and innovators) hired by the principal, which means retaining rules of agents play an essential role. Relative incentives of hiring different agents, the retaining rules, depend on the relative importance of technology adoption and innovation, which further depend on a country's distance to world technology frontier. In the absence of countries' interactions (which constitute the world technology frontier), there is no technology adoption and innovation is the only channel left, whose consequence is the triviality of retaining rules (old low skill innovators are always fired and retaining rules about adopters become irrelevant). So even if there is no international trade, the implicit interactions among countries indirectly determine the technological progress of a country. In future research, the model could be simplified to a two-country case while having two final goods, and international trade in goods is allowed which results in changes in production patterns compared to autarky.

For tractability, this chapter assumes the probabilities of abilities change (namely, p^A and p^I) are exogenous and vary across countries. A more realistic model is to consider them endogenously determined by agents' educational level and what jobs they do when young. Evidence from psychology suggests education and job characteristics are important in determining the magnitude of individual aging's impacts on abilities. In the future research, factors affecting p^A and p^I could be explicitly put in the analysis, and focus is on optimal fiscal policy (education subsidy) and trade policy (which direct workers to certain job positions). These complicate the current model very much and they are left for further research.

3.8 Appendix

3.8.1 Participation constraints

This section solves the sufficient conditions for agents' participation constraints to hold. Each type of agents is analysed below.

Old, high skill adopters (OHA)

For OHA, the participation constraint requires his net profit from doing adoption must be larger than doing manual work in the final production. If he works as an adopter, he gets S_t^A, but he has to pay $\widehat{RE}_t^A$ to finance the adoption project. If he works as a manual labor, he gets the wage rate w_t. Participation constraint for him requires the following to hold

$$PC_t^A(O,H) \equiv S_t^A - \widehat{RE}_t^A - w_t \geq 0; \tag{PC-OHA}$$

$$S_t^A \geq \mu^A \delta N \left((\theta^A + \gamma^A)\bar{A}_{t-1} + \theta^I A_{t-1}\right);$$

$$\widehat{RE}_t^A \leq \mu^A \delta N \left(\sigma^A(\theta^A + \gamma^A)\bar{A}_{t-2} + (\theta^I + \sigma^I \gamma^I) A_{t-2}\right).$$

A sufficient condition for (PC-OHA) is

$$\mu^A\delta N\left((\theta^A+\gamma^A)\bar{A}_{t-1}+\theta^I A_{t-1}\right)-\mu^A\delta N\left(\sigma^A(\theta^A+\gamma^A)\bar{A}_{t-2}+(\theta^I+\sigma^I\gamma^I)A_{t-2}\right)$$
$$\geq\frac{1-\alpha}{\alpha}\chi^{-\frac{\alpha}{1-\alpha}}\bar{A}_t$$

$\Longleftarrow$

$$\mu^A\delta N(\theta^A+\gamma^A)\bar{A}_{t-1}-\mu^A\delta N\left(\sigma^A(\theta^A+\gamma^A)+(\theta^I+\sigma^I\gamma^I)\right)\bar{A}_{t-2}\geq\frac{1-\alpha}{\alpha}\chi^{-\frac{\alpha}{1-\alpha}}\bar{A}_t$$

$\Longleftarrow$

$$\mu^A\delta N\left[(\theta^A+\gamma^A)-\frac{1}{1+g}\left(\sigma^A(\theta^A+\gamma^A)+(\theta^I+\sigma^I\gamma^I)\right)\right]\geq(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}.$$

Assuming $(\theta^A+\gamma^A)>\frac{1}{1+g}\left(\sigma^A(\theta^A+\gamma^A)+(\theta^I+\sigma^I\gamma^I)\right)$, which requires σ^A and σ^I to be small enough, it follows that

$$N\geq\frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}}{\mu^A\delta\left((1-\frac{\sigma^A}{1+g})(\theta^A+\gamma^A)-\frac{1}{1+g}(\theta^I+\sigma^I\gamma^I)\right)}\equiv N^A_{OH}<\infty.$$

Old, low skill adopters (OLA)

A sufficient condition for the participation constraint of OLA is

$$\mu^A\delta N\left(\sigma^A\theta^A\bar{A}_{t-1}+\theta^I A_{t-1}\right)-\mu^A\delta N\left(\sigma^A\theta^A\bar{A}_{t-2}+(\theta^I+\sigma^I\gamma^I)A_{t-2}\right)$$
$$\geq\frac{1-\alpha}{\alpha}\chi^{-\frac{\alpha}{1-\alpha}}\bar{A}_t$$

$\Longleftarrow$

$$\mu^A\delta N\sigma^A\theta^A-\frac{1}{1+g}\mu^A\delta N\left(\sigma^A\theta^A+(\theta^I+\sigma^I\gamma^I)\right)\geq(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}.$$

Assuming $\sigma^A\theta^A>\frac{1}{1+g}\left(\sigma^A\theta^A+(\theta^I+\sigma^I\gamma^I)\right)$, it follows that

$$N\geq\frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}}{\mu^A\delta\left(\frac{g}{1+g}\sigma^A\theta^A-\frac{1}{1+g}(\theta^I+\sigma^I\gamma^I)\right)}\equiv N^A_{OL}<\infty.$$

Young adopters (YA)

For YA, the following holds:

$$PC^A_t(Y)\equiv S^A_t+\mathbb{E}_t Rent_{t+1}-\widehat{RE}^A_t-w_t\geq 0,\qquad\text{(PC-YA)}$$

where

$$\mathbb{E}_t Rent_{t+1} = \lambda^A R_t^A(H) PC_{t+1}^A(O,H) \\ + (1-\lambda^A) R_t^A(L) PC_{t+1}^A(O,L),$$

where

$$R_t^A(H) = \begin{cases} 1 & \text{if } OHA \text{ is retained in period } t+1, \\ 0 & \text{if } OHA \text{ is fired in period } t+1; \end{cases}$$

and

$$R_t^A(L) = \begin{cases} 1 & \text{if } OLA \text{ is retained in period } t+1, \\ 0 & \text{if } OLA \text{ is fired in period } t+1. \end{cases}$$

Since $\mathbb{E}_t Rent_{t+1} \geq 0$ and $\widehat{RE}_t^A = 0$ for a young agent (young agents are born with no wealth), a sufficient condition for (PC-YA) is

$$\mu^A \delta N(\sigma^A \theta^A \bar{A}_{t-1} + \theta^I A_{t-1}) \geq \frac{1-\alpha}{\alpha} \chi^{-\frac{\alpha}{1-\alpha}} \bar{A}_t$$

$\Longleftarrow$

$$\mu^A \delta N \sigma^A \theta^A \bar{A}_{t-1} \geq \frac{1-\alpha}{\alpha} \chi^{-\frac{\alpha}{1-\alpha}} \bar{A}_t$$

$\Longleftrightarrow$

$$N \geq \frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}}{\mu^A \delta \sigma^A \theta^A} \equiv N_Y^A < \infty.$$

Old, high skill innovators (OHI)

For OHI,

$$PC_t^I(O,H) \equiv S_t^I - \widehat{RE}_t^I - w_t \geq 0, \tag{PC-OHI}$$
$$S_t^I \geq \mu^I \delta N \left(\sigma^A \theta^A \bar{A}_{t-1} + (\theta^I + \gamma^I) A_{t-1}\right),$$
$$\widehat{RE}_t^I \leq \mu^I \delta N \left(\sigma^A (\theta^A + \gamma^A) \bar{A}_{t-2} + (\theta^I + \sigma^I \gamma^I) A_{t-2}\right),$$

$\Longleftarrow$

$$\mu^I \delta N \left[\left(\sigma^A \theta^A \bar{A}_{t-1} + (\theta^I + \gamma^I) A_{t-1}\right) - \left(\sigma^A (\theta^A + \gamma^A) \bar{A}_{t-2} + (\theta^I + \sigma^I \gamma^I) A_{t-2}\right)\right] \\ \geq \frac{1-\alpha}{\alpha} \chi^{-\frac{\alpha}{1-\alpha}} \bar{A}_t.$$

Note $(\theta^I+\gamma^I)A_{t-1} > (\theta^I+\sigma^I\gamma^I)A_{t-2}$ because $A_{t-1} \geq A_{t-2}$, $\sigma^I \in (0,1)$ and $\gamma^I > 0$. Assuming $\sigma^A\theta^A(1+g) \geq \sigma^A(\theta^A+\gamma^A)$, a sufficient condition for (PC-OHI) is

$$N \geq \frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}}{\mu^I\delta\left(\sigma^A\theta^A - \frac{1}{1+g}\sigma^A(\theta^A+\gamma^A)\right)} \equiv N^I_{OH} < \infty.$$

Old, low skill innovators (OLI)

Similarly, a sufficient condition for the participation constraint of OLI is

$$\mu^I\delta N\left[\left(\sigma^A\theta^A\bar{A}_{t-1}+\theta^I A_{t-1}\right)-\left(\sigma^A(\theta^A+\gamma^A)\bar{A}_{t-2}+\theta^I A_{t-2}\right)\right] \geq \frac{1-\alpha}{\alpha}\chi^{-\frac{\alpha}{1-\alpha}}\bar{A}_t.$$

Assuming $\sigma^A\theta^A(1+g) \geq \sigma^A(\theta^A+\gamma^A)$ (same as in getting N^I_{OH}), the following is required

$$N \geq \frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}}{\mu^I\delta\left(\sigma^A\theta^A - \frac{1}{1+g}\sigma^A(\theta^A+\gamma^A)\right)} \equiv N^I_{OL} = N^I_{OH} < \infty.$$

Young innovators (YI)

For YI, the participation constraint (PC-YI) requires (similar logic as (PC-YA))

$$\mu^I\delta N\left(\sigma^A\theta^A\bar{A}_{t-1}+\theta^I A_{t-1}\right) \geq \frac{1-\alpha}{\alpha}\chi^{-\frac{\alpha}{1-\alpha}}\bar{A}_t$$

$\Longleftarrow$

$$N \geq \frac{(1+g)(1-\alpha)\alpha^{-1}\chi^{-\alpha/(1-\alpha)}}{\mu^I\delta\sigma^A\theta^A} \equiv N^I_Y < \infty.$$

3.8.2 Project sizes

The analysis below shows how the project sizes are determined.

Notation

$\vec{s}_t(v) = \left(s^A_t(v), s^I_t(v)\right) \in \{\sigma^A, 1\} \times \{\sigma^I, 1\}$ denotes project size choices; $\vec{e} = \left(e^A, e^I\right) \in \{Y, O\}^2$ to denote if agents are old (O) or young (Y). $\vec{z} = \left(z^A, z^I\right) \in \{H, L\}^2$ denotes ability levels, where $\vec{z} = (L, H)$ means the adopter has low adoption ability and the innovator has high innovation ability.

YA

Concerning the project size for a young adopter, $s^A_t(v|Y)$, this subsection compares $\mathbb{E}_t V_t(v|\sigma^A, s^I_t(v), Y, e^I, \vec{z}$ with $\mathbb{E}_t V_t(v|1, s^I_t(v), Y, e^I, \vec{z})$, namely, the expected value of the firm of conducting different adoption project sizes, *given* whatever the choice is in the innovation sector.

$$\begin{aligned}
&\mathbb{E}_t V_t(v|\sigma^A, s_t^I(v), Y, e^I, \vec{z}) \\
&= (1-\mu^A-\mu^I)\delta N\left[\sigma^A\left(\theta^A+\lambda^A\gamma^A\right)\bar{A}_{t-1}+s_t^I(v)\left(\theta^I+\gamma_t^I(v)\right)A_{t-1}\right] \\
&-\max\left\{\phi^A\kappa^A\bar{A}_{t-1}-\widehat{RE}_t^A,0\right\}-\max\left\{k_t^I-\widehat{RE}_t^I,0\right\}.
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_t V_t(v|1, s_t^I(v), Y, e^I, \vec{z}) \\
&= (1-\mu^A-\mu^I)\delta N\left[\left(\theta^A+\lambda^A\gamma^A\right)\bar{A}_{t-1}+s_t^I(v)\left(\theta^I+\gamma_t^I(v)\right)A_{t-1}\right] \\
&-\max\left\{\kappa^A\bar{A}_{t-1}-\widehat{RE}_t^A,0\right\}-\max\left\{k_t^I-\widehat{RE}_t^I,0\right\}.
\end{aligned}$$

As shown at the end of A2, all old agents can *fully* finance their projects (either large or small), so it follows that

$$\max\left\{\phi^A\kappa^A\bar{A}_{t-1}-\widehat{RE}_t^A,0\right\}=\max\left\{\kappa^A\bar{A}_{t-1}-\widehat{RE}_t^A,0\right\}=0.$$

Young adopters run small projects ($s_t^A=\sigma^A$) *iff*

$$\mathbb{E}_t V_t(v|\sigma^A, s_t^I(v), Y, e^I, \vec{z}) \geq \mathbb{E}_t V_t(v|1, s_t^I(v), Y, e^I, \vec{z})$$

$$\Longleftrightarrow$$

$$(1-\mu^A-\mu^I)\delta N(\sigma^A-1)(\theta^A+\lambda^A\gamma^A)\bar{A}_{t-1} \geq (\phi^A-1)\kappa^A\bar{A}_{t-1}$$

$$\Longleftrightarrow$$

$$\delta \leq \frac{(1-\phi^A)\kappa^A}{(1-\mu^A-\mu^I)N(1-\sigma^A)(\theta^A+\lambda^A\gamma^A)} \equiv \delta_Y^A < \infty.$$

YI

This subsection compares $\mathbb{E}_t V_t(v|s_t^A, \sigma^I, e^A, Y, \vec{z})$ with $\mathbb{E}_t V_t(v|s_t^A, 1, e^A, Y, \vec{z})$ to determine the project size choice in the innovation sector.

$$\begin{aligned}
&\mathbb{E}_t V_t(v|s_t^A, \sigma^I, e^A, Y, \vec{z}) \\
&= (1-\mu^A-\mu^I)\delta N\left[s_t^A(v)\left(\theta^A+\gamma_t^A(v)\right)\bar{A}_{t-1}+\left(\theta^I+\sigma^I\lambda^I\gamma^I\right)A_{t-1}\right] \\
&-\max\left\{k_t^A-\widehat{RE}_t^A,0\right\}-\max\left\{\phi^I\kappa^I\bar{A}_{t-1}-\widehat{RE}_t^I,0\right\}.
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_t V_t(v|s_t^A, 1, e^A, Y, \vec{z}) \\
&= (1-\mu^A-\mu^I)\delta N\left[s_t^A(v)\left(\theta^A+\gamma_t^A(v)\right)\bar{A}_{t-1}+\left(\theta^I+\lambda^I\gamma^I\right)A_{t-1}\right] \\
&-\max\left\{k_t^A-\widehat{RE}_t^A,0\right\}-\max\left\{\kappa^I\bar{A}_{t-1}-\widehat{RE}_t^I,0\right\}.
\end{aligned}$$

Young innovators run small projects ($s_t^I = \sigma^I$) *iff*

$$\mathbb{E}_t V_t(v|s_t^A, \sigma^I, e^A, Y, \vec{z}) \geq \mathbb{E}_t V_t(v|s_t^A, 1, e^A, Y, \vec{z})$$

$$\Longleftrightarrow$$

$$\delta \leq \frac{(1-\phi^I)\kappa^I}{(1-\mu^A-\mu^I)N(1-\sigma^I)(\theta^I+\lambda^I\gamma^I)} \frac{1}{a_{t-1}}.$$

Since $a_{t-1} \in [0,1]$, a sufficient condition for the previous inequality is

$$\delta \leq \frac{(1-\phi^I)\kappa^I}{(1-\mu^A-\mu^I)N(1-\sigma^I)(\theta^I+\lambda^I\gamma^I)} \equiv \delta_Y^I < \infty.$$

OLA

This section compares $\mathbb{E}_t V_t(v|\sigma^A, s_t^I(v), O, e^I, L, z^I)$ with $\mathbb{E}_t V_t(v|1, s_t^I(v), O, e^I, L, z^I)$ to determine the project size for an old, low skill adopter, namely, $s_t^A(v|O, L)$.

$$\begin{aligned}
&\mathbb{E}_t V_t(v|\sigma^A, s_t^I(v), O, e^I, L, z^I) \\
&= (1-\mu^A-\mu^I)\delta N \left[\sigma^A \left(\theta^A\right) \bar{A}_{t-1} + s_t^I(v) \left(\theta^I + \gamma_t^I(v)\right) A_{t-1}\right] \\
&- \max\left\{\phi^A \kappa^A \bar{A}_{t-1} - \widehat{RE}_t^A, 0\right\} - \max\left\{k_t^I - \widehat{RE}_t^I, 0\right\}.
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_t V_t(v|1, s_t^I(v), O, e^I, L, z^I) \\
&= (1-\mu^A-\mu^I)\delta N \left[\left(\theta^A\right) \bar{A}_{t-1} + s_t^I(v) \left(\theta^I + \gamma_t^I(v)\right) A_{t-1}\right] \\
&- \max\left\{\kappa^A \bar{A}_{t-1} - \widehat{RE}_t^A, 0\right\} - \max\left\{k_t^I - \widehat{RE}_t^I, 0\right\}.
\end{aligned}$$

Note that

$$\widehat{RE}_t^A \geq \mu^A \delta N \left(\sigma^A \theta^A \bar{A}_{t-2} + \theta^I A_{t-2}\right). \tag{3.A.2.1}$$

OLI

This subsection compares $\mathbb{E}_t V_t(v|s_t^A, \sigma^I, e^A, O, z^A, L)$ with $\mathbb{E}_t V_t(v|s_t^A, 1, e^A, O, z^A, L)$ to determine the project size for an old, low skill innovator, namely, $s_t^I(v|O, L)$.

$$\begin{aligned}
&\mathbb{E}_t V_t(v|s_t^A, \sigma^I, e^A, O, z^A, L) \\
&= (1-\mu^A-\mu^I)\delta N \left[s_t^A(v) \left(\theta^A + \gamma_t^A(v)\right) \bar{A}_{t-1} + \theta^I A_{t-1}\right] \\
&- \max\left\{k_t^A - \widehat{RE}_t^A, 0\right\} - \max\left\{\phi^I \kappa^I \bar{A}_{t-1} - \widehat{RE}_t^I, 0\right\}.
\end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}_t V_t(v|s_t^A, 1, e^A, O, z^A, L) \\
&= (1-\mu^A-\mu^I)\delta N\left[s_t^A(v)\left(\theta^A+\gamma_t^A(v)\right)\bar{A}_{t-1}+\theta^I A_{t-1}\right] \\
&- \max\left\{k_t^A-\widehat{RE}_t^A, 0\right\} - \max\left\{\kappa^I\bar{A}_{t-1}-\widehat{RE}_t^I, 0\right\}.
\end{aligned}
$$

Note that

$$
\widehat{RE}_t^I \geq \mu^I\delta N\left(\sigma^A\theta^A\bar{A}_{t-2}+\theta^I A_{t-2}\right), \tag{3.A.2.2}
$$

namely, the payment to an innovator in the previous period, if he happened to be working with a low skill adopter (this is the worst situation).

Combining (3.A.2.1) and (3.A.2.2), if $N \geq \max\left\{\frac{(1+g)\kappa^I}{\mu^I\delta\sigma^A\theta^A}, \frac{(1+g)\kappa^A}{\mu^A\delta\sigma^A\theta^A}\right\} < \infty$, then both OLI and OLA can fully finance large projects, in their respective sector. And under this condition imposed upon N, it follows that $s_t^I(v|O,L) = s_t^A(v|O,L) = 1$. If old low skill managers run large projects, their high skill counterparts will obviously run large projects, both because the principal can have higher profit, and because the high skill managers have higher ability to finance the large projects, ceteris paribus.

3.8.3 Proofs for Propositions 1 and 2

The analysis below shows the full proofs for Propositions 1 and 2, namely how the retains rules for old adopters and innovators are derived.

OHA

Recall from above that if $N \geq \frac{(1+g)\kappa^A}{\mu^A\delta\sigma^A\theta^A}$, old adopters, either high or low skill, can fully finance a large adoption project, which implies OHA are always be retained, namely $R_t^A(H) = 1$.[23]

OLA

From lemma 4, OLA run large projects if retained, namely, the expected value of a firm with OLA is

$$
\begin{aligned}
\mathbb{E}_t V_t =& (1-\mu^A-\mu^I)\delta N\left[\left(\theta^A+p^A\gamma^A\right)\bar{A}_{t-1}+s_t^I(v)\left(\theta^I+\mathbb{E}_t\gamma_t^I(v)\right)A_{t-1}\right] \\
&- \max\left\{\kappa^A\bar{A}_{t-1}-\widehat{RE}_t^A, 0\right\} - \max\left\{k_t^I-\widehat{RE}_t^I, 0\right\}.
\end{aligned}
$$

[23]Since retaining rules only apply to old managers, $R_t^A(O,H)$ is simplified to $R_t^A(H)$, omitting O in the notation.

While the expected value of hiring YA (who runs small projects) is

$$\begin{aligned}\mathbb{E}_t V_t =&(1-\mu^A-\mu^I)\delta N\left[\sigma^A\left(\theta^A+\lambda^A\gamma^A\right)\bar{A}_{t-1}+s_t^I(v)\left(\theta^I+\mathbb{E}_t\gamma_t^I(v)\right)A_{t-1}\right]\\ &-\max\left\{\phi^A\kappa^A\bar{A}_{t-1}-\widehat{RE}_t^A,0\right\}-\max\left\{k_t^I-\widehat{RE}_t^I,0\right\}.\end{aligned}$$

Comparing these two expressions, it is easy to check that retaining OLA is worthwhile if

$$(1-\mu^A-\mu^I)\delta N\left(\theta^A+p^A\gamma^A-\sigma^A\theta^A-\sigma^A\lambda^A\gamma^A\right)\bar{A}_{t-1}+\phi^A\kappa^A\bar{A}_{t-1}\geq 0$$

$$\Longleftrightarrow$$

$$p^A\geq\sigma^A\lambda^A-\frac{(1-\sigma^A)\theta^A}{\gamma^A}-\frac{\phi^A\kappa^A}{\gamma^A(1-\mu^A-\mu^I)\delta N}\equiv p^A(L)<1.$$

The retaining rule for an OLA is given below

$$R_t^A(L)=\begin{cases}1 & \text{if} \quad p^A\geq p^A(L),\\ 0 & \text{if} \quad p^A<p^A(L).\end{cases}$$

OLI

From lemma 4, OLI run large projects if retained, namely, the expected value of a firm with OLI is

$$\begin{aligned}\mathbb{E}_t V_t =&(1-\mu^A-\mu^I)\delta N\left[s_t^A(v)\left(\theta^A+\mathbb{E}_t\gamma_t^A(v)\right)\bar{A}_{t-1}+\left(\theta^I\right)A_{t-1}\right]\\ &-\max\left\{k_t^A-\widehat{RE}_t^A,0\right\}-\max\left\{\kappa^I\bar{A}_{t-1}-\widehat{RE}_t^I,0\right\}.\end{aligned}$$

While the expected value of hiring YI (who runs small projects) is

$$\begin{aligned}\mathbb{E}_t V_t =&(1-\mu^A-\mu^I)\delta N\left[s_t^A(v)\left(\theta^A+\mathbb{E}_t\gamma_t^A(v)\right)\bar{A}_{t-1}+\left(\theta^I+\sigma^I\lambda^I\gamma^I\right)A_{t-1}\right]\\ &-\max\left\{k_t^A-\widehat{RE}_t^A,0\right\}-\max\left\{\phi^I\kappa^I\bar{A}_{t-1}-\widehat{RE}_t^I,0\right\}.\end{aligned}$$

Comparing these two expressions, it is easy to check that dismissing OLI is worthwhile if

$$(1-\mu^A-\mu^I)\delta N a_{t-1}\sigma^I\lambda^I\gamma^I\geq\phi^I\kappa^I$$

$$\Longleftrightarrow$$

$$a_{t-1}\geq\frac{\phi^I\kappa^I}{(1-\mu^A-\mu^I)\delta N\sigma^I\lambda^I\gamma^I}\equiv a(L,I).$$

The retaining rule for OLI is

$$R_t^I(L)=\begin{cases}1 & \text{if} \quad a_{t-1}\leq a(L,I),\\ 0 & \text{if} \quad a_{t-1}\geq a(L,I).\end{cases}$$

OHI

This subsection solves the retaining rule of OHI, which turns out to depend both on a country's distance to frontier, and the severity of individual aging (this already hints that the impacts of individual aging are related with a country's distance to frontier, to some extent).

From lemma 4, OHI run large projects if retained, namely, the expected value of a firm with OHI is

$$\begin{aligned}\mathbb{E}_t V_t =&(1-\mu^A-\mu^I)\delta N\left[s_t^A(v)\left(\theta^A+\mathbb{E}_t\gamma_t^A(v)\right)\bar{A}_{t-1}+\left(\theta^I+(1-p^I)\gamma^I\right)A_{t-1}\right]\\&-\max\left\{k_t^A-\widehat{RE}_t^A,0\right\}-\max\left\{\kappa^I\bar{A}_{t-1}-\widehat{RE}_t^I,0\right\}.\end{aligned}$$

While the expected value of hiring YI (who run small projects) is

$$\begin{aligned}\mathbb{E}_t V_t =&(1-\mu^A-\mu^I)\delta N\left[s_t^A(v)\left(\theta^A+\mathbb{E}_t\gamma_t^A(v)\right)\bar{A}_{t-1}+\left(\theta^I+\sigma^I\lambda^I\gamma^I\right)A_{t-1}\right]\\&-\max\left\{k_t^A-\widehat{RE}_t^A,0\right\}-\max\left\{\phi^I\kappa^I\bar{A}_{t-1}-\widehat{RE}_t^I,0\right\}.\end{aligned}$$

Comparing these two expressions, it is easy to check that retaining OHI is worthwhile if

$$\phi^I\kappa^I\geq(1-\mu^A-\mu^I)\delta N a_{t-1}\left[\sigma^I\lambda^I\gamma^I-(1-p^I)\gamma^I\right]. \tag{R-OHI}$$

If $\sigma^I\lambda^I\gamma^I-(1-p^I)\gamma^I\leq 0$, then (R-OHI) always holds, implying OHI are always retained if $p^I\leq 1-\sigma^I\lambda^I$. This is intuitive since OHI will be dismissed *only if* individual aging lowers innovation skill of an OHI severely enough.

If $p^I>1-\sigma^I\lambda^I$, then

$$\text{(R-OHI)}$$

$$\Longleftrightarrow$$

$$a_{t-1}<\frac{\phi^I\kappa^I}{(1-\mu^A-\mu^I)\delta N\left(\sigma^I\lambda^I\gamma^I-(1-p^I)\gamma^I\right)}\equiv a(H,I),$$

hence the retaining rule for an OHI is summarized below

$$R_t^I(H)=\begin{cases}1 & \text{if} \quad p^I\leq 1-\sigma^I\lambda^I \quad \text{OR} \quad a_{t-1}<a(H,I),\\ 0 & \text{if} \quad p^I>1-\sigma^I\lambda^I \quad \text{AND} \quad a_{t-1}>a(H,I).\end{cases}$$

3.8.4 Endogenized law of motion

This section shows the details about how equations (3.8) to (3.13) are derived.

Productivity of firms owned by young principals

For a firm owned by a young principal (called young firm), only young agents will be hired and small projects are chosen, so $s_t^A = \sigma^A$ and $s_t^I = \sigma^I$, moreover the adopter's expected adoption ability is $\lambda^A \gamma^A + (1-\lambda^A)\cdot 0 = \lambda^A \gamma^A$ and the innovator's expected innovation ability is $\lambda^I \gamma^I$. Putting all of these into equation (3.6) gives the productivity of a young firm as

$$A_t^Y(v) = \sigma^A(\theta^A + \lambda^A\gamma^A)\bar{A}_{t-1} + (\theta^I + \sigma^I\lambda^I\gamma^I)A_{t-1},$$

where the subscript Y denotes the firm is owned by a young principal. Assuming all firms (of the unit mass) are owned by young principals, the previous equations gives

$$A_t^Y \equiv \int_0^1 A_t^Y(v)dv = \sigma^A(\theta^A + \lambda^A\gamma^A)\bar{A}_{t-1} + (\theta^I + \sigma^I\lambda^I\gamma^I)A_{t-1},$$

namely, A_t^Y is the aggregate productivity assuming all firms are owned by young agents. A_t^Y and $A_t^Y(v)$ have the same expression due to the unit mass of firms and symmetry across all young firms, ex ante.

Productivity of firms owned by old principals

For an old principal, he can choose to retain his managers or hire young managers.[24]

The productivity of old firms is expressed succinctly as:

$$\begin{aligned}
A_t^O \equiv & \int_0^1 A_t^O(v)dv \\
=&(1-\lambda^A)\left[R_t^A(L)\cdot 1\cdot(\theta^A + p^A\gamma^A) + (1-R_t^A(L))\sigma^A(\theta^A+\lambda^A\gamma^A)\right]\bar{A}_{t-1} \\
&+ \lambda^A\cdot R_t^A(H)\cdot 1\cdot(\theta^A+\gamma^A)\bar{A}_{t-1} \\
&+ (1-\lambda^I)\left[R_t^I(L)\cdot(\theta^I + 1\cdot 0) + (1-R_t^I(L))(\theta^I+\sigma^I\lambda^I\gamma^I)\right]A_{t-1} \\
&+ \lambda^I\left[R_t^I(H)\cdot(\theta^I + 1\cdot(1-p^I)\gamma^I) + (1-R_t^I(H))(\theta^I+\sigma^I\lambda^I\gamma^I)\right]A_{t-1},
\end{aligned}$$

where the first two lines of the above formula denote contribution from the adoption sector. With probability λ^A the adopter turns out to be OHA, in which case he is retained (proposition 1) and runs a large project (lemma 4). With probability $1-\lambda^A$ the adopter turns out to be OLA, in which case he could be dismissed or retained. If retained, he runs a large

[24] In this chapter, if a principal dismisses his old agent, the principal will only hire a young agent. The principal will never 'steal' an old agent from other principals, nor hires the other old agent from his own firm. These two cases are covered in appendix A5.

project (lemma 4), but his expected adoption ability is $p^A\gamma^A$ instead of γ^A (due to individual aging); if dismissed, a YI is hired, who runs a small project (lemma 4) with expected adoption ability $\lambda^A\gamma^A$. Similar logic applies to the innovation sector. Summing up both adoption and innovation yields the above formula.

Substitute in the values for retaining rules to get A_t^O in each case gives

- Case **I(i)**: $R_t^A(L) = 0, R_t^I(L) = R_t^I(H) = 1$,

$$A_t^O(v) = \lambda^A(\theta^A + \gamma^A)\bar{A}_{t-1} + (1-\lambda^A)\sigma^A(\theta^A + \lambda^A\gamma^A)\bar{A}_{t-1} + \lambda^I\left(\theta^I + (1-p^I)\gamma^I\right)A_{t-1} + (1-\lambda^I)\theta^I A_{t-1}.$$

- Case **I(ii)**: $R_t^A(L) = 0, R_t^I(L) = 0, R_t^I(H) = 1$,

$$A_t^O(v) = \lambda^A(\theta^A + \gamma^A)\bar{A}_{t-1} + (1-\lambda^A)\sigma^A(\theta^A + \lambda^A\gamma^A)\bar{A}_{t-1} + \lambda^I\left(\theta^I + (1-p^I)\gamma^I\right)A_{t-1} + (1-\lambda^I)\left(\theta^I + \sigma^I\lambda^I\gamma^I\right)A_{t-1}.$$

- Case **I(iii)**: $R_t^A(L) = 0, R_t^I(L) = R_t^I(H) = 0$,

$$A_t^O(v) = \lambda^A(\theta^A + \gamma^A)\bar{A}_{t-1} + (1-\lambda^A)\sigma^A(\theta^A + \lambda^A\gamma^A)\bar{A}_{t-1} + (\theta^I + \sigma^I\lambda^I\gamma^I)A_{t-1}.$$

- Case **II(i)**: $R_t^A(L) = 1, R_t^I(L) = R_t^I(H) = 1$,

$$A_t^O(v) = \lambda^A(\theta^A + \gamma^A)\bar{A}_{t-1} + (1-\lambda^A)(\theta^A + p^A\gamma^A)\bar{A}_{t-1} + \lambda^I\left(\theta^I + (1-p^I)\gamma^I\right)A_{t-1} + (1-\lambda^I)\theta^I A_{t-1}.$$

- Case **II(ii)**: $R_t^A(L) = 1, R_t^I(L) = 0, R_t^I(H) = 1$,

$$A_t^O(v) = \lambda^A(\theta^A + \gamma^A)\bar{A}_{t-1} + (1-\lambda^A)(\theta^A + p^A\gamma^A)\bar{A}_{t-1} + \lambda^I\left(\theta^I + (1-p^I)\gamma^I\right)A_{t-1} + (1-\lambda^I)\left(\theta^I + \sigma^I\lambda^I\gamma^I\right)A_{t-1}.$$

- Case **II(iii)**: $R_t^A(L) = 1, R_t^I(L) = R_t^I(H) = 0$,

$$A_t^O(v) = \lambda^A(\theta^A + \gamma^A)\bar{A}_{t-1} + (1-\lambda^A)\sigma^A(\theta^A + \lambda^A\gamma^A)\bar{A}_{t-1} + (\theta^I + \sigma^I\lambda^I\gamma^I)A_{t-1}.$$

Since half of the firms are young and half are old,

$$\int_0^1 A_t(v)dv = \frac{1}{2}\int_0^1 A_t^Y(v)dv + \frac{1}{2}\int_0^1 A_t^O(v)dv = \frac{A_t^Y + A_t^O}{2}.$$

Combining this with $a_t \equiv A_t/\bar{A}_t$ gives the following:

- Case **I(i)**: $R_t^A(L) = 0, R_t^I(L) = R_t^I(H) = 1,$

$$\begin{aligned}2(1+g)a_t = &\left[\left(\lambda^A + (1-\lambda^A)\sigma^A + \sigma^A\right)\theta^A + \left(1 + (1-\lambda^A)\sigma^A + \sigma^A\right)\lambda^A\gamma^A\right]\\ &+ \left[2\theta^I + (\sigma^I + 1 - p^I)\lambda^I\gamma^I\right].\end{aligned}$$

- Case **I(ii)**: $R_t^A(L) = 0, R_t^I(L) = 0, R_t^I(H) = 1,$

$$\begin{aligned}2(1+g)a_t = &\left[\left(\lambda^A + (1-\lambda^A)\sigma^A + \sigma^A\right)\theta^A + \left(1 + (1-\lambda^A)\sigma^A + \sigma^A\right)\lambda^A\gamma^A\right]\\ &+ \left[2\theta^I + \left(\sigma^I + 1 - p^I + (1-\lambda^I)\sigma^I\right)\lambda^I\gamma^I\right].\end{aligned}$$

- Case **I(iii)**: $R_t^A(L) = 0, R_t^I(L) = R_t^I(H) = 0,$

$$\begin{aligned}2(1+g)a_t = &\left[\left(\lambda^A + (1-\lambda^A)\sigma^A + \sigma^A\right)\theta^A + \left(1 + (1-\lambda^A)\sigma^A + \sigma^A\right)\lambda^A\gamma^A\right]\\ &+ \left[2\theta^I + 2\sigma^I\lambda^I\gamma^I\right].\end{aligned}$$

- Case **II(i)**: $R_t^A(L) = 1, R_t^I(L) = R_t^I(H) = 1,$

$$\begin{aligned}2(1+g)a_t = &\left[\left(1 + \sigma^A\right)\theta^A + \left(\sigma^A\lambda^A + \lambda^A + (1-\lambda^A)p^A\right)\gamma^A\right]\\ &+ \left[2\theta^I + (\sigma^I + 1 - p^I)\lambda^I\gamma^I\right].\end{aligned}$$

- Case **II(ii)**: $R_t^A(L) = 1, R_t^I(L) = 0, R_t^I(H) = 1,$

$$\begin{aligned}2(1+g)a_t = &\left[\left(1 + \sigma^A\right)\theta^A + \left(\sigma^A\lambda^A + \lambda^A + (1-\lambda^A)p^A\right)\gamma^A\right]\\ &+ \left[2\theta^I + \left(\sigma^I + 1 - p^I + (1-\lambda^I)\sigma^I\right)\lambda^I\gamma^I\right].\end{aligned}$$

- Case **II(iii)**: $R_t^A(L) = 1, R_t^I(L) = R_t^I(H) = 0,$

$$\begin{aligned}2(1+g)a_t = &\left[\left(1 + \sigma^A\right)\theta^A + \left(\sigma^A\lambda^A + \lambda^A + (1-\lambda^A)p^A\right)\gamma^A\right]\\ &+ \left[2\theta^I + 2\sigma^I\lambda^I\gamma^I\right].\end{aligned}$$

3.8.5 Two issues about employment rules

In this section, two issues about employment rules are studied with more details. Firstly, in the model it is assumed that if an old manager is dismissed then he must work as a manual labor. He can not work as a manager hired by other principals. Secondly, if the old manager was an adopter (innovator) when young, he will not work as an innovator (adopter) when old. The analysis below explains these two issues.

No principal will hire agents from other principals

If an old principal finds his innovator (similar logic applies to adopter as well) to be low skill, he could dismiss the low skill innovator, and 'steal' a high skill innovator from other principals by offering a high wage. This is also mentioned in AAZ (2006) where they *assume* the value of a firm hiring a low skill innovator is not smaller than the value of a firm hiring a high skill innovator (equation (15) in their paper). First of all, one extra (very restricted) assumption can solve the problem. However, in this model a very mild assumption can also solve this issue.

In section 3.2, it is assumed that if a young agent works as an innovator, at the end of the first period, his innovation skill is revealed to him and his principal only, but *not to other agents or principals*. Under this mild assumption, the next paragraph shows that no principal will hire other principals' managers by offering a higher wage.

Suppose principal 1 finds his innovator (say innovator 1) to be low skill, and offers a high wage and suppose an innovator (say innovator 2) from some other principal (say principal 2) comes to take the offer. Now the question becomes: does innovator 2 have high innovator skill or not? If yes, then obviously principal 2 knows this, so principal 2 should have more incentives to retain innovator 2. If innovator 2 is high skill, then the first period profit for principal 2 is higher than that for principal 1, which means principal 2 has more resources (money) to retain innovator 2. Therefore, principal 2 has both more incentives and more resources to retain innovator 2, if innovator is really high ability; and if principal does not retain innovator 2, innovator 2 must have low innovation skill. Keeping this in mind, neither should principal 1 has any incentive to hire innovator 2 (let alone at a higher wage offered by principal 2). As a result, no principal will hire managers from other principals (due to information asymmetry and resource asymmetry).

Conditions ruling out 'cross employment'

If a principal finds his adopter and innovator happen to be low skill simultaneously, it is possible that the principal will dismiss his OLA and OLI. However, instead of hiring a young agent, the principal could hire his OLI as the new adopter, since the principal only knows the OLI's innovation ability, and the OLI could have high adoption ability. This situation is named "cross employment"in this model. This section gives conditions ruling out cross employment, hence a principal will always hire a young agent if he dismisses his current old agents.

There are two types of cross employment: replacing OLI with OLA, and replacing OLA with OLI. Since the conditions ruling out each of them are very similar, only the first case is presented: replacing OLI with OLA.

If the principal dismisses OLI and OLA, and hires the OLA as the new innovator, the firm's expected value is:

$$[\text{contribution from adoption sector}] + (1-\mu^I)(\delta N)\left[\theta^I + (1-\lambda^I)p^I\right]A_{t-1} - 0\ ;$$

where the firm's expected value of hiring a young agent is:

$$[\text{contribution from adoption sector}] + (1-\mu^I)(\delta N)\left[\theta^I + (1-\lambda^I)\right]A_{t-1} - \phi^I\kappa^I\bar{A}_{t-1}\ .$$

The principal will prefer hiring a young agent rather than the OLA if

$$(1-\mu^I)(\delta N)\left[\theta^I + (1-\lambda^I)p^I\right]A_{t-1} < (1-\mu^I)(\delta N)\left[\theta^I + (1-\lambda^I)\right]A_{t-1} - \phi^I\kappa^I\bar{A}_{t-1}$$

$$\Leftrightarrow$$

$$\frac{\phi^I\kappa^I}{(1-\mu^I)\delta(1-\lambda^I)(1-p^I)a_{t-1}} < N^2.$$

Under the assumption that $\exists\bar{a} > 0$ such that $a_t \geq \bar{a}$ for all t and all countries, then $\exists\bar{N} < \infty$ such that for $N > \bar{N}$, then there is no 'cross-employment'.

3.8.6 Endogenize world technical growth rate

In this chapter's model, it is assumed that the world technical growth rate g is exogenous. This subsection shows how to endogenize g. Working with an endogenous g does not qualitatively change the model results. Because of this, g is assumed exogenous in the main model.

From section 3.4, intercept **II** > intercept **I** if $p^A = 1$, and it follows that

$$\text{intercept } \mathbf{II}|_{p^A=1} = (1+\sigma^A)\theta^A + (1+\sigma^A\lambda^A)\gamma^A,$$

$$\text{slope } \mathbf{ii}|_{p^I=0} = 2\theta^I + \left[\sigma^I + 1 + (1-\lambda^I)\sigma^I\right]\lambda^I\gamma^I > \text{slope } \mathbf{iii},$$
$$\text{slope } \mathbf{ii}|_{p^I=1} = 2\theta^I + \left[\sigma^I + (1-\lambda^I)\sigma^I\right]\lambda^I\gamma^I < \text{slope } \mathbf{iii}.$$

It is assumed that the world technology frontier is determined by the most advanced country, namely $a_t = a_{t-1} = 1$, then using intercept $\mathbf{II}|_{p^A=1}$ and slope $\mathbf{ii}|_{p^I=0}$, g can be endogenized as

$$2(1+g) = (1+\sigma^A)\theta^A + (1+\sigma^A\lambda^A)\gamma^A + \left[2\theta^I + \sigma^I + 1 + (1-\lambda^I)\sigma^I\right]\lambda^I\gamma^I,$$
$$g = \frac{1}{2}\left[(1+\sigma^A)\theta^A + (1+\sigma^A\lambda^A)\gamma^A + (2\theta^I + 2\sigma^I + 1 - \lambda^I\sigma^I)\lambda^I\gamma^I\right] - 1\ .$$

3.8.7 Parameters and endogenous variables in used in the numerical analysis in section 3.6

This section contains the three tables mentioned in 3.6.2. They list the numbers for parameters and endogenous variables, used in 3.6.3 to test Propositions 3, 4 and 5.

Table 1: Exogenous parameters

Case	λ^A	λ^I	σ^A	σ^I	θ^A	θ^I	γ^A	γ^I	ϕ^A	κ^A	μ^A	μ^I	δ	N
1	0.5	0.5	0.6	0.6	0.7	0.7	1	1	0.8	0.7	0.2	0.3	0.5	1000
2	0.8	0.9	0.7	0.65	0.63	0.8	2	3	0.8	0.9	0.5	0.1	0.6	10
3	0.9	0.3	0.99	0.5	0.4	0.74	1.53	0.49	0.3	0.67	0.23	0.43	0.78	23
4	0.87	0.9	0.5	0.8	1	2	5	7	0.6	0.75	0.12	0.09	0.1	20
5	0.65	0.72	0.8	0.65	0.58	0.63	1.5	2.8	0.38	0.45	0.08	0.05	0.42	200
6	0.78	0.75	0.45	0.85	0.79	0.81	1.32	2.04	0.45	0.62	0.1	0.095	0.58	156

Table 2: Values for three variables in Proposition 3

Case	$p^A(L)$	$\sigma^A\lambda^A - \frac{(1-\sigma^A)\theta^A}{\gamma^A}$	$1 - \sigma^I\lambda^I$
1	0.01776	0.02	0.7
2	0.3155	0.4655	0.415
3	0.86685	0.88839	0.85
4	0.27804	0.335	0.28
5	0.44111	0.44267	0.532
6	0.01893	0.02183	0.3625

Table 3: Values for p^A, p^I, two intercepts and three slopes

Case	p^A	p^I	Intercept **I**	Intercept **II**	Slope **i**	Slope **ii**	Slope **iii**
1	0.01	0.5	1.93	1.925	1.95	2.1	2
	0.0185	0.6	1.93	1.92925	1.9	2.05	2
	0.025	0.8	1.93	1.9325	1.8	1.95	2
	0.03	0.85	1.93	1.935	1.775	1.925	2
2	0.2	0.2	3.9772	3.871	5.515	5.6905	5.11
	0.4	0.3	3.9772	3.951	5.245	5.4205	5.11
	0.5	0.5	3.9772	3.991	4.705	4.8805	5.11
	0.6	0.6	3.9772	4.031	4.435	4.6105	5.11
3	0.7	0.7	3.67215	3.64333	1.5976	1.64905	1.627
	0.88	0.8	3.67215	3.67087	1.5829	1.63435	1.627
	0.91	0.9	3.67215	3.67546	1.5682	1.61965	1.627
	0.93	0.95	3.67215	3.67852	1.56085	1.6123	1.627
4	0.25	0.2	8.24275	8.1875	14.08	14.584	14.08
	0.3	0.25	8.24275	8.22	13.765	14.269	14.08
	0.35	0.3	8.24275	8.2525	13.45	13.954	14.08
	0.4	0.35	8.24275	8.285	13.135	13.639	14.08
5	0.4	0.45	3.0314	3.009	3.6792	4.04611	3.8808
	0.442	0.5	3.0314	3.03105	3.5784	3.94531	3.8808
	0.45	0.55	3.0314	3.03525	3.4776	3.84451	3.8808
	0.452	0.6	3.0314	3.0363	3.3768	3.74371	3.8808
6	0.015	0.25	2.64476	2.64278	4.068	4.39313	4.221
	0.02	0.3	2.64476	2.64423	3.9915	4.31663	4.221
	0.022	0.4	2.64476	2.64481	3.8385	4.16363	4.221
	0.025	0.45	2.64476	2.64568	3.762	4.08713	4.221

Chapter 4

Population aging, educational effort and economic growth

4.1 Introduction

Population aging has become one of the most important demographic phenomena facing many countries in the world. In many countries, the distribution of the population is shifted toward older ages (Weil, 2006).

Numerous studies have investigated the macroeconomic implications of demographic change. Early studies use calibrated general equilibrium overlapping-generation models to study the relationship between demographic structure, saving and capital accumulation. This method was first introduced by Auerbach and Kotlikoff (1987) and later followed by studies such as Auerbach et al. (1989), Miles (1999) and Hviding and Merette (1998). These papers all use large-scale simulations, and to analyze the same question, but from a simpler and more theoretical point of view, the first objective of this chapter is to investigate the impact of population aging upon the macroeconomic performance such as capital accumulation and output growth, in a two-period overlapping-generation model.

The literature mentioned above is silent on how economic activities affect the process of technological change and the connection between demography and technological change. The missing of modeling technological progress could yield limitations on implications of the models (Cutler et al. 1990). Though assuming away endogenous technological progress in their simulated models, Hviding and Merette (1998, p.30) admits that a purely neo-classical production function fails to include any positive spillover effects from increased human capital investment.

The modern economic growth theory, developed by Romer (1986) and Lucas (1988), has emphasized factors such as learning by doing, spillover effects, especially human capital investment and research and development (R&D) as key components of economic growth. There are numerous studies focusing on the connection between population aging, economic growth, education investment and human capital accumulation, such as de la Croix and Licandro (1999), Hu (1999), Kalemli-Ozcan et al. (2000), Boucekkine et al. (2002), Echevarria and Iza (2006), and Heijdra and Romp (2009). They find that an increase in longevity, as a key feature of population aging, encourages human capital investment and this contributes to economic growth.

Despite of making technological progress endogenous, many of the models mentioned in the previous paragraph assume that the final production sector uses only one factor (either human capital or effective labor) or they work with only small open economies to pin down the interest rate.[1] Another problem with the models is that they assume that all agents are identical and will all accumulate some human capital. In reality, on the other hand, both unskilled labor (such as manual workers) and skilled labor (such as scientists) exist simultaneously, doing different kinds of jobs (such as manual work and academic research respectively).

The second objective of the chapter is to investigate how population aging affects educational effort, in a general equilibrium model. In contrast with the current literature, the model in this chapter features both unskilled and skilled labor at the same time, and moreover, this chapter results show that this is a possible outcome even if all new born agents are identical.

Early literature with exogenous technological progress finds a negative impact of population aging upon physical capital accumulation, output and economic growth, while recent literature emphasizing the endogeneity of technological change finds the opposite. This seems to suggest that having endogenous technological progress is necessary for population aging to boom economic growth. However, there is no paper formally showing that having endogenous technological progress is necessary for population aging to have positive effects on economic growth. To this end, the third objective of the thesis is to analyze, in a model with essentially neo-classical growth progress, whether population aging has positive effects on physical capital accumulation, output and economic growth.

As emphasized earlier on, this chapter features human capital accumulation and endogenous technological progress. To obtain an essentially neo-classical growth feature, an insight is borrowed from Acemoglu (2009, chapter 13). The marginal costs of research machines are introduced in a special way such that the resulting model is neo-classical in essence (more details in section 4.3). This is novel compared to the current literature and the benefits obtained

[1]For tractability, the literature has eliminated the endogeneity of interest rate. For example, Boucekkine et al. (2002) achieve the constancy of the interest rate by assuming that the felicity function is linear, i.e. that the inter-temporal substitution elasticity is infinite. Heijdra and Romp (2009), closely related to Boucekkine et al. (2002), assume that the economy is of a small size and has access to well-functioning markets including the world capital market thus the interest rate is exogenously given and constant. The latter has the advantage that they can postulate a concave felicity function, giving rise to well-defined consumption profiles.

from doing so will be highlighted in section 4.4.

The rest of the chapter is organized as follows. Section 4.2 describes the model structure, including the problems in the household, production and research and development sectors. Section 4.3 solves the the optimization problems of all sectors for each period and derives the steady state equilibrium. Under the steady state equilibrium, section 4.4 analyzes the impacts of population aging on education, capital accumulation, output growth and welfare. As an extension to the baseline model in 4.2, section 4.5 introduces agents' ability heterogeneity in doing education and studies the impacts of population aging. In the end, section 4.6 concludes the chapter with a summary of the main results and some future research ideas.

4.2 Model structure

This section describes the structure of the model. This chapter aims to study the impacts of population aging upon education, economic growth and welfare. In doing so, this chapter utilizes the overlapping-generation model and extends it to allow for population aging. This model consists of the household sector, final production sector and research and development sector and the model is general equilibrium. Below, the details of each sector are described.

4.2.1 Household sector

The model in this chapter is a variant of the classic overlapping-generation model by Diamond (1965). In this economy where time is discrete, all agents can live for at most two periods, young and old. Compared to the textbook model where all agents can survive into the old age for sure, uncertainties are introduced to agents' survival into the old age. Instead of surviving with certainty, each agent has probability $s \in (0, 1)$ of surviving into the old age, and s is called the survival rate. Similar to Ludwig and Vogel (2010), population aging is defined via an increase in the survival rate s. If s increases, there are more survivals each period and hence a larger population size. Suppose each period the size of new-born agents is normalized to one, then during each period of time, the ratio of old people to young people is equal to s. As s increases, there are more old people per unit of young agent in the economy, reflecting a shift in the demographic structure towards older ages. This is one of the key features of population aging.

In this model, it is assumed that the young agents give birth to the next generation, and the birth takes place at the end of their young age, *before* they enter the old age, before any sudden death (dying before entering the old age) takes place. When the next generation are born, the parents immediately become old and the previous old generation die. This specification ensures that the number of new-born agents is not affected by the survival rate s. This also rules out any confusing interaction between survival rate and population growth rates, permitting a focus on the impacts of a higher survival rate upon the economy.

When young agents are born, they are identical and unskilled. An unskilled labor can choose to go to school, do education, and become skilled labor later on, or he can simply stay unskilled. If he chooses to stay as an unskilled labor, during youth, he works in the final production sector, earns income, makes consumption decision for his youth, and saves for his old age.[2] If he succeeds in surviving into the old age, he will consume his previous-period saving (plus any interest earned). Since survival into the old (retirement) period is uncertain, one problem naturally rising here is how to treat the savings of the agents who die before reaching old age (those failing to surviving into the old age). In the literature, there are generally two methods. Ehrlich and Lui (1999) assume that these savings are simply wasted. Zhang et al. (1995), Ludwig and Vogel (2010) and many others assume there are actuarially fair annuity markets and the savings (plus any interest earned) are divided evenly among those survivals to their old age.

The literature that assumes the existence of actuarially fair annuity markets for savings of young agents all follow the seminal works of Blanchard (1985) and Yaari (1965). The idea is to assume that there are perfect annuity markets against survival risks of unskilled labor. Namely, if an unskilled labor fails to survive into the old age, his wealth is given to the annuity company at his death. In return, if he does survive, he will get an interest higher than the pure market interest (this also provides incentives for him to hold some of his wealth when alive, even if he faces the risk of not spending all his asset in the case of sudden death). In this setting, the assumption of a perfect annuity market is equivalent to: the savings of all unskilled labor born at at t, plus the interest earned at the pure market rate, are distributed evenly among those survived and old unskilled labor at $t+1$. In this model, since all unskilled labor born at the same time are identical, ex ante (before they know if they actually survive into old), they can be treated as one representative agent.

Alternative to staying unskilled, a new-born agent can choose to do education when young.

[2]His saving will go to investment and accumulation of physical capital, K.

In this model, if an unskilled labor goes to school when young and if he does survive into the old age, he will become a skilled labor and work in the research and development (R&D) sector, called a scientist for simplicity. A schooling agent cannot work when young (he does research and gets payoff only in the old age), but he still has to consume, so if an unskilled labor decides to do education, he must borrow during young to finance his consumption.[3] When young, he decides the amount to borrow for consumption. When old, he works as a scientist and gets payoff, pays back his previous-period borrowing (plus interest accrued) and consumes the amount left.

Besides the survival risks for unskilled labor (some of their assets are not consumed at death), this model also has the survival risks of students. This problem is raised also by Kalemli-Ozcan, Ryder and Weil (2001), who include another annuity market. If a schooling agent[4] does not survive to the old age, his borrowing during young cannot be paid back by himself and this will cause loss to the agent or company who lends him the money. Therefore, to ensure the market of such annuity company perfect (zero profit and zero loss), the following is assumed: all borrowings of students at t, plus the interest at the pure market rate, are distributed evenly and paid back by those students surviving to $t+1$.[5] This sounds a little "unfair" to the students who do survive into the old age, since they must pay back the borrowing of some other agents. However, this policy is set exogenously by the company lending money in the first place, so all students must already take this into account when deciding whether to do education or not. Moreover, as in Kalemli-Ozcan et al. (2000), this is the only way to ensure all survival risks are taken into account.

Upon being born, an unskilled agent must choose between staying unskilled or going to school.

[3]In this model skilled labor get payoff in the old rather than young age. This is to have a natural analogy between physical and human capital accumulation. If an agent saves in terms of physical capital accumulation, he gets return in the old period, so it is natural to have the payoff of accumulating human capital also in the old period. Since doing education is often thought as accumulating human capital, the payoff of doing education is also realized in the old period, instead of young.

[4]The term schooling agent refers to an young agent doing education. Among the students, those who survive into their old (only a fraction s of them survive) are called skilled labor (H) or scientists. Namely, being a skilled labor (or scientist) is contingent upon the success to survive into the old age.

[5]The idea of ensuring perfect annuity market against survival risks is similar to Kalemli-Ozcan et al. (2000), even though the form of such annuities is different from theirs, due to this different model specification.

Unskilled labor (denoted by L)

If a new born agent chooses to stay unskilled, his problem is expressed by

$$\max_{c_{Lt}^t, c_{Lt+1}^t} U_L^t = c_{Lt}^t \left(c_{Lt+1}^t\right)^{\delta s}, \tag{4.1}$$

$$s.t. \quad c_{Lt}^t + s_{Lt}^t = w_t, \tag{4.2}$$

$$c_{Lt+1}^t = \frac{(1+r_{t+1})s_{Lt}^t}{s}, \tag{4.3}$$

where U_L^t is the (discounted) lifetime utility of an unskilled labor born during period t, c_{Lt}^t the consumption when young (during period t), c_{Lt+1}^t the consumption when old (during period $t+1$), s_{Lt}^t the saving (which goes to physical capital accumulation) when young, δ the pure time discount factor, s the survival rate into the old age, w_t the wage rate when young, r_{t+1} the pure market interest rate between t and $t+1$. The lifetime utility is expressed as a Cobb-Douglas function. Taking logarithmic operation of the utility function gives

$$log\left(U_L^t\right) = log\left(c_{Lt}^t\right) + \delta s \cdot log\left(c_{Lt+1}^t\right),$$

which shows clearly that the effective discount factor of utility from old consumption is δs, incorporating both time discounting and survival uncertainty.

It is assumed that the annuity market against survival risks of unskilled labor is perfect, hence the savings of those unskilled labor who fail to survive into the old age are evenly distributed to those unskilled labor who indeed survive (which allows to use the representative agent method within each generation). This mechanism is reflected by the second constraint equation above. To understand it, suppose that a continuum of size 1 of unskilled labor save s_{Lt}^t when young, due to the law of large numbers (which applies here due the continuum of unskilled labor), the size of survivors is s. Total saving plus interest earned at the pure market rate is equal to $(1+r_{t+1})s_{Lt}^t$, and this is divided evenly among those size of s survived unskilled labor, with each survivor getting an amount equal to $\frac{(1+r_{t+1})s_{Lt}^t}{s}$, which appears on the right hand side of equation (4.3).

Skilled labor (denoted by H)

If a young unskilled labor chooses to do education in order to become skilled labor, upon surviving into the old age, he will work as a scientist in the R&D sector. Taking this lifetime process into account, when young, he solves the problem expressed by

$$\max_{c_{Ht}^t, c_{Ht+1}^t} U_H^t = f \cdot c_{Ht}^t \left(c_{Ht+1}^t\right)^{\delta s}, \tag{4.4}$$

$$s.t. \quad c_{Ht}^t = b_t, \tag{4.5}$$

$$c_{Ht+1}^t + \frac{1+r_{t+1}}{s} b_t = \frac{\Pi_{t+1}}{H_{t+1}}, \tag{4.6}$$

where U_H^t is the discounted lifetime utility of a schooling agent, $f \in (0,1)$ an exogenous constant, c_{Ht}^t and c_{Ht+1}^t his consumption plans when young and old respectively, b_t his borrowing when young to finance his early consumption (recall that he has no income when young). Π_{t+1} is the total profit for R&D sector during time $t+1$, while H_{t+1} is the total number of (surviving) scientists during $t+1$. Scientists are assumed to fully cooperate in the R&D process and they share the R&D profit equally among themselves, hence $\frac{\Pi_{t+1}}{H_{t+1}}$ is the net profit to each scientist, namely his payoff during the old age. Equation (4.4) is quite similar to equation (4.1), both having the same Cobb-Douglas specification.

The term $f < 1$ in the discounted lifetime utility of a schooling agent plays an important role in this model. First of all, there is no monetary cost of education in the model. The assumption that $f < 1$ captures the idea that education is costly, with the welfare lower than if f is absent. In this sense, $f < 1$ (lifetime welfare is multiplied by a factor less than unitary) can be thought of a dis-utility effect. Education cost is modelled in terms of direct dis-utility instead of monetary cost in the constraints. This makes the model much more tractable in the following analysis (Acemoglu (2009) chapter 10 has a similar treatment). Conventionally it is thought that education cost is (only) in terms of forgone wage (see, for example, Ben-Porath (1967) and de la Croix and Licandro (1999)). However, for most students, doing education is a painful learning process and costly at a psychological level.[6] This psychological cost cannot be captured merely in a monetary sense. The dis-utility term $f \in (0,1)$ can include either monetary cost or psychological cost, or both and so it reflects reality better.

In this section, all agents are assumed to have the same ability to do education and become skilled. Technically, f is not agent specific in the baseline model. However, agents with higher abilities should incur less cost in education, so they should have a higher f than agents with

[6] From the free open online course, Positive Psychology at Harvard University (the most popular course at Harvard University), a recent study finds that about 47% of college students across U.S. have depression, to the point of not functioning.

lower abilities. In section 4.5, the baseline model is extended to incorporate heterogeneities in agents' innate abilities to become skilled labor. As shown in section 4.5, this modification does not change main conclusions, but it gives a better justification for the resulting the equilibrium.

In this model the borrowing b_t takes the form of dis-saving. Students must borrow to finance their consumption when young, and this will take away some amount of final output that could otherwise be invested for physical capital accumulation. Therefore, borrowing of the students when young is treated as dis-saving in terms of physical capital. Physical capital accumulation will be discouraged by borrowing from students. This is reflected in the market clearing condition for physical capital presented below in equation (4.14).

4.2.2 Final production sector

In this model there is a final production sector. The final production sector requires unskilled labor (from young agents who decide to stay unskilled), research machines (from the R&D sector) and physical capital.[7]

The final good for consumption is produced using unskilled labor, physical capital and research machine, according to

$$Y_t = (L_t)^{\alpha}(K_t)^{\beta}(q_t)^{\theta}(x_t)^{1-\alpha-\beta}, \qquad \theta, \alpha, \beta, (1-\alpha-\beta) \in (0,1), \tag{4.7}$$

where L_t is the amount of unskilled labor, K_t the amount of physical capital, and x_t the machine used, q_t is the quality of machine. Quality q_t enters the final production function in a multiplicative fashion and serves as a productivity augmenting term. The final producer hires unskilled labor and rents physical capital from unskilled labor (each period it is the unskilled labor who save so they own the physical capital and can rent it out to the final producer). It rents research machines from the R&D firm (owned by scientists).[8]

[7] This modeling strategy is common in endogenous growth literature, as summarized in Acemoglu (2009), chapter 14,

[8] As the final production function illustrates, *research machine* is just the stuff the final producer rents from the R&D sector, and any names can be used for machine (such as *thing*, *stuff* etc.). Using machines from R&D sector in the final production is standard in modern endogenous growth literature (see, for example, Acemoglu (2009)). One reason is it simplifies significantly how scientists make profit. When a scientist discovers a knowledge, he can charge a price from those using the knowledge and at the same, the user(s) of the knowledge must be willing to pay that price. The willingness to pay can be very easily solved if the knowledge is treated

To simplify the analysis, research machines and physical capital are both assumed to fully depreciate after use, which implies that $1 + r_{t+1} = R_{t+1}$, where R_t denotes the rental rate of physical capital during period t. Moreover, the final production sector is assumed perfectly competitive, with the final producer taking all prices as given (input and output prices). The final production has constant returns to scale in inputs (labor L, capital K and research machine x), and the final production sector makes zero profit and the ownership of the final production firm is irrelevant.

Following the tradition in dynamic macroeconomic models under general equilibrium, the price of the final good is set to unitary in all periods. During any given period t, the final producer solves the following problem

$$\max_{L_t, K_t, x_t} Y_t - w_t L_t - R_t K_t - p_t^x x_t, \tag{4.8}$$

where w_t is the wage rate, R_t the rental rate of physical capital, and p_t^x is the price of research machines.[9]

4.2.3 R&D sector

This model has the feature of endogenous technological progress and how R&D takes place is explicitly modelled. This model builds on an early seminal work by Aghion and Howitt (1992) and the recent works of Acemoglu (2009) and Acemoglu et al. (2011). In the models of Aghion and Howitt (1992), Acemoglu (2009) and Acemoglu et al. (2011), technology is incorporated in the research machines produced by the R&D sector. Research inputs (either scientists or final output or both) are used to increase the machine quality. The outcome of any single scientist is stochastic. The simplest formulation is that the research result is either success or failure (binary), and each scientist has some probability to succeed or fail. To have deterministic research outcome on the overall level, as in Acemoglu (2009) and Acemoglu et al. (2001), they employ the law of large numbers, by assuming all the scientists are on a continuum. This model also needs the overall research outcome to be deterministic, and to achieve this, this model simply assumes the research outcome of each scientist is certain and

as or embodied in a production factor, whose price can be obtained just from a first order condition. Therefore, knowledge is embodied in the factor called *research machine*. Research machine is owned by scientists in a monopoly manner and is an input in the final production sector.

[9] The price of research machine is denoted by p_t^x instead of p_t, and the superscript x is used to emphasize this is the price for the research machine (research machine is denoted x_t).

with no risk of failure (so their payoff from doing research is also certain). This assumption achieves essentially the same goal (certainty on the overall research outcome) as Acemoglu (2009) and Acemoglu et al. (2001), and it simplifies analyzing the model. Another difference of this model from Aghion and Howitt (1992), Acemoglu (2009) and Acemoglu et al. (2011) is that, in this model agents endogenously choose to be unskilled or skilled and so the size of scientists is endogenous. In contrast, the literature most often works with an exogenously given constant size of scientists and so the formation of scientists in the first place is unclear.[10]

The students born during t, upon survival, will do research at the beginning of period $t+1$. They work on the machine with quality q_t and instantaneously improve the quality to q_{t+1} (hence machines with quality q_{t+1} are readily available at the beginning of period $t+1$). The scientists have monopoly power over the machines with quality q_{t+1} and rent them out to the final production sector during period $t+1$. This process is creatively destructive, in that the machines with quality lower than q_{t+1} will not be used during $t+1$.[11] Each scientist gets payoff from the final producer during time $t+1$, and he pays back his previous borrowing (to finance his consumption at youth) and consumes the amount left.

Scientists are assumed to fully cooperate during each time period and they share the total R&D profits equally among them. This explains why in equation (4.6), the total R&D profit Π_{t+1} is divided by the total number of scientists, H_{t+1}.

The evolution of machine qualities is assumed to take the following form

$$q_{t+1} = g(q_t) \cdot h(H_{t+1}). \tag{4.9}$$

The only restriction imposed on the functions $g(\cdot)$ and $h(\cdot)$ is

$$g(q) \cdot h(H) \geq q, \qquad \text{for all } q \text{ and } H \geq 0, \tag{4.10}$$

which says machine qualities can never decrease over time.

The input to the R&D sector is the final output, whereas the production of one unit of machine during t with quality q_t requires MC_t units of final output. The marginal cost can be written as $MC_t(q_t)$ to emphasize the fact that it depends upon q_t. However, since only

[10]One exception is Acemoglu (1998), who briefly suggests a method to endogenize the number of scientist.

[11]The inventors of machine with quality lower than q_{t+1} is dead during $t+1$. Since out-dated machines are not used anymore, no concern is given to who gets profits from those low-quality machines. This presents a perfect matching: those one-period active scientists have only one-period monopoly power over the machines they invent.

machines with quality q_t are used during t, q_t is suppressed for simplicity. Skilled labor, the monopolistic owner of research machine in the R&D sector, solves the following problem at time $t+1$

$$\max_{x_{t+1}} \quad \Pi_{t+1} \equiv \left(p^x_{t+1}(x_{t+1}) - MC_{t+1}\right) \cdot x_{t+1}. \tag{4.11}$$

where $p^x_{t+1}(x_{t+1})$ is the price of research machines during $t+1$. Skilled labor have some monopoly powers over research machines and so they will not take the machine price as given. They will affect the machine price by choosing quantities to supply and this explains the dependence of p^x_{t+1} upon x_{t+1}.

In terms of the marginal cost specification, it is reasonable to assume that the marginal cost is increasing in the machine quality, since more advanced machines are usually more expensive to build. This gives the final producer a trade-off between higher quality and higher cost, namely he could prefer using low quality machines because high quality machines are too expensive. While most often costs are not high enough to offset positive effects from higher quality, sometimes it is indeed the case.[12] In this chapter, for simplicity, these two opposing forces (high costs associated with high quality) are assumed to exactly cancel each other, in that the quality of research machines does not contribute directly into final output growth (see later equation (4.43)). Another reason for choosing this specification, as mentioned in the introduction, is to contrast this model with those in the literature. Studies such as Ludwig and Vogel (2010), de la Croix and Licandro (1999), Hu (1999) and Kalemli-Ozcan et al. (2000), utilizing endogenous growth models, find that population aging indirectly increases economic growth, via its direct positive effect on educational effort and human capital accumulation encourages economic growth. This model here essentially removes the positive impacts of human capital accumulation on output growth, and it will be shown below that population aging can still increase output growth. This provides another resolution to the conflicts between early simulation studies such as Auerbach and Kotlikoff (1987) and Hviding and Merette (1999), and the notable empirical study by Barro and Sala-I-Martin (1995).

The marginal cost of research machine with quality q_t is assumed to take the following form

$$MC_t = \psi q_t^{\overline{\frac{\theta}{1-\alpha-\beta}}}, \qquad \psi > 0. \tag{4.12}$$

[12] One such example can be found in computer science industry in the military service. There are some highly advanced computers, much more powerful than the standard, but too expensive for daily home and office usage.

The specification (4.12) ensures this model has a neo-classical nature. More details are discussed in 4.3.2, after the capital accumulation equation (4.35) is derived.

4.2.4 Market clearing

There are market clearing conditions for five variables and they are unskilled labor (L), skilled labor (H), research machines (x), final output (Y) and physical capital (K). For the markets of H, L and x, the 'demand=supply' conditions are trivially satisfied via appropriate usage of notation. During each period t, the final output is used in six ways and they are consumption by the students born during t, consumption by the unskilled labor born during t, consumption by the unskilled labor born during $t-1$ and surviving to t, consumption by the skilled labor during t (who are those students born during $t-1$ and surviving to t), production of research machines in the R&D sector and saving by the unskilled labor born at t. This is reflected by the following equation

$$Y_t = c_{Ht}^t \frac{H_{t+1}}{s} + c_{Lt}^t L_t + c_{Lt+1}^t s L_{t-1} + c_{Ht-1}^t H_t + x_t MC_t + s_{Lt}^t L_t. \tag{4.13}$$

For the physical capital accumulation from t to $t+1$, there are savings from the unskilled labor born at t, as well as dis-savings (borrowings) from the students born at t. This is reflected by the following equation (recall that is full depreciation in this model)

$$K_{t+1} = s_{Lt}^t L_t - b_t \frac{H_{t+1}}{s}, \tag{4.14}$$

which is also the law of motion for physical capital. Since students at t become skilled labor only after surviving into period $t+1$, the subscript for H is $t+1$ rather than t. Moreover, since only a fraction s of students at t can survive into period $t+1$, $\frac{H_{t+1}}{s}$ gives the amount of agents doing education (and borrow) at t.

4.3 Steady state equilibrium

4.3.1 Model solution details

In this section the model is solved for each period. During period t, optimization by unskilled labor (expressed by equations from (4.1) to (4.3)) gives the following Euler equation

$$c_{Lt+1}^t = \delta R_{t+1} c_{Lt}^t, \tag{4.15}$$

where the fact $1 + r_{t+1} = R_{t+1}$ (due to the full depreciation assumption) is used. This characterizes the inter-temporal decision of unskilled labor. The Euler equation, together with the two constraint equations (4.2) and (4.3), yields the consumption and saving solutions as

$$c^t_{Lt} = \frac{1}{1+\delta s} w_t, \tag{4.16}$$

$$s^t_{Lt} = \frac{\delta s}{1+\delta s} w_t, \tag{4.17}$$

$$c^t_{Lt+1} = \frac{(1+r_{t+1})\,\delta}{1+\delta s} w_t. \tag{4.18}$$

Substituting consumption plans for the young and old periods into the utility function, the maximized discounted lifetime welfare of an unskilled labor born at time t is given by

$$\omega^t_L = (w_t)^{1+\delta s} (R_{t+1})^{\delta s} \delta^{\delta s} (1+\delta s)^{-(1+\delta s)}. \tag{4.19}$$

One thing to note (for the analysis of steady-state welfare later) is that the lifetime welfare is positively related to both the wage rate and rental rate. This is quite intuitive: a higher wage rate means more resources to allocate during young, and a higher rental rate gives higher return for any given saving.

Having solved the optimization problem of unskilled labor, the problem for skilled labor is solved similarly. Using equations from (4.4) to (4.6), the optimization problem of a schooling agent born at t is solved, which yields

$$c^t_{Ht+1} = \delta R_{t+1} c^t_{Ht}, \tag{4.20}$$

$$c^t_{Ht} = b_t = \frac{s\Pi_{t+1}}{(1+\delta s) R_{t+1} H_{t+1}}, \tag{4.21}$$

$$c^t_{Ht+1} = \frac{\delta s \Pi_{t+1}}{(1+\delta s) H_{t+1}}. \tag{4.22}$$

Substituting the consumption plans for the young and old into the utility function, the maximized discounted lifetime utility of a schooling agent born at time t is given by

$$\omega^t_H = f \cdot \left(\frac{\Pi_{t+1}}{H_{t+1}}\right)^{1+\delta s} \frac{s}{1+\delta s} \frac{1}{R_{t+1}} \left(\frac{\delta s}{1+\delta s}\right)^{\delta s}. \tag{4.23}$$

The welfare of students, equation (4.23), is negatively related with the rental rate, so it is also negatively related with the interest. The reason is that students do not get a payoff from saving during young age and they only get a payoff from the R&D sector during the old age, so a higher interest rate just means a lower present discounted value of their old-age payoff, hence a lower discounted lifetime welfare.

After solving the problem of household, the final producer's problem, expressed by equation (4.8), is solved. This optimization with respect to unskilled labor, capital and research machines gives the functions for wage rate, rental rate and machine price as below

$$w_t = \alpha(L_t)^{\alpha-1}(K_t)^{\beta}(q_t)^{\theta}(x_t)^{1-\alpha-\beta}, \tag{4.24}$$

$$R_t = \beta(L_t)^{\alpha}(K_t)^{\beta-1}(q_t)^{\theta}(x_t)^{1-\alpha-\beta}, \tag{4.25}$$

$$p_t^x = (1-\alpha-\beta)(L_t)^{\alpha}(K_t)^{\beta}(q_t)^{\theta}(x_t)^{-\alpha-\beta}. \tag{4.26}$$

where w_t is the wage rate of unskilled labor, R_t is the rental rate of capital and p_t^x is the price of research machine during period t.

After the final producer's problem, the problem for the R&D sector can be solved. Equation (4.26) tells how the machine price depends on the amount of research machine (provided by the R&D firm). Anticipating the dependence of machine price on the quantity supplied, the skilled labor in R&D sector have monopoly power on research machines. They will consider the quantity supplied when solving problem (4.11), which then becomes

$$\max_{x_{t+1}} \Big((1-\alpha-\beta)(L_{t+1})^{\alpha}(K_{t+1})^{\beta}(q_{t+1})^{\theta}(x_{t+1})^{-\alpha-\beta} - MC_{t+1}\Big)\, x_{t+1}. \tag{4.26}$$

This optimization in turn yields the demand for machines expressed as

$$x_{t+1} = \left(\frac{(1-\alpha-\beta)^2(L_{t+1})^{\alpha}(K_{t+1})^{\beta}(q_{t+1})^{\theta}}{MC_{t+1}}\right)^{\frac{1}{\alpha+\beta}}. \tag{4.28}$$

Substituting equation (4.28) into equation (4.26), the price of research machines can be expressed as

$$p_{t+1}^x = \frac{MC_{t+1}}{1-\alpha-\beta}. \tag{4.29}$$

The price is a constant markup of the marginal cost, which follows naturally from the fact that the demand for research machine is iso-elastic, as expressed by equation (4.28). Substituting equations (4.28) and (4.29) into equation (4.11), the maximized total profit of the R&D sector during period t can be expressed as

$$\Pi_{t+1} = (\alpha+\beta)(1-\alpha-\beta)^{\frac{2-\alpha-\beta}{\alpha+\beta}}(L_{t+1})^{\frac{\alpha}{\alpha+\beta}}(K_{t+1})^{\frac{\beta}{\alpha+\beta}}(q_{t+1})^{\frac{\theta}{\alpha+\beta}}(MC_{t+1})^{\frac{\alpha+\beta-1}{\alpha+\beta}}. \tag{4.30}$$

The R&D profit is positively related with the machine quality and marginal cost. This is intuitive, as machines with higher quality receive a higher demand (as can be seen from equation (4.28)), and the R&D sector needs a higher profit to cover the machine production with higher marginal costs.

Each agent born at time t chooses to do education or stay unskilled. Under perfect foresight, his career decision must yield the highest discounted lifetime welfare. When there are both students and unskilled labor during each period,[13] an unskilled labor and a schooling agent must achieve the same discounted lifetime welfare, namely

$$\omega_L^t = \omega_H^t \quad \text{for each period } t. \tag{4.31}$$

Substituting the lifetime welfare of unskilled and skilled labor, expressed by equations (4.19) and (4.23), into the condition (4.31) gives

$$f^{\frac{1}{1+\delta s}} \frac{\Pi_{t+1}}{H_{t+1}} s = w_t R_{t+1}, \tag{4.32}$$

which is a condition required for the equality between skilled and unskilled labor's welfare.

To solve the model, another condition needs to be satisfied, which is the law of motion for physical capital. From the law of motion for physical capital (4.14), and denoting the size of all new born agents at t as l_t (l_t contains both unskilled labor and students at time t), it follows that

$$\begin{aligned} K_{t+1} &= s_{Lt}^t L_t - b_t \frac{H_{t+1}}{s} \\ &= \frac{\delta s}{1+\delta s} w_t L_t - \frac{1}{1+\delta s} w_t f^{-\frac{1}{1+\delta s}} (l_t - L_t) \\ &= \left(\frac{\delta s}{1+\delta s} L_t - \frac{1}{1+\delta s} f^{-\frac{1}{1+\delta s}} (l_t - L_t) \right) \alpha (1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}} L_t^{-\frac{\beta}{\alpha+\beta}} K_t^{\frac{\beta}{\alpha+\beta}} q_t^{\frac{\theta}{\alpha+\beta}} MC_t^{\frac{\alpha+\beta-1}{\alpha+\beta}}. \end{aligned} \tag{4.33}$$

This subsection derives the model solution details and the next subsection will solve for the steady state equilibrium, which is the key focus for analysis of population aging impacts.

4.3.2 Deriving the steady state solutions

In the steady state equilibrium, this chapter focuses on the following variables, which are

1. relative fraction of students versus unskilled labor, $\frac{s_H}{s_L}$;

[13] In this model, there are both unskilled labor and students during each period. Suppose not, that during some t there is no L, then the production of final good is zero, which is not a possible equilibrium. Suppose on the other hand that there is no H during some t, then there is no research machine at $t+1$ so the final production is zero at $t+1$, again not an equilibrium.

2. physical capital per unskilled labor $\frac{K}{L}$ and physical capital per capita $\frac{K}{l}$;

3. output per unskilled labor $\frac{Y}{L}$ and output per capita $\frac{Y}{l}$;

4. common welfare of unskilled labor and students, ω.

To study the impacts of population aging, the above steady state variables are solved, then comparative statistic analysis is conducted with respect to the survival rate s. The ratio of students to unskilled labor, $\frac{s_H}{s_L}$, is a measure of educational effort. An increase in $\frac{s_H}{s_L}$ means more students per unit of unskilled labor, and this represents a higher overall educational effort.

In the steady state, the fraction of unskilled labor to new born, $\frac{L_t}{l_t}$, is constant.[14] Dropping time subscripts, $s_L \equiv \frac{L}{l}$ is used to denote the (steady-state) fraction of new born agents who decide to stay unskilled. Moreover, $s_H \equiv 1 - s_L$ is used to denote the (steady-state) fraction of students. Since $\frac{H_{t+1}}{s} = l_t - L_t = s_H l_t = \frac{s_H}{s_L} L_t$, equation (4.33) becomes

$$\begin{aligned} K_{t+1} &= \left(\frac{\delta s}{1+\delta s} L_t - (\frac{s_H}{s_L}) L_t \cdot \frac{1}{1+\delta s} \cdot f^{-\frac{1}{1+\delta s}} \right) \alpha(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}} L_t^{-\frac{\beta}{\alpha+\beta}} K_t^{\frac{\beta}{\alpha+\beta}} q_t^{\frac{\theta}{\alpha+\beta}} MC_t^{-\frac{1-\alpha-\beta}{\alpha+\beta}} \\ &= \left(\frac{\delta s}{1+\delta s} - (\frac{s_H}{s_L}) \frac{1}{1+\delta s} f^{-\frac{1}{1+\delta s}} \right) \alpha(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}} L_t^{\frac{\alpha}{\alpha+\beta}} K_t^{\frac{\beta}{\alpha+\beta}} q_t^{\frac{\theta}{\alpha+\beta}} MC_t^{-\frac{1-\alpha-\beta}{\alpha+\beta}}. \end{aligned} \tag{4.34}$$

Substituting the research machine marginal cost, expressed by (4.12), into (4.34) yields

$$K_{t+1} = \left(\frac{\delta s}{1+\delta s} - (\frac{s_H}{s_L}) \frac{1}{1+\delta s} f^{-\frac{1}{1+\delta s}} \right) \alpha(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}} L_t^{\frac{\alpha}{\alpha+\beta}} K_t^{\frac{\beta}{\alpha+\beta}} \psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}. \tag{4.35}$$

Now, the machine quality and marginal cost terms disappear (except a new constant term), which gives the model a resemblance with the models in neo-classical growth literature. During each period, physical capital is positively related with the capital in the previous period, not related with machine qualities. This is why this model uses the specification of MC in (4.12), eliminating the positive effect of higher machine quality upon output. From equation (4.35), it can be seen that machine quality is eliminated from the capital accumulation so higher machine quality does not contribute directly to more output or faster capital accumulation. In this model, even if economic growth is directly related with research machine qualities, population aging still has positive impacts upon economic growth, as shown later.

With the capital accumulation in (4.35), capital per unskilled labor can be solved by dividing both sides of equation (4.35) by L_{t+1}. Denote the population growth rate by n

[14] Such a steady state exists, since all equilibrium conditions hold if s_{Ht} and s_{Lt} are constant through time.

($\frac{l_{t+1}}{l_t} = 1+n$), capital accumulation per unskilled labor is

$$\begin{aligned}\frac{K_{t+1}}{L_{t+1}} \equiv k_{t+1} &= \left(\frac{\delta s}{1+\delta s} - (\frac{s_H}{s_L})\frac{1}{1+\delta s} f^{-\frac{1}{1+\delta s}}\right)\alpha(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}}\psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}\frac{L_t^{\frac{\alpha}{\alpha+\beta}}K_t^{\frac{\beta}{\alpha+\beta}}}{L_t^{\frac{\alpha}{\alpha+\beta}}L_t^{\frac{\beta}{\alpha+\beta}}}\frac{L_t}{L_{t+1}}\\ &= \left(\frac{\delta s}{1+\delta s} - (\frac{s_H}{s_L})\frac{1}{1+\delta s} f^{-\frac{1}{1+\delta s}}\right)\alpha(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}}\psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}\frac{L_t}{L_{t+1}}k_t^{\frac{\beta}{\alpha+\beta}}\\ &= \left(\frac{\delta s}{1+\delta s} - (\frac{s_H}{s_L})\frac{1}{1+\delta s} f^{-\frac{1}{1+\delta s}}\right)\alpha(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}}\psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}\frac{1}{1+n}k_t^{\frac{\beta}{\alpha+\beta}},\end{aligned} \tag{4.36}$$

where the fact that $\frac{L_t}{L_{t+1}} = \frac{s_L l_t}{s_L l_{t+1}} = \frac{l_t}{l_{t+1}} = \frac{1}{1+n}$ is used. Since the power for k_t, $\frac{\beta}{\alpha+\beta} < 1$, equation (4.36) implies that k_t will converge to a constant value (in the steady state), denoted by k. Putting both k_{t+1} and k_t as k in the above equation gives the following

$$k = \left(\frac{\left(\frac{\delta s}{1+\delta s} - (\frac{s_H}{s_L})\frac{1}{1+\delta s} f^{-\frac{1}{1+\delta s}}\right)\alpha(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}}\psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}}{1+n}\right)^{\frac{\alpha+\beta}{\alpha}}. \tag{4.37}$$

This gives the value of physical capital per unskilled labor in the steady state. This expression is not in reduced form yet, since the ratio of students to unskilled labor, namely $\frac{s_H}{s_L}$, is not a primitive but a derived value. $\frac{s_H}{s_L}$ will be solved below to get the reduced form of k.

In the steady state physical capital per unskilled labor, expressed by (4.38), the only endogenous variable is the fraction of students to unskilled labor, namely $\frac{s_H}{s_L}$, which a measure of education effort. To solve for $\frac{s_H}{s_L}$, the welfare equality condition, (4.32) is used. Substituting into (4.32) the wage rate (4.24), rental rate (4.35), R&D sector profit (4.30) and marginal cost (4.12), equation (4.32) gives

$$f^{\frac{1}{1+\delta s}} s\left(\frac{\delta s}{1+\delta s} + \frac{s_H}{s_L}\frac{-1}{1+\delta s} f^{\frac{-1}{1+\delta s}}\right) = \frac{s_H}{s_L}\frac{\beta}{(\alpha+\beta)(1-\alpha-\beta)}, \tag{4.38}$$

and equation (4.38) allows to solve the the ratio $\frac{s_H}{s_L}$ as a function of the survival rate s, expressed as

$$\frac{s_H}{s_L} = \frac{1-s_L}{s_L} = f^{\frac{1}{1+\delta s}}\frac{\delta s^2(\alpha+\beta)(1-\alpha-\beta)}{\beta(1+\delta s) + s(\alpha+\beta)(1-\alpha-\beta)}. \tag{4.39}$$

Equation (4.39) is very important since it expressed the relative employment bias in a reduced form. This model is particularly interested in the relationship between $\frac{s_H}{s_L}$ and the survival rate s. Later comparative static analysis will be done, with respect to the survival rate s.

Substituting the ratio of students to unskilled labor (4.37) into the steady state physical capital per unskilled labor (4.37), the reduced form of physical capital per unskilled labor is expressed as

$$k = \left(\frac{\delta s \beta \alpha (1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}} \psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}}{(1+n)\left[\beta(1+\delta s) + s(\alpha+\beta)(1-\alpha-\beta)\right]} \right)^{\frac{\alpha+\beta}{\alpha}}. \tag{4.40}$$

Having solved the physical capital per unskilled labor, physical capital per capita can also be solved. Note that in contrast to most neo-classical growth models, in this model, even within one generation, physical capital per unskilled labor is not the same as physical capital per capita. This is due to the fact that not every new born agent stays unskilled, and some become students. To solve for physical capital per capita, the ratio of unskilled labor, s_L is needed first. From the ratio $\frac{s_H}{s_L}$, since $s_H + s_L = 1$, s_L can be expressed as the following

$$s_L = \left(1 + f^{\frac{1}{1+\delta s}} \frac{\delta s^2(\alpha+\beta)(1-\alpha-\beta)}{\beta(1+\delta s) + s(\alpha+\beta)(1-\alpha-\beta)}\right)^{-1}. \tag{4.41}$$

The steady-state physical capital per capita within each generation, namely the steady-state value of $\frac{K_t}{l_t}$, denoted by $\frac{K}{l}$, can be solved as

$$\begin{aligned}\left(\frac{K}{l}\right) =& \frac{K}{L}\frac{L}{l} = k s_L \\ =& \left(\frac{\delta \alpha \beta (1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}} \psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}}{1+n} \right)^{\frac{\alpha+\beta}{\alpha}} \left[\beta(1+\delta s) + s(\alpha+\beta)(1-\alpha-\beta)\right]^{-\frac{\beta}{\alpha}} \\ & \times \frac{s^{\frac{\alpha+\beta}{\alpha}}}{\beta(1+\delta s) + s(\alpha+\beta)(1-\alpha-\beta) + f^{\frac{1}{1+\delta s}} \delta s^2 (\alpha+\beta)(1-\alpha-\beta)}.\end{aligned} \tag{4.42}$$

Having finished the part on physical capital, the next step is to investigate the final output. Similar to K_t and $\frac{K_t}{L_t}$, Y_t and $\frac{Y_t}{L_t}$ can be solved as functions of physical capital and physical capital per unskilled labor, which are

$$Y_t = (1-\alpha-\beta)^{\frac{2}{\alpha+\beta}-2} \psi^{\frac{\alpha+\beta-1}{\alpha+\beta}} L_t L_t^{-\frac{\beta}{\alpha+\beta}} K_t^{\frac{\beta}{\alpha+\beta}}, \tag{4.43}$$

$$y_t \equiv \frac{Y_t}{L_t} = (1-\alpha-\beta)^{\frac{2}{\alpha+\beta}-2} \psi^{\frac{\alpha+\beta-1}{\alpha+\beta}} k_t^{\frac{\beta}{\alpha+\beta}}. \tag{4.44}$$

Since this model has chosen the research machine marginal cost specification to offset the impacts of higher machine quality on output, it is no surprise that $\frac{Y_t}{L_t}$ is unrelated with machine quality q.

With the output per unskilled labor (4.44) and the fraction of unskilled labor s_L (4.41), output per capita can be solved as a function of $\frac{s_H}{s_L}$, which is

$$\begin{aligned}\left(\frac{Y}{l}\right) = s_L \cdot \left(\frac{Y}{L}\right) &= (\alpha\beta)^{\frac{\beta}{\alpha}}(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha}}\psi^{\frac{\alpha+\beta-1}{\alpha}}(1+n)^{-\frac{\beta}{\alpha}}\delta^{\frac{\beta}{\alpha}} \\ &\times \left(\frac{s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}\right)^{\frac{\beta}{\alpha}} \\ &\times \frac{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)+f^{\frac{1}{1+\delta s}}\delta s^2(\alpha+\beta)(1-\alpha-\beta)} \\ =&(\alpha\beta)^{\frac{\beta}{\alpha}}(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha}}\psi^{\frac{\alpha+\beta-1}{\alpha}}(1+n)^{-\frac{\beta}{\alpha}}\delta^{\frac{\beta}{\alpha}} \\ &\times \left(\frac{s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}\right)^{\frac{\beta}{\alpha}} \frac{1}{1+(s_H/s_L)^{-1}}. \end{aligned} \tag{4.45}$$

Having solved for capital accumulation, output, next it comes to the steady-state values for discounted lifetime welfare.Since $\omega_L^t = \omega_H^t$ in equilibrium, this common welfare is denoted ω_t. Substituting the wage rate (4.24), rental rate (4.25), machine quantity demanded (4.28) and marginal cost (4.12) into the lifetime welfare (4.19), the welfare is expressed as

$$\begin{aligned}\omega_t &= w_t^{1+\delta s}R_{t+1}^{\delta s}\delta^{\delta s}(1+\delta s)^{-(1+\delta s)} \\ &= \left(\alpha^{1+\delta s}(1-\alpha-\beta)^{(\frac{2}{\alpha+\beta}-2)(1+2\delta s)}\psi^{(\frac{\alpha+\beta-1}{\alpha+\beta})(1+2\delta s)}\beta^{\delta s}\right)\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}k_t^{(1+\delta)\frac{\beta}{\alpha+\beta}}k_{t+1}^{-\frac{\alpha}{\alpha+\beta}\delta s}. \end{aligned} \tag{4.46}$$

Evaluating k_t and k_{t+1} at their steady-state value (expressed by equation (4.40)), the steady-state welfare, denoted by ω, is expressed as

$$\begin{aligned}\omega = &\left(\alpha^{1+\delta s}(1-\alpha-\beta)^{(\frac{2}{\alpha+\beta}-2)(1+2\delta s)}\psi^{(\frac{\alpha+\beta-1}{\alpha+\beta})(1+2\delta s)}\beta^{\delta s}\right) \\ &\times \left(\frac{\delta\beta s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}\right)^{\frac{\beta(1+\delta s)-\alpha\delta s}{\alpha+\beta}}\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}. \end{aligned} \tag{4.47}$$

The ratio of students to unskilled labor, (4.39), physical capital accumulation (4.40) and (4.42), final output (4.44) and (4.45) and lifetime welfare (4.47) describe all important variables of interest in the steady state and they are the focus in the next section when analyzing the impacts of population aging. The exogenous variable of particular interest is s, the survival rate. Comparative statistic exercises are done below, via an increase in s, and this gives the impacts of population aging upon macroeconomic performances.

4.4 The effects of population aging

The above section solves the model in the steady state and this section investigates the impacts of population aging, for the steady state variables. Population aging is defined by an increase in the survival rate s. When young students are more likely to survive into the old age to get their payoff, they have more incentives to do education when young and this could increase the fraction of people doing education. A higher educational effort increases the rate of technological progress. When young unskilled labor have a higher survival probability, they are more willing to save when young and this could increase the total amount of capital. Due to the specification of marginal costs of research machines, this model is neo-classical in essence, which means a higher rate of technological progress does not mean a higher rate of economic growth. However, more accumulation of physical capital implies faster economic growth.

Below, the impacts of population aging upon the ratio of students (educational efforts), physical capital accumulation, output growth and steady-state welfare are analyzed. This is done via comparative statistic analysis, with respect to an increase in the survival rate s.

4.4.1 Ratio of students

The fraction of students versus unskilled labor is denoted by $\frac{s_H}{s_L}$, and it represents the overall educational efforts, with a higher $\frac{s_H}{s_L}$ meaning higher educational efforts. The key equation to analyze here is (4.38), which implicitly defines $\frac{s_H}{s_L}$ as a function of the survival rate s. Totally differentiating equation (4.38) with respect to s gives

$$\frac{\partial(f^{\frac{1}{1+\delta s}})}{\partial s}\left(\frac{\delta s}{1+\delta s}-\frac{s_H}{s_L}\frac{f^{-\frac{1}{1+\delta s}}}{1+\delta s}\right)+f^{\frac{1}{1+\delta s}}\frac{\partial(\frac{\delta s}{1+\delta s})}{\partial s}+f^{\frac{1}{1+\delta s}}\frac{s_H}{s_L}\frac{\partial(-\frac{1}{1+\delta s}f^{-\frac{1}{1+\delta s}})}{\partial s}$$
$$=\frac{\partial s_H/s_L}{\partial s}\left(\frac{\beta}{(\alpha+\beta)(1-\alpha-\beta)}+\frac{1}{1+\delta s}\right),$$

which can be further simplified to

$$\frac{\partial s_H/s_L}{\partial s}\left(\frac{\beta}{(\alpha+\beta)(1-\alpha-\beta)}+\frac{1}{1+\delta s}\right)$$
$$=\frac{s_H}{s_L}\frac{\delta}{(1+\delta s)^2}+f^{\frac{1}{1+\delta s}}\frac{\delta}{(1+\delta s)^2}\left(1+\frac{(-log(f))\delta s}{1+\delta s}\right).$$

In the above equation, $\left(\frac{\beta}{(\alpha+\beta)(1-\alpha-\beta)}+\frac{1}{1+\delta s}\right)>0$ and $\frac{s_H}{s_L}\frac{\delta}{(1+\delta s)^2}+f^{\frac{1}{1+\delta s}}\frac{\delta}{(1+\delta s)^2}\left(1+\frac{(-log(f))\delta s}{1+\delta s}\right)>$

0. In order for the above equation to hold, the following inequality must hold

$$\frac{\partial s_H/s_L}{\partial s} > 0. \tag{4.48}$$

Inequality (4.48) implies that, after population aging via a higher survival probability, during each period there will be relatively more agents choosing to do education (in order to become skilled labor) versus staying unskilled.

Intuitively, with a higher probability of surviving into the old age, the present value of future payoff for a scientist, $\frac{s}{R_{t+1}}\frac{\Pi_{t+1}}{H_{t+1}}$, is higher, other things fixed. This raises the effective payoff of doing research, hence the incentive to do education and become skilled labor is higher. Consequently, more agents choose to go to school instead of staying unskilled.[15]

4.4.2 Physical capital, final output and welfare

In this subsection the impacts of population aging upon the physical capital accumulation, final output and welfare are investigated.

In terms of physical capital, both physical capital per unskilled labor (k) and per capita ($\frac{K}{l}$) are analyzed. It follows immediately from equation (4.40) that k increases in s. Therefore, population aging increases steady-state physical capital per unskilled labor. After population aging, there will be more students versus unskilled labor, and so a higher physical capital per unskilled labor does not guarantee a higher physical capital per capita. To analyze $\left(\frac{K}{l}\right)$, equation (4.42) implies

$$\left(\frac{K}{l}\right) = \left(\frac{\delta\alpha\beta(1-\alpha-\beta)^{\frac{2(1-\alpha-\beta)}{\alpha+\beta}}\psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}}{1+n}\right)^{\frac{\alpha+\beta}{\alpha}}\left(\frac{s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}\right)^{\beta/\alpha}$$
$$\times\frac{s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)+f^{\frac{1}{1+\delta s}}\delta s^2(\alpha+\beta)(1-\alpha-\beta)},$$

[15]Besides the intuition presented here, there is another tempting but wrong reasoning. According to this reasoning, the return of being a scientist can be realized only if he survives into the old age, therefore students face the risk of dying before their payoff is realized. With a higher survival probability, they are less worried about dying during youth and this provides higher incentives for going to school. The problem with this reasoning is that, even if skilled labor get payoffs only in the old age, they can borrow to finance their consumption when young. If they cannot survive into the old age, they do not need to pay back the debt (this is also why other survived scientists must pay back the debt at a rate higher than the pure interest rate, to ensure zero profit, no loss in this case, for the company or agents willing to lend), hence the risk of early death should not be in their concern and should not affect their decision to go to school.

where the second line above is $\frac{\frac{s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}}{1+(s_H/s_L)^{-1}}$. Since s_H/s_L increases in s, so is $\left(\frac{K}{l}\right)$.

Having investigated the impacts of population aging upon physical capital accumulation, attention is drawn to final output. Analyzing the steady-state version (4.44), output per unskilled labor, it is clear that output is increasing in physical capital, which further implies $\frac{\partial y}{\partial s} > 0$ and this is because $\frac{\partial k}{\partial s} > 0$. From equation (4.45) the output per capita, output per capita $\left(\frac{Y}{l}\right)$ is increasing in s_H/s_L and because s_H/s_L increases in s, so does $\left(\frac{Y}{l}\right)$. The results concerning final output are as expected, due to the resemblance of this model with neoclassical growth models. Because population aging encourages physical capital accumulation, it should also raise output per unskilled labor and output per capita.

Next, the effect of population aging upon steady state welfare is studied. From the steady state welfare, equation (4.47),

$$\omega = \text{constant} \times \left(\frac{\delta\beta s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}\right)^{\frac{\beta(1+\delta s)-\alpha\delta s}{\alpha+\beta}} \delta^{\delta s}(1+\delta s)^{-(1+\delta s)}. \quad (4.49)$$

The presence of $\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}$ in (4.49) depends on the survival rate s. The expression $\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}$ exists here because of choice of Cobb-Douglas utility function, and this expression is not related to the consumption or saving decisions of households. In order to focus on the impacts of population aging via consumption and saving decisions, the expression $\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}$ can be removed by multiplying (4.49) by the reciprocal of $\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}$. The usage of this method is following Acemoglu (2009, chapter 10).

After multiplying (4.49) by the reciprocal of $\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}$, the effect of population aging upon welfare is fully determined by the expression $\left(\frac{\delta\beta s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}\right)^{\frac{\beta(1+\delta s)-\alpha\delta s}{\alpha+\beta}}$. The base, $\frac{\delta\beta s}{\beta(1+\delta s)+s(\alpha+\beta)(1-\alpha-\beta)}$, is increasing in s, so if the power is positive, ω is increasing in s. The power is positive if $\beta(1+\delta s) > \alpha\delta s \Longleftrightarrow \beta > \alpha\frac{\delta s}{1+\delta s}$. Since $\alpha\frac{\delta s}{1+\delta s}$ depends on s, a sufficient condition for $\beta > \alpha\frac{\delta s}{1+\delta s}$ is that $\beta > \alpha$. Roughly speaking, if β is large enough, $\frac{\partial \omega}{\partial s} > 0$.

To get some insights for this result, welfare is expressed as

$$\omega = \text{constant} \cdot w^{1+\delta s} R^{\delta s}, \quad (4.50)$$

where welfare ω increases in both wage rate w and rental rate R. Combing equations (4.24), (4.25), (4.28) and (4.12) and evaluating variables in the steady state, it follows that

$$w = \alpha(1-\alpha-\beta)^{\frac{2}{\alpha+\beta}-2} \psi^{\frac{\alpha+\beta-1}{\alpha+\beta}} k^{\frac{\beta}{\alpha+\beta}}, \quad (4.51)$$

$$R = \beta(1-\alpha-\beta)^{\frac{2}{\alpha+\beta}-2}\psi^{\frac{\alpha+\beta-1}{\alpha+\beta}}k^{-\frac{\alpha}{\alpha+\beta}}. \tag{4.52}$$

As already established, k increases in s, so an increase in s will raise w but lower R.[16] With a large β, there is less diminishing return to K in the final production function and so the decline in R due to larger s is smaller in magnitude, while the increase in w is larger in magnitude. The net result is the steady-state welfare is more likely to increase in s with a larger β.

The following proposition summarizes the impacts of population aging in this section.

Proposition 1 *There is an equilibrium where s_H, s_L, w, R, $\frac{K}{L}$, $\frac{Y}{L}$, $\omega_L = \omega_H$ all converge to their steady state values. Population aging, defined via an increase in s, has the following effects:*

1. $\frac{\partial s_H/s_L}{\partial s} > 0$. *After population aging, there will be more agents choosing to do education and become skilled labor. There are more skilled versus unskilled labor;*

2. *After population aging, wage rate w increases while rental rate R decreases;*

3. *Physical capital and output per unskilled labor and per capita within each generation – $\frac{K}{L}$, $\frac{Y}{L}$, $\frac{K}{l}$ and $\frac{Y}{l}$ – all increase in s;*

4. *The welfare of both a unskilled labor and a schooling agent, ω_L and ω_H, will increase in s if $\beta > \alpha$. Population aging will increase steady-state welfare of all agents if the diminishing return to K in the final production function is of small magnitude.*

Proposition 1 summarizes the impacts of population aging upon educational efforts and macroeconomic performances. After population aging, the present value of expected future R&D payoff is higher and this causes more agents to do education and become skilled labor.

There are three channels underlying the result that $\frac{\partial(K/l)}{\partial s} > 0$. After population aging,

1. each unskilled labor L saves (in terms of physical capital accumulation) $\frac{\delta s}{1+\delta s}w$. The propensity to save, $\frac{\delta s}{1+\delta s}$; and the amount saved, $\frac{\delta s}{1+\delta s}w = \text{constant} \cdot \frac{\delta s}{1+\delta s}k^{\frac{\beta}{\alpha+\beta}}$, also increases in s since k increases in s;

[16]This result is similar to Kalemli-Ozcan et al. (2000), who argues that interest rate varies positively with mortality, “as would be expected from the simple intuition that shorter lives lead to lower wealth accumulation” (page 11). Moreover, Fougere and Merette (1999) and Hviding and Merette (1998) also find that population aging tends to lower interest rate.

2. combing equations (4.21) and (4.32), each schooling agent's dis-saving is equal to $b_t = \frac{1}{1+\delta s}\left(f^{\frac{1}{1+\delta s}}\right)^{-1} w = \text{constant} \cdot \left(f^{\frac{1}{1+\delta s}}\right)^{-1} \frac{1}{1+\delta s} k^{\frac{\beta}{\alpha+\beta}}$. $\left(f^{\frac{1}{1+\delta s}}\right)^{-1}$ is decreasing in s while $k^{\frac{\beta}{\alpha+\beta}}$ is increasing in s, hence the borrowing of each schooling agent can increase or down after population aging;

3. since $\frac{\partial(s_H/s_L)}{\partial s} > 0$, there will be more students dis-saving and less unskilled labor saving.

While 3 and possibly 2 discourage accumulation of physical capital, 1 is an opposite force. The analytical result above shows that 1 dominates. This finding contributes to the literature on population aging, physical and human capital accumulation. In this model, the specification for marginal costs of research machine is chosen in a way that this model model resembles neo-classical growth models and research machine quality is removed from the derived final output. As a result, after population aging, even if having more scientists leads to faster evolution of research machine qualities, it will not directly contribute to output per capita. However, the analytical results above show that population aging will still contribute to physical capital accumulation, and this is through a higher amount of saving from each unskilled labor. This highlights the importance of population aging for physical capital accumulation. Because this model is neo-classical in nature, more physical capital per unskilled labor (capita) implies more output per unskilled labor (capita).

The intuition for the welfare effects of population aging is presented in the two paragraphs immediately above Proposition 1.

In the next section, an extension with ability heterogeneity to the baseline model is studied.

4.5 Extension with ability heterogeneity

4.5.1 Model variation

In the above analysis, all new-born agents are assumed to have the same innate ability to do education (and become scientists). Proposition 1 shows that there exists one equilibrium where in each period, there are both unskilled labor and students.

In this section the above analysis is extended by introducing agents' abilities heterogeneities in doing education. Under this extension, unskilled labor and students are differentiated by their abilities, where only agents with abilities higher than a certain level choose to do education and become scientists.

New-born agents' ability to do education is denoted by a. The range of ability a is normalized to $[0,1]$ and the probability distribution of a is assumed to be uniform on $[0,1]$. The factor f in equation (4.4) is now assumed to depend on a, namely $f(a)$, and equation (4.4) becomes

$$U_t^H(a) = f(a) \cdot c_{Ht}^t(a) \cdot \left(c_{Ht+1}^t(a)\right)^{\delta s}. \tag{4.53}$$

Education will be easier for agents with higher abilities (or abilities can be interpreted as intelligence), and so education is less painful to them. One simple way to model this is to assume that $f(a)$ is continuous and increasing in a. This implies that, other things (lifetime consumptions) equal, agents with higher a will have higher lifetime utilities. Since lifetime utilities depend on ability a, in equation (4.53) this is emphasized by using $U_t^H(a)$ instead of U_t^H.

Accordingly, $f(a)$ is assumed continuous in a and $f'(a) > 0$ for all $a \in [0,1]$. This is not restrictive since agents' abilities can be ranked in a monotonic manner so that agents with higher abilities a always have higher lifetime utilities (other things being equal). Moreover, this model uses the normalization that $f(1) = 1$. The fact that $f < 1$ represents the dis-utility effect from doing education. If $f(a = 1) = 1$, there is no dis-utility for agents with the top ability ($a = 1$). This normalization makes the model more tractable without affecting the qualitative results. Also, $f(a = 0) = 0$, namely, the dis-utility term for agents with the bottom ability is extreme (they suffer from schooling so much that they will simply get zero utility if they go to school). This assumption ensures that agents with ability a around 0 will never do education and there is always a strictly positive amount of unskilled labor during each period.

Similar to the analysis in section 4.3, optimization problems can be solved for both unskilled labor and students. For those unskilled labor, their optimization problem is still fully represented by equations (4.1) to (4.3) and the solution is unchanged, with their maximized lifetime utility still represented by equation (4.19).

If an agent with ability a chooses to do education, he will solve the problem of maximizing (4.53) subject to constraints (4.5) and (4.6). Similar to the analysis in section 4.3, his maximized lifetime utility, in analogy with equation (4.23), is

$$\omega_H^t(a) = f(a) \cdot \left(\frac{\Pi_{t+1}}{H_{t+1}}\right)^{1+\delta s} \frac{s}{1+\delta s} \frac{1}{R_{t+1}} \left(\frac{\delta s}{1+\delta s}\right)^{\delta s}. \tag{4.54}$$

Since $f(a)$ is assumed increasing in a, (4.54) implies that students with higher a will enjoy higher lifetime utilities than students with lower a. This is different from the baseline model where all agents have the same ability. When all agents have the same ability, all students achieve the same lifetime utilities. During time t, among the students, the lowest ability can be denoted a_t. Since lifetime utilities of students decrease in a, the lifetime utility of agents with ability a_t will be the smallest lifetime utility among all students.

In the equilibrium with both unskilled labor and students existing during each period, the following must be true

$$\omega_L^t = \omega_H^t(a_t). \tag{4.55}$$

During period t, if an agent with ability lower than a_t chooses to go to school, his lifetime utility will be less than $\omega_L^t = \omega_H^t(a_t)$, which means he is better-off if he chooses to stay unskilled. Consequently, agents with ability $a < a_t$ will stay as unskilled labor, while agents with ability $a > a_t$ will go to school. Agents with ability equal to a_t are indifferent between staying unskilled and going to school. Since the size of new born in each period is normalized to unitary, the amount of unskilled labor during t is simply $a_t = L_t$ and the amount of students is $1 - a_t$, while the size of skilled labor during $t+1$ is $s \cdot (1 - a_t) = H_{t+1}$ (only a fraction s can survive into the old stage).[17]

Using the fact that f depends on a and the fact that $s \cdot (1 - a_t) = H_{t+1}$, the condition that ensures unskilled and skilled labor achieve the same welfare, expressed by equation (4.32), becomes

$$(f(a))^{\frac{1}{1+\delta s}} \frac{\Pi_{t+1}}{(1 - a_t)} = w_t R_{t+1}. \tag{4.56}$$

Similar to the baseline model, attention is drawn to the steady state where the fractions of students and unskilled labor are all constant across time. In the steady state where a_t is constant across time, and it is denoted this value by $\bar{a}$. $\bar{a}$ depends on s (and other exogenous parameters as well).

The steady-state fraction of unskilled labor is $s_L = \bar{a}$ and the fraction of students is $s_H = (1 - \bar{a})$. The equation expressing $\frac{s_H}{s}{}_L$ as a function of the survival rate s, equation

[17]The random variable a is continuous on [0,1], so the mass of those agents with ability equal to a_t is equal to zero.

(4.38), can be modified to

$$f(\bar{a})^{\frac{1}{1+\delta s}} s\left(\frac{\delta s}{1+\delta s}+\frac{1-\bar{a}}{\bar{a}}\frac{-1}{1+\delta s}f(\bar{a})^{-\frac{1}{1+\delta s}}\right)=\frac{1-\bar{a}}{\bar{a}}\frac{\beta}{(\alpha+\beta)(1-\alpha-\beta)}. \tag{4.57}$$

Equation (4.57) is the key equation for analyzing the impacts of population aging below.

4.5.2 The impacts of population aging

Similar to the analysis in section 4.4, the steady state impacts of population aging (via a higher s) upon the ratio of $\frac{s_H}{s_L}$, capital accumulation, final output and welfare are analyzed.

Equation (4.57) implicitly defines $\bar{a}$ as a function of s. Totally differentiating it with respect to s and rearranging gives the following important expression

$$\frac{\delta-log(f(\bar{a}))}{(1+\delta s)^2}+\frac{1}{s}+\frac{\frac{\delta}{(1+\delta s)^2}+\frac{1-\bar{a}}{\bar{a}}\frac{\delta f(\bar{a})^{-\frac{1}{1+\delta s}}}{(1+\delta s)^2}+\frac{1-\bar{a}}{\bar{a}}\frac{\delta(-logf(\bar{a}))}{(1+\delta s)^3}}{\frac{\delta s}{1+\delta s}+\frac{1-a}{a}\frac{-1}{1+\delta s}f(a)^{-\frac{1}{1+\delta s}}}$$
$$=\bar{a}'(s)\left\{\frac{-f'(\bar{a})}{f(\bar{a})(1+\delta s)}+\frac{-\frac{f(\bar{a})^{-\frac{1}{1+\delta s}}}{\bar{a}^2(1+\delta s)}-\frac{(1-\bar{a})f'(\bar{a})}{\bar{a}(1+\delta s)^2 f(\bar{a})}}{\frac{\delta s}{1+\delta s}+\frac{1-\bar{a}}{\bar{a}}\frac{-1}{1+\delta s}f(\bar{a})^{-\frac{1}{1+\delta s}}}-\frac{1}{\bar{a}(1-\bar{a})}\right\}. \tag{4.58}$$

Since $f(0)=0$, agents with ability 0 will never go to school and there will always be a strictly positive amount of unskilled labor during each period, hence $\bar{a}>0$. Moreover, following the same logic as in section 4.3, during no period can there be zero amount of students, hence $\bar{a}<1$. In the open interval $a\in(0,1)$, since $f(a)\in(0,1)$, $logf(a)<0$, and this implies that the LHS of (4.58) is negative. Moreover, since $f'(a)>0$, all the terms in the bracket of the RHS of (4.58) are positive. For equation (4.58) to hold, it follows that $\bar{a}'(s)<0$. Hence, with population aging (via a higher s), there are more students $1-\bar{a}$ and less unskilled agents $\bar{a}$ during each period in the new steady state. The impacts of population aging upon educational efforts are similar to the baseline model in section 4.3, where a is assumed to be equal across all agents. Therefore, introducing agents' heterogeneities in abilities does not invalidate the earlier results concerning the impacts of population aging upon educational effort.

Next the impacts of population aging upon physical capital accumulation, final output and welfare are analyzed. The steady state physical capital per unskilled labor expressed by equation(4.40) still applies here, and (4.40) is not affected by $f(a)$ or a. Therefore, the steady-state ratio of physical capital per unskilled labor, namely $k\equiv\frac{K}{L}$, is unchanged with ability heterogeneity and the comparative statics analysis in section 4.4.2 still applies here, which means an increase in s will raise the steady-state value of k. Similar logic applies to $\frac{K}{l}$ (represented by (4.49)) and $y\equiv\frac{Y}{L}$ (represented by (4.44)).

In terms of output per capita, represented by (4.45), s_H/s_L is replaced by $\frac{1-\bar{a}}{\bar{a}}$. Since $\bar{a}$ is decreasing in s, $\frac{1-\bar{a}}{\bar{a}}$ is increasing in s, so $\frac{Y}{l}$ is still increasing in s, same as the results in section 4.4.

With ability heterogeneity, the impacts of aging upon $\frac{s_H}{s_L}$, k, $\frac{K}{l}$, y and $\frac{Y}{l}$ are summarized in the following proposition:

Proposition 2 *Under the baseline previous model, but assuming that agents' abilities to do education, represented by a, differ across agents and follow a uniform distribution on $[0,1]$, there exists an equilibrium where there is a value of ability level $\bar{a} \in (0,1)$ such that all new-born agents with abilities higher than $\bar{a}$ will go to school and all agents with abilities lower than a will choose to remain unskilled. In this equilibrium, s_H, s_L, $\frac{K}{L}$, $\frac{K}{l}$, $\frac{Y}{L}$ and $\frac{Y}{l}$ all converge to their steady-state values. Population aging defined as a higher s has the following impacts:*

1. $\frac{\partial \bar{a}}{\partial s} < 0$ *and* $\frac{\partial s_H/s_L}{\partial s} > 0$. *After population aging, there will be more agents choosing to do education and become skilled labor. Population aging increases overall educational efforts;*

2. *Physical capital and output per unskilled labor and per capita within each generation, namely $\frac{K}{L}$, $\frac{Y}{L}$, $\frac{K}{l}$ and $\frac{Y}{l}$, all increase after population aging.*

Comparing Proposition 2 with Proposition 1, it can be seen that introducing ability heterogeneities does not qualitatively change the impacts of population aging upon educational efforts, physical capital accumulation and final output.

Next attention is drawn to the impacts of population aging upon the steady-state welfare, and here it is more complicated than the baseline model without ability heterogeneity. From Proposition 2, the ability 'threshold', $\bar{a}$, decreases in s. After s increases, $\bar{a}$ will fall and these two $\bar{a}$'s are denoted by a_1 and a_2 where $a_2 < a_1$ and $a_1 \in (0,1)$ and $a_2 \in (0,1)$. Before s rises, the steady-state welfares of agents with ability a_1 and a_2 are denoted by ω_1 and ω_2 respectively. After s rises, they are denoted by ω_1' and ω_2' respectively. Figure 4.1 is a simple illustration, where after s rises, more agents will choose to go to school.

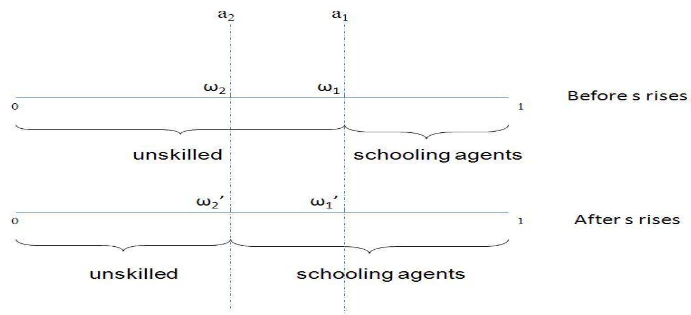


Figure 4.1. Career decisions before and after population aging under ability heterogeneity

The impacts of population aging upon steady state welfare are analyzed in two steps, first for agents with ability $a < a_2$ then for agents with ability $a > a_2$. These are done from the next paragraph. Before doing so, it is convenient for the analysis below to notice some relations between ω_1 and ω_2, and between ω_1' and ω_2'. Before s increases, in the steady state, all unskilled labor (those with $a < a_1$) achieve the same lifetime utility. Moreover, a_1 is the minimum ability for students and agents with a_1 are indifferent between going to school and remaining unskilled, which implies that $\omega_1 = \omega_2$. After s increases, all unskilled labor have the same welfare as ω_2' (welfare of the schooling agent with the lowest ability). Since $f(a)$ is increasing in a, students with higher a will enjoy higher welfare, and this means $\omega_2' < \omega_1'$.

Agents with $a < a_2$

To analyze the welfare effects of population aging, this model starts from ω_2 and ω_2'. ω_2 and ω_2' are equal to the welfare of unskilled labor before and after s rises, so they can be represented by equation (4.47) with different s. Proposition 1 implies that $\omega_2' \gtreqless \omega_2 \iff \beta \gtreqless \frac{\alpha\delta s}{1+\delta s}$. Therefore, if $\beta > \frac{\alpha\delta s}{1+\delta s}$, agents with ability $a < a_2$ are better off after population aging; if $\beta < \frac{\alpha\delta s}{1+\delta s}$, agents with ability $a < a_2$ are worse off after population aging; if $\beta = \frac{\alpha\delta s}{1+\delta s}$, agents with ability $a < a_2$ have the same welfare before and after population aging.

Agents with $a > a_2$

Next attention is drawn to the welfare change of agents with $a > a_2$. After s increases, all agents with $a < a_2$ are unskilled and they have the same welfare. After s increases, agents with $a > a_2$ are skilled and their welfare increase with a. The analysis is done under different possibilities.

If $\omega_2' = \omega_2$ (the case where $\beta = \frac{\alpha\delta s}{1+\delta s}$): for those with $a \in (a_2, a_1]$, they have welfare ω_2

before s rises, but higher than ω_2 after s rises (they are students after s rises, whose welfare increases with a). In particular, $\omega_1' > \omega_1 = \omega_2$. If an agent has ability $a > a_1$, his welfare is equal to $\frac{f(a)}{f(a_1)} \times$ *welfare of agents with ability* a_1. Since agents with ability a_1 enjoy higher welfare after s rises, all agents have higher welfare after s rises. Therefore, an increase in s benefits agents with ability $a > a_2$.

If $\omega_2' > \omega_2$ (the case where $\beta > \frac{\alpha\delta s}{1+\delta s}$): following a similar argument as above, the welfare of all agents with $a > a_1$ are higher after s rises. In this case, an increase in s benefits all agents with $a > a_2$in the economy.

If $\omega_2' < \omega_2$ (the case where $\beta < \frac{\alpha\delta s}{1+\delta s}$): if $\omega_1' < \omega_1$, all agents with $a > a_2$ have lower welfare after s rises. If $\omega_1' = \omega_1$, all agents with $a \geq a_1$ have the same welfare as before s rises, and all agents with $a \in (a_2, a_1)$ have lower welfare. If $\omega_1' > \omega_1$, all agents with $a \geq a_1$ have higher welfare, and since welfare of students is continuous in a, these must exist some ability $a^* \in (a_2, a_1)$ such that all students with $a > a^*$ are better off after s increases, while all students with $a < a^*$ are worse off after s increases and those with $a = a^*$ have the same welfare after s increases.

The above analysis about welfare change is summarized below.

Proposition 3 *Denote the lowest ability of students by a_1 before s rises and a_2 after s rises, population aging via an increase in s has the following welfare impacts:*

1. *If $\beta = \frac{\alpha\delta s}{1+\delta s}$, population aging increases the welfare of agents with $a > a_2$ and has no welfare impacts upon agents with $a < a_2$;*

2. *If $\beta > \frac{\alpha\delta s}{1+\delta s}$, population aging raises the welfare of all agents in the economy;*

3. *If $\beta < \frac{\alpha\delta s}{1+\delta s}$, all agents with $a \leq a_2$ are worse off after population aging. For the agents with $a > a_2$ (namely the students after population aging): (i) if $\omega_1' < \omega_1$, all agents have lower welfare after population aging; (ii) if $\omega_1' = \omega_1$, all agents with $a \geq a_1$ are not affected, while all agents with $a \in (a_2, a_1)$ have lower welfare; (iii) if $\omega_1' > \omega_1$, all agents with $a \geq a_1$ have higher welfare, and these exists some ability $a^* \in (a_2, a_1)$ such that all students with $a > a^*$ are better off after population aging, while all students with $a < a^*$ are worse off after population aging and those with $a = a^*$ are not affected.*

Comparing the above welfare results with sections 4.4, after introducing ability heterogeneities in education, an increase in s makes weakly more agents to gain. Intuitively, section

4.5 allows agents' welfare to increase with ability a, and so if unskilled labor achieve the same welfare before and after population aging, those skilled labor with higher a should enjoy higher lifetime welfare.

4.6 Conclusion

Population aging has become a world-wide phenomenon, characterized by a larger fraction of old people per unit of young agent. This chapter utilizes the discrete version of the overlapping generation model to investigate the impacts of population aging upon educational effort, physical capital accumulation, final output and welfare.

Some significant results are found. First, consistent with the empirical findings and similar to current literature, there is a positive relationship between survival rate and overall educational effort. After population aging, there are more agents choosing to go to school. As a result, there are more skilled labor doing research during each period. This is because a higher survival rate increases the payoff of doing education, which makes more people go to school.

Secondly, it is found that population aging encourages physical capital accumulation. After population aging, there are more students, and students must borrow to finance their youth consumption. Moreover, each student borrows more than before population aging. These two factors tend to lower physical capital accumulation. However, population aging increases the saving of each unskilled labor, and this factor dominates. Therefore, population aging encourages physical capital accumulation. At the same time, because this model has a strong resemblance to the neo-classical models, output growth is positively related with population aging through faster physical capital accumulation.

Thirdly, the impacts of population aging upon welfare are ambiguous in general. Population aging, by decreasing the amount of unskilled labor during each period, increases the wage rate while decreases the rental rate. These two income effects have opposite impacts upon welfare. This chapter presents necessary and sufficient conditions for either population aging is welfare-increasing or welfare-decreasing.

The one-sector model in this chapter can be treated as a natural benchmark for a further two-sector analysis. A two-sector model allows not only to study the impacts of population aging upon the total size of scientists versus unskilled labor, but also the relative amount of scientists in different research sectors, which is a key factor determining the direction of technical change. This problem is analyzed in the next chapter of the thesis.

Chapter 5

Population aging and directed technical change

5.1 Introduction

One of the most important demographic issues facing many countries nowadays is population aging. Numerous studies investigate the macroeconomic implications of demographic change. Among its various impacts, the connection between demography and technical change has drawn major attention. Prettner and Prskawetz (2010) provide a survey investigating the impacts of demographic change upon economic growth under various frameworks.

Chapter 4 investigated the impacts of population aging upon educational efforts and economic growth. In a one-sector growth model with endogenous technological progress and human capital investment through purposeful education, Chapter 4 finds that population aging tends to raise educational effort since an increase in longevity increases the effective benefit of schooling. In multi-sector growth models, interactions exist among different research sectors and this would potentially change the insight of Chapter 4. Therefore, the first question this chapter investigates is whether the finding of Chapter 4 can be generalized to two-sector growth models.

In this chapter's model, there is a unique final good produced using two intermediate goods. Two extreme cases will be investigated where the two intermediates are either perfect complements or perfect substitutes. Any other general degree of substitutability with general multiple sectors is just between these two extreme cases.

Most of the current literature finds a positive relationship between population aging and education effort, but there is a common limitation, that only the effect of aging upon overall technological progress is analyzed, while the relative performance of different research sectors is ignored. In most of the economic growth models, technologies of different sectors evolve in the same fashion (Romer, 1996; Aghion and Howitt, 1992; Jones, 1995 and Grossman and Helpman, 1991).

As pointed out by Acemoglu (1998), empirical studies find a significant bias of technical change in various sectors (for example, skill labor intensive versus unskilled labor intensive sectors), with more research and development (R&D) investment put in certain sectors. This triggered the directed technical change (DTC) literature, initiated by Acemoglu (1998, 2002, 2003, 2007, 2009), among others. The key mechanism underlying directed technical change is the relative profit of doing R&D in various sectors. R&D is conducted by profit-maximizing firms, who devote more research investment in the most profitable technology.

Among current literature, the first paper (and the only theoretical paper till now) studying population aging and directed technical change is Irmen (2009). Irmen models endogenous growth in a competitive framework, while in the DTC literature the research sector has some monopoly powers. This key difference makes comparisons between them difficult. Therefore, this chapter aims to investigate the impacts of population aging upon the direction of technical change, in a framework with some market imperfection and this is consistent the standard DTC literature. This gives the model a better resemblance with the DTC literature.

Even in this chapter's simple two-period overlapping-generation model, the algebra becomes very messy and the analytic solutions can be obtained only in two special cases, where the two intermediates are perfect complements, or perfect substitutes. With two intermediates being perfect complements, the price effect dominates the market size effect and this chapter's results are very similar to Irmen (2009). However, when two intermediates are perfect substitutes, there is no price effect and only the market size effect exists, and the results are exactly opposite to those of Irmen (2009).

The rest of the chapter is organized as follows. Section 5.2 describes the two-sector model (which is an extension of the model in Chapter 4 section 4.2), including the problems in the household, education and production sectors. Section 5.3 solves the optimization problems of all sectors and derives the steady state solutions. Using the steady state solutions, sections 5.4 and 5.5 study the impacts of population aging when two intermediate goods are perfect complements (section 5.4) or perfect substitutes (section 5.5). In the end, section 5.6 concludes the chapter with a summary of the main results and some future research ideas.

5.2 Two-sector growth model

5.2.1 Household sector

The model in this chapter is an extension of the model in Chapter 4. Chapter 4 utilizes a one-sector specification, where this chapter utilizes a two-sector growth model. The household sector is exactly the same as expressed in Chapter 4 section 4.2.

Unskilled labor

The problem of unskilled labor is the same as in Chapter 4 section 4.2.1, and equations (4.1) to (4.3) apply here, with the labels changed into (5.1)-(5.3).

$$\max_{c_{Lt}^t, c_{Lt+1}^t} U_L^t = c_{Lt}^t \left(c_{Lt+1}^t\right)^{\delta s}, \tag{5.1}$$

$$s.t. \quad c_{Lt}^t + s_{Lt}^t = w_t, \tag{5.2}$$

$$c_{Lt+1}^t = \frac{(1+r_{t+1})s_{Lt}^t}{s}. \tag{5.3}$$

Schooling agents

Same as in Chapter 4 section 4.2.1, if a young unskilled labor chooses to do education in order to become skilled labor, upon surviving into the old age, he will work as a scientist in the R&D sector. There are two R&D sectors, and they are denoted q_L and q_K sectors (full details are covered in subsections 5.2.2 and 5.2.3). A schooling agent, upon surviving into the old age, decides whether to work in the q_L sector or q_K sector. If he chooses to work in the q_L sector, he solves the problem expressed by

$$\max_{c_{Ht}^t(L), c_{Ht+1}^t(L)} U_H^t(L) = f \cdot c_{Ht}^t(L) \left(c_{Ht+1}^t(L)\right)^{\delta s}, \tag{5.4}$$

$$s.t. \quad c_{Ht}^t(L) = b_t(L), \tag{5.5}$$

$$c_{Ht+1}^t(L) + \frac{1+r_{t+1}}{s} b_t(L) = \frac{\Pi_{Ht+1}(L)}{H_{Lt+1}}, \tag{5.6}$$

where $c_{Ht}^t(L)$ is his consumption when young, $c_{Ht+1}^t(L)$ his consumption when old, $b_t(L)$ his borrowing to finance the consumption when young. Similar to the one-sector model in Chapter 4, $f \in (0,1)$ again captures the effect of dis-utility from schooling. $\frac{\Pi_{Ht+1}(L)}{H_{Lt+1}}$ denotes the payoff a scientist gets from doing research in the q_L sector (details about research payoffs are covered in subsection 5.2.3).

Similar to students working in the q_L sector, if a schooling agent chooses to work in the q_K sector, he solves the problem expressed by

$$\max_{c_{Ht}^t(K), c_{Ht+1}^t(K)} U_H^t(K) = f \cdot c_{Ht}^t(K) \left(c_{Ht+1}^t(K)\right)^{\delta s}, \tag{5.7}$$

$$s.t. \quad c_{Ht}^t(K) = b_t(K), \tag{5.8}$$

$$c_{Ht+1}^t(K) + \frac{1+r_{t+1}}{s} b_t(K) = \frac{\Pi_{Ht+1}(K)}{H_{Kt+1}}. \tag{5.9}$$

The explanations for notations in (5.7) to (5.9) are very similar to those in (5.4) to (5.6), with all K (for q_K sector) replaced by L (for q_L sector).

5.2.2 Final and intermediate goods

This model has a final production sector producing the consumption good. In contrast to the one-sector growth model in Chapter 4 section 4.2.2, the final good is produced using two intermediate goods. There are two research sectors, one for each intermediate good (details are presented below). One advantage of having two research sectors is that this allows for investigating the effects of population aging upon the *relative* amount of scientists across these two research sectors and so the direction of technical change.

The final good is produced using two intermediate goods. The final good production function is expressed by

$$Y_t = F(Y_{Kt}, Y_{Lt}), \tag{5.10}$$

where Y_K and Y_L denote two intermediates. $F(\cdot)$ features constant returns to scale. In the literature of directed technical change, the constant elasticity of substitution (CES) function is often used (Acemoglu, 2009), and the general form is the following

$$Y_t = \left((Y_{Lt})^{\frac{\varepsilon-1}{\varepsilon}} + (Y_{Kt})^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}} \quad \varepsilon \geq 0, \tag{5.11}$$

where ε is the elasticity of substitution between two intermediates. While $\varepsilon = 0$, Y_L and Y_K are perfect complements, and the production function becomes Leontief. With $\varepsilon = 1$, the production function is Cobb-Douglas. With $\varepsilon \to \infty$, Y_L and Y_K become perfect substitutes, and the production function becomes linear. An immediate consequence of this CES function is a relation between prices of two intermediates (details in the appendix 5.7.1) expressed as

$$\left((p_{Lt})^{1-\varepsilon} + (p_{Kt})^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} = 1. \tag{5.12}$$

Solving the model with a general value of ε turns out to be very difficult and in order to get reduced-form solutions, some restrictions on ε are needed. There are two extreme cases for ε, namely $\varepsilon = 0$ and $\varepsilon \to \infty$.

This chapter first analyzes the case where $\varepsilon = 0$ and it turns out that this case has a large resemblance with Irmen (2009). If $\varepsilon = 0$, the substitution between two intermediates are zero and this is the case where the intermediates are perfect complements, namely

$$Y_t = \min\{Y_{Lt}, Y_{Kt}\}. \tag{5.13}$$

Evaluating $\varepsilon = 0$ in equation (5.12) gives the following simple relation between the intermediate prices

$$p_{Lt} + p_{Kt} = 1. \tag{5.14}$$

Since the two intermediates are perfect complements, when they both have strictly positive prices, their quantities must satisfy the relation expressed by

$$Y_{Lt} = Y_{Kt}. \tag{5.15}$$

Productions of the two intermediate goods require labor, physical capital and research machines. Y_K and Y_L are assumed to differ in factor intensities and as denoted by the subscripts, Y_K is assumed to be relatively more K-intensive while Y_L is assumed to be relatively more L-intensive. As an extreme case for different factor intensity, this model assumes that the production of Y_K (Y_L) does not require any L (K). In analogy with the one-sector growth model in Chapter 4 of the thesis, Cobb-Douglas production functions are used for Y_K and Y_L, namely

$$Y_{Lt} = (L_t)^{\alpha}(q_{Lt})^{\theta_L}(x_{Lt})^{1-\alpha}, \quad \alpha \in (0,1), \theta_L > 0, \tag{5.16}$$

$$Y_{Kt} = (K_t)^{\beta}(q_{Kt})^{\theta_K}(x_{Kt})^{1-\beta}, \quad \beta \in (0,1), \theta_K > 0, \tag{5.17}$$

where x_{Lt} (x_{Kt}) denotes the research machine produced by q_L (q_K) sector, and q_{Lt} and q_{Kt} are qualities of the machines at t.

In this model the markets for the final good and two intermediates are assumed to be perfectly competitive. Due to constant returns to scale, the final production sector and the two intermediate sectors all make zero profits and so the ownerships of the final production and two intermediate sectors are all irrelevant. Setting the price of the final good to one in all periods, during each time period t the final producer solves the problem expressed by

$$\max_{Y_{Lt}, Y_{Kt}} \pi_t \equiv 1 \cdot Y_t - p_{Lt}Y_{Lt} - p_{Kt}Y_{Kt}, \tag{5.18}$$

where p_{Lt} (p_{Kt}) is the price of intermediate Y_L (Y_K) during time period t. The intermediates producers solve the problems expressed by

$$\max_{L_t, x_{Lt}} \pi_{Lt} \equiv p_{Lt}Y_{Lt} - w_tL_t - p^x_{Lt}x_{Lt}, \tag{5.19}$$

$$\max_{K_t, x_{Kt}} \pi_{Kt} \equiv p_{Kt}Y_{Kt} - R_tK_t - p^x_{Kt}x_{Kt}, \tag{5.20}$$

where p^x_{Lt} (p^x_{Kt}) is the price for machine x_L (x_K), w_t is the wage rate and R_t is the rental rate. Assuming physical capital K, research machines x_L and x_K all depreciate fully after use, it follows that $1 + r_t = R_t$.

5.2.3 R&D sector

In contrast to the one-sector model in Chapter 4 section 4.2.3 where there is only one R&D sector, in this two-sector model there are two R&D sectors, one for each intermediate good. For simplicity, they are named q_L sector (producing research machine x_L) and q_K sector (producing research machine x_K). The research result of q_L (q_K) sector increases the productivity of machine x_L (x_K). In the next paragraph, for expositional simplicity, only the q_L sector is described and the specification for q_K sector is identical with change of notations (from L to K).

If a schooling agent born during t survives into period $t + 1$, he chooses to work in the q_L or q_K sector (but not both at the same time). If he chooses to work in the q_L sector, he does research at the beginning of $t + 1$. Scientists in q_L sector work on the machine with quality q_{Lt} and instantaneously improve the quality to q_{Lt+1}. During time $t+1$ the scientists in q_L sector have monopoly power over the machine with quality q_{Lt+1} and rent it out to the intermediate Y_L sector. This process is creative destruction, and the machines with quality lower than q_{Lt+1} will not be used during $t+1$. Each scientist gets research payoff during $t+1$, pays back his previous borrowing (with interest) and consumes the amount left.

All scientists within q_L sector are assumed to fully cooperate during each time period, sharing the R&D profits equally among them. The payoff for each scientist in q_L sector at $t + 1$ is the total profits from q_L sector divided by total number of H in q_L. The notation H_{t+1} is used for the total number of scientists during period $t + 1$, H_{Lt+1} (H_{Kt+1}) for the number of scientists in q_L (q_K) sector during $t + 1$. All scientists are distributed between q_L and q_K sectors, so $H_{t+1} = H_{Lt+1} + H_{Kt+1}$. Denoting the total profits at q_L sector at time $t + 1$ by $\Pi_{Ht+1}(L)$, each scientist in q_L sector gets $\frac{\Pi_{Ht+1}(L)}{H_{Lt+1}}$.

The evolution of machine qualities $\{q_{Lt}\}_{t=0}^{\infty}$ is represented by

$$q_{Lt+1} = g_L(q_{Lt}) \cdot h_L(H_{Lt+1}). \tag{5.21}$$

In terms of the function forms of $g_L(\cdot)$ and $h_L(\cdot)$, it is assumed that

$$g_L(q_L) \cdot h_L(H_L) \geq q_L, \qquad \text{for all } q_L \text{ and } H_L \geq 0, \tag{5.22}$$

which says machine qualities $\{q_{Lt}\}_{t=0}^{\infty}$ can never decrease over time.

The input to R&D in q_L sector is the final output, whereas the production of one unit of machine at t with quality q_{Lt} requires MC_{Lt} units of final output (the form of MC_{Lt} is specified later).[1] During time period $t+1$, skilled labor in the q_L research sector solves the problem represented by

$$\max_{x_{Lt+1}} \quad \Pi_{Ht+1}(L) \equiv \left(p^x_{Lt+1}(x_{Lt+1}) - MC_{Lt+1}\right) \cdot x_{Lt+1}, \tag{5.23}$$

Skilled labor has monopoly powers over research machines, so they will take the dependence of price p^x_L upon quantity x_L into account. That is, they do not take price p^x_L as given. It is this monopoly power over the machines that gives skilled labor the initial motivation to do education.

As in Chapter 4, the following assumptions are made on the marginal costs of research machines (the counterpart of equation (4.12))

$$MC_{Lt} = \psi(q_{Lt})^{\frac{\theta_L}{1-\alpha}}, \qquad \psi > 0, \tag{5.24}$$

$$MC_{Kt} = \phi(q_{Kt})^{\frac{\theta_K}{1-\beta}}, \qquad \phi > 0, \tag{5.25}$$

which again captures the idea that machines with higher quality are more expensive (in terms of final output).

5.2.4 Market clearing

Similar to section 4.2.4, this subsection describes the market clearing conditions for the final output and physical capital accumulation.

During each period t, the final output is used in nine ways and they are saving by the unskilled labor born at t, production of research machines in the q_L sector, production of research machines in the q_K sector, consumption by the unskilled labor born during t, consumption by the students born during t who will work in the q_L sector during $t+1$, consumption by the students born during t who will work in the q_K sector during $t+1$, consumption by the unskilled labor born during $t-1$ and surviving to t, consumption by the skilled labor during t working in the q_L sector and consumption by the skilled labor during t working in the q_K

[1] The marginal cost can be written as $MC_{Lt}(q_{Lt})$ to emphasize the fact that marginal costs depend on q_{Lt}.

sector. This is reflected by the following equation

$$
\begin{aligned}
Y_t =& s_{Lt}^t L_t + MC_{Lt} x_{Lt} + MC_{Kt} x_{Kt} \\
&+ c_{Lt}^t L_t + c_{Ht}^t(L) \frac{H_{Lt+1}}{s} + c_{Ht}^t(K) \frac{H_{Kt+1}}{s} + c_{Lt}^{t-1} s L_{t-1} + c_{Ht}^{t-1}(L) H_{Lt} + c_{Ht}^{t-1}(K) H_{Kt}.
\end{aligned}
\tag{5.26}
$$

In terms of the market clearing for physical capital, the following equation holds

$$
K_{t+1} = s_{Lt}^t L_t - b_t(L) \frac{H_{Lt+1}}{s} - b_t(K) \frac{H_{Kt+1}}{s}, \tag{5.27}
$$

which is also the law of motion for physical capital. In each period, due to full depreciation, physical capital this period does not contribute to the physical capital next period. Within the young generation of each period, unskilled labor save (increasing physical capital of next period) while students borrow (increasing physical capital of next period). Similar to the one-sector growth model in Chapter 4 of the thesis, borrowing of the schooling agents ($b_t(L)$ and $b_t(K)$) is reflected as dis-saving of the physical capital.

5.3 Steady state equilibrium

5.3.1 Model solution details

Unskilled labor solve the problem represented by equations (5.1) to (5.3), and the solution is given by

$$
c_{Lt}^t = \frac{1}{1+\delta s} w_t, \tag{5.28}
$$

$$
s_{Lt}^t = \frac{\delta s}{1+\delta s} w_t, \tag{5.29}
$$

$$
c_{Lt+1}^t = \frac{(1+r_{t+1})\,\delta}{1+\delta s} w_t. \tag{5.30}
$$

Substituting the above consumption plans for the young and old periods into the utility function (5.1), the maximized discounted lifetime welfare of an unskilled labor born at time t is

$$
\omega_L^t = (w_t)^{1+\delta s} (R_{t+1})^{\delta s} \delta^{\delta s} (1+\delta s)^{-(1+\delta s)}. \tag{5.31}
$$

A schooling agent choosing to work in the q_L (q_K) sector solves the problem expressed by equations (5.4) to (5.6) ((5.7) to (5.9)), and the solution is given by

$$c^t_{Ht}(L) = b_t(L) = \frac{s}{(1+\delta s)R_{t+1}} \frac{\Pi_{Ht+1}(L)}{H_{Lt+1}}, \tag{5.32}$$

$$c^t_{Ht+1}(L) = \frac{\delta s}{1+\delta s} \frac{\Pi_{Ht+1}(L)}{H_{Lt+1}}, \tag{5.33}$$

$$c^t_{Ht}(K) = b_t(K) = \frac{s}{(1+\delta s)R_{t+1}} \frac{\Pi_{Ht+1}(K)}{H_{Kt+1}}, \tag{5.34}$$

$$c^t_{Ht+1}(K) = \frac{\delta s}{1+\delta s} \frac{\Pi_{Ht+1}(K)}{H_{Kt+1}}. \tag{5.35}$$

Substituting the above expressions into the lifetime welfare equations (5.4) and (5.7), the maximized lifetime welfare of a schooling agent working in the q_L or q_K sector is given by

$$\omega^t_H(L) = f \cdot s(\delta s)^{\delta s}(1+\delta s)^{-(1+\delta s)} \left(\frac{\Pi_{Ht+1}(L)}{H_{Lt+1}}\right)^{1+\delta s} (R_{t+1})^{-1}, \tag{5.36}$$

$$\omega^t_H(K) = f \cdot s(\delta s)^{\delta s}(1+\delta s)^{-(1+\delta s)} \left(\frac{\Pi_{Ht+1}(K)}{H_{Kt+1}}\right)^{1+\delta s} (R_{t+1})^{-1}. \tag{5.37}$$

Two intermediate sectors solve the problem expressed by equations (5.19) and (5.20), and the solution is given by

$$w_t = p_{Lt}\alpha(L_t)^{\alpha-1}(q_{Lt})^{\theta_L}(x_{Lt})^{1-\alpha}, \tag{5.38}$$

$$p^x_{Lt} = p_{Lt}(1-\alpha)(L_t)^{\alpha}(q_{Lt})^{\theta_L}(x_{Lt})^{-\alpha}, \tag{5.39}$$

$$R_t = p_{Kt}\beta(K_t)^{\beta-1}(q_{Kt})^{\theta_K}(x_{Kt})^{1-\beta}, \tag{5.40}$$

$$p^x_{Kt} = p_{Kt}(1-\beta)(K_t)^{\beta}(q_{Kt})^{\theta_K}(x_{Kt})^{-\beta}. \tag{5.41}$$

Substituting the wage rate (5.38) into the problem of q_L sector expressed by (5.23) with all terms evaluated at t, the solution of problem (5.23) is given by

$$x_{Lt} = \left((1-\alpha)^2 p_{Lt}(L_t)^{\alpha}(q_{Lt})^{\theta_L}(MC_{Lt})^{-1}\right)^{\frac{1}{\alpha}}, \tag{5.42}$$

$$\Pi_{Ht}(L) = \alpha(1-\alpha)^{\frac{2}{\alpha}-1}(p_{Lt})^{\frac{1}{\alpha}} L_t (q_{Lt})^{\frac{\theta_L}{\alpha}} (MC_{Lt})^{1-\frac{1}{\alpha}}, \tag{5.43}$$

while a similar argument in the q_K sector gives

$$x_{Kt} = \left((1-\beta)^2 p_{Kt}(K_t)^{\beta}(q_{Kt})^{\theta_K}(MC_{Kt})^{-1}\right)^{\frac{1}{\beta}}, \tag{5.44}$$

$$\Pi_{Ht}(K) = \beta(1-\beta)^{\frac{2}{\beta}-1}(p_{Kt})^{\frac{1}{\beta}} K_t (q_{Kt})^{\frac{\theta_K}{\beta}} (MC_{Kt})^{1-\frac{1}{\beta}}. \tag{5.45}$$

An agent born at time t chooses do education or stay unskilled. Moreover, if he goes to school, upon surviving into the old age, he chooses to work in the q_L or q_K sector. Under perfect foresight, he will try to obtain the highest discounted lifetime welfare. When there are unskilled labor as well as schooling agents aiming for both R&D sectors in each period, the unskilled labor, the schooling agents choosing to work in the q_L sector, and the school agents choosing to work in the q_K sector must have the same lifetime welfare, namely

$$\omega_L^t = \omega_H^t(K) = \omega_H^t(L). \tag{5.46}$$

The notation $s_{Lt} \equiv L_t/l_t$ is used for the share of new-born agents who stay unskilled and $s_{Ht} = 1 - s_{Lt}$ denotes the share of new-born agents going to school. Moreover, $s_{HKt} \equiv H_{Kt}/H_t$ denotes the share of schooling agents who choose to work in the q_K sector, and $s_{HLt} \equiv H_{Lt}/H_t = 1 - s_{HKt}$ denotes the share of schooling agents working in the q_L sector.

From $\omega_H^t(K) = \omega_H^t(L)$, and using equations (5.36) and (5.37), it follows that

$$\begin{aligned} &\frac{\alpha(1-\alpha)^{2/\alpha-1}(p_{Lt+1})^{1/\alpha}L_{t+1}(q_{t+1})^{\theta_L/\alpha}(MC_{Lt+1})^{1-1/\alpha}}{s_{HLt}} \\ &= \frac{\beta(1-\beta)^{2/\beta-1}(p_{Kt+1})^{1/\beta}K_{t+1}(q_{t+1})^{\theta_K/\beta}(MC_{Kt+1})^{1-1/\beta}}{s_{HKt}}. \end{aligned} \tag{5.47}$$

From $\omega_L^t = \omega_H^t(L)$, using two welfare equations (5.31) and (5.36), it follows that

$$w_t R_{t+1} = f^{\frac{1}{1+\delta s}} s \frac{\Pi_{Ht+1}(L)}{H_{Lt+1}}, \tag{5.48}$$

which can be thought of the two-sector model counterpart of the equation ensuring welfare equality (4.32) in Chapter 4 of the thesis. Using the equation (5.27) for physical capital accumulation, it follows that

$$\begin{aligned} K_{t+1} &= s_{Lt}^t L_t - b_t(L)\frac{H_{Lt+1}}{s} - b_t(K)\frac{H_{Kt+1}}{s} \\ &= s_{Lt}^t L_t - b_t(L) s_{HLt+1}\frac{s_{Ht}}{s_{Lt}} L_t - b_t(K) s_{HKt+1}\frac{s_{Ht}}{s_{Lt}} L_t. \end{aligned} \tag{5.49}$$

This chapter considers the steady state where $\{s_{Ht}\}_{t=0}^{\infty}$, $\{s_{Lt}\}_{t=0}^{\infty}$, $\{s_{HKt}\}_{t=0}^{\infty}$ and $\{s_{HLt}\}_{t=0}^{\infty}$ are all constant sequences. Substituting equations saving decisions (5.29), (5.32) and (5.34) into (5.49) gives

$$\frac{K_{t+1}}{L_{t+1}} = \frac{1}{1+n}\left(\frac{\delta s}{1+\delta s}w_t - s_{HL}\frac{s_H}{s_L}\frac{s}{(1+\delta s)R_{t+1}}\frac{\Pi_{Ht+1}(L)}{H_{Lt+1}} - s_{HK}\frac{s_H}{s_L}\frac{s}{(1+\delta s)R_{t+1}}\frac{\Pi_{Ht+1}(K)}{H_{Kt+1}}\right). \tag{5.50}$$

From equations (5.30), (5.45) and (5.48), it follows that

$$\frac{s\Pi_{Ht+1}(L)}{H_{Lt+1}R_{t+1}} = f^{-\frac{1}{1+\delta s}} w_t = \frac{s\Pi_{Ht+1}(K)}{H_{Kt+1}R_{t+1}}. \tag{5.51}$$

Substituting (5.51) into equation (5.50) gives

$$\begin{aligned}(1+n)\frac{K_{t+1}}{L_{t+1}} =&\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}(p_{Lt})^{1/\alpha}(q_{Lt})^{\theta/\alpha}(MC_{Lt})^{\frac{\alpha-1}{\alpha}} \\ &\times\left(\frac{\delta s}{1+\delta s} - s_{HL}\frac{s_H}{s_L}\frac{1}{1+\delta s}f^{-\frac{1}{1+\delta s}} - s_{HK}\frac{s_H}{s_L}\frac{1}{1+\delta s}f^{-\frac{1}{1+\delta s}}\right). \end{aligned} \tag{5.52}$$

Using research machine marginal cost (5.24), the capital accumulation (5.52) in the steady state becomes

$$\left(\frac{K_{t+1}}{L_{t+1}}\right)^* \equiv k^* = \frac{1}{1+n}\alpha(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}\psi^{\frac{\alpha-1}{\alpha}}(p_L)^{\frac{1}{\alpha}}\left(\frac{\delta s}{1+\delta s} - \frac{s_H}{s_L}\frac{1}{1+\delta s}f^{-\frac{1}{1+\delta s}}\right). \tag{5.53}$$

Using equations (5.25), (5.38), (5.40) and (5.48), the second part of equation (5.46) gives

$$f^{\frac{1}{1+\delta s}} = \left(\frac{\beta(1-\beta)^{\frac{2}{\beta}-2}\phi^{\frac{\beta-1}{\beta}}}{1-\alpha}\right)s_{HL}\frac{s_H}{s_L}\frac{1}{1+n}(p_K)^{\frac{1}{\beta}}. \tag{5.54}$$

From equation (5.47), it follows that

$$\frac{\alpha(1-\alpha)^{\frac{2}{\alpha}-1}(p_L)^{\frac{1}{\alpha}}L_t(q_{Lt})^{\frac{\theta_L}{\alpha}}(MC_{Lt})^{\frac{\alpha-1}{\alpha}}}{s_{HL}} = \frac{\beta(1-\beta)^{\frac{2}{\beta}-1}(p_K)^{\frac{1}{\beta}}K_t(q_{Kt})^{\frac{\theta_L}{\beta}}(MC_{Kt})^{\frac{\beta-1}{\beta}}}{s_{HK}}, \tag{5.55}$$

which gives the expression

$$(p_K)^{\frac{1}{\beta}} = \frac{(1-\alpha)(1+n)}{\beta(1-\beta)^{\frac{2}{\beta}-1}\phi^{\frac{\beta-1}{\beta}}}\frac{s_{HK}}{s_{HL}}\left(\frac{\delta s}{1+\delta s} - \frac{s_H}{s_L}\frac{f^{-\frac{1}{1+\delta s}}}{1+\delta s}\right)^{-1}. \tag{5.56}$$

Substituting equation (5.56) into equation (5.54) gives the following equation

$$f^{\frac{1}{1+\delta s}} = \frac{1}{1-\beta}s_{HK}\frac{s_H}{s_L}\left(\frac{\delta s}{1+\delta s} - \frac{s_H}{s_L}\frac{f^{-\frac{1}{1+\delta s}}}{1+\delta s}\right)^{-1}. \tag{5.57}$$

To solve for the prices for intermediates, putting (5.42) and (5.44) into (5.16) and (5.17), which gives

$$Y_{Lt} =(p_{Lt})^{\frac{1-\alpha}{\alpha}}(1-\alpha)^{\frac{2(1-\alpha)}{\alpha}}L_t\psi^{\frac{\alpha-1}{\alpha}}, \tag{5.58}$$

$$Y_{Kt} =(p_{Kt})^{\frac{1-\beta}{\beta}}(1-\beta)^{\frac{2(1-\beta)}{\beta}}K_t\phi^{\frac{\beta-1}{\beta}}. \tag{5.59}$$

5.3.2 Steady state equilibrium

In the following analysis, attention is drawn to the steady state where s_{HKt}, s_{HLt}, s_{Ht} and s_{Lt} are all constant through time. Consequently, the t subscripts are dropped hereafter.

Combining equations (5.14), (5.15), (5.54), (5.58) and (5.59), the price for intermediate Y_L can be expressed as

$$p_L = (1-\alpha)^{\beta-1}(1+n)^{\beta}\alpha^{-1}\beta^{1-\beta}(1-\beta)^{\beta-1}\phi^{1-\beta}\left(\frac{s_{HK}}{s_{HL}}\right)^{\beta-1}\left(\frac{\delta s}{1+\delta s}-\frac{s_H}{s_L}\frac{f^{-\frac{1}{1+\delta s}}}{1+\delta s}\right)^{-\beta}. \tag{5.60}$$

Equation (5.60) follows after the assumption $\varepsilon = 0$ and it does not apply for a general ε (not applicable later when the case with $\varepsilon \to \infty$ is analyzed later). Equation (5.60), together with equations (5.56) and (5.14), yield

$$(1-\alpha)^{\beta}(1+n)^{\beta}\beta^{-\beta}(1-\beta)^{\beta-2}\phi^{1-\beta}\left(\frac{s_{HK}}{s_{HL}}\right)^{\beta}\left(\frac{\delta s}{1+\delta s}-\frac{s_H}{s_L}\frac{f^{-\frac{1}{1+\delta s}}}{1+\delta s}\right)^{-\beta}\left(\frac{\beta(1-\beta)}{\alpha(1-\alpha)}\frac{s_{HL}}{s_{HK}}+1\right)=1. \tag{5.61}$$

Next the steady-state values of physical capital per unskilled labor, final output per unskilled labor and welfare can be solved. Combining equations (5.14), (5.15), (5.53) and (5.54), physical capital per unskilled labor can be expressed as

$$\left(\frac{K}{L}\right)^{*} = \text{constant}\cdot\left(\frac{s_{HK}}{s_{HL}}\right)^{\frac{\beta-1}{\alpha}}\left(\frac{\delta s}{1+\delta s}-\frac{s_H}{s_L}\frac{f^{-\frac{1}{1+\delta s}}}{1+\delta s}\right)^{\frac{\alpha-\beta}{\alpha}}. \tag{5.62}$$

Equation (5.62) is not in reduced form, since s_{HK} and s_{HL} are endogenous. In contrast to Chapter 4, the current two-sector model is not tractable enough to give reduced-form solutions for s_{HK} and s_{HL}. However, comparative statistic exercises can still be conducted in section 5.4.

Since $Y_{Lt} = Y_{Kt}$ in equilibrium, $Y_t = Y_{Lt}$. Using equations (5.58) and (5.60), the output per unskilled labor can be expressed as

$$\left(\frac{Y}{L}\right)^{*} = \text{constant}\cdot\frac{(1-s_{HK})^{\frac{(1-\alpha)(1-\beta)}{\alpha}}}{(s_{HK})^{\frac{1-\alpha}{\alpha}}}\left(f^{\frac{1}{1+\delta s}}\right)^{\frac{(1-\alpha)\beta}{\alpha}}\left(\frac{s_H}{s_L}\right)^{\frac{(\alpha-1)\beta}{\alpha}}. \tag{5.63}$$

From equation (5.46), the common steady state welfare to all agents is denoted ω^*. Using equations (5.30), (5.37), (5.39), (5.41), (5.43), (5.54), (5.58) and (5.62), the steady state welfare is expressed as

$$\omega^* = \text{constant}\cdot\delta^{\delta s}(1+\delta s)^{-(1+\delta s)}\left(\frac{s_{HK}}{s_{HL}}\right)^{\delta s+\frac{(1+\delta s)(\beta-1)}{\alpha}}\left(\frac{\delta s}{1+\delta s}-\frac{s_H}{s_L}\frac{f^{-\frac{1}{1+\delta s}}}{1+\delta s}\right)^{-\left(\delta s+\frac{\beta(1+\delta s)}{\alpha}\right)}. \tag{5.64}$$

In the next section, comparative statistic analysis with respect to the survival rate is conducted to study the impacts of population aging upon educational effort $\frac{s_H}{s_L}$, directed technical change s_{HK}, physical capital per unskilled labor($\left(\frac{K}{L}\right)^*$, output per unskilled labor $\left(\frac{Y}{L}\right)^*$ and welfare ω^*.

5.4 The effect of population aging

The above section solves the model in the steady state and this section investigates the impacts of population aging, for the steady state variables. Population aging is defined by an increase in the survival rate s. When young students are more likely to survive into the old age to get their payoff, they have more incentives to do education when young and this could increase the fraction of people doing education so there are more scientists in the q_L and q_K sectors. However, the number of scientists in the two research sectors may not increase at the same rate and this is a bias in technical change and this will depend on the relative strength of the price effect and market size effect, as will be discussed below.

Below, the impacts of population aging upon the direction of technical change, ratio of students (educational efforts), physical capital accumulation, output growth and steady-state welfare are analyzed. This is done via comparative statistic analysis, with respect to an increase in the survival rate s.

5.4.1 The direction of technical change

In equations (5.57) and (5.61) there are two endogenous variables, s_H (which determines $s_L = 1 - s_H$) and s_{HK} (which determines $s_{HL} = 1 - s_{HK}$), and one exogenous variable s. This determines s_L, s_K, s_{HK} and s_{HL} all as functions of s. From equation (5.57) it follows that

$$\frac{s_H}{s_L} = \delta s f^{\frac{1}{1+\delta s}} \left(1 + \frac{(1+\delta s)s_{HK}}{(1-\beta)}\right)^{-1}. \tag{5.65}$$

Substituting equation (5.65) into equation (5.61), taking *log* operation, and differentiating the resulting equation with respect to s, yields

$$\frac{\partial s_{HK}}{\partial s} \left\{ \frac{\beta}{1 - s_{HK}} + \frac{\beta(\beta-1)}{\alpha(1-\alpha)(s_{HK})^2 + \beta(1-\beta)(1-s_{HK})s_{HK}} + \frac{\beta(1+\delta s)}{(1-\beta)+(1+\delta s)s_{HK}} \right\}$$
$$= \frac{\beta(1-\beta+s_{HK})}{s\left((1-\beta)+(1+\delta s)s_{HK}\right)} > 0, \tag{5.66}$$

and rearranging (5.66) gives

$$\begin{aligned}\frac{\partial s_{HK}}{\partial s} =& \frac{\beta(1-\beta+s_{HK})}{s\left((1-\beta)+(1+\delta s)s_{HK}\right)} \\ &\times \left\{\frac{\beta}{1-s_{HK}} + \frac{\beta(\beta-1)}{\alpha(1-\alpha)(s_{HK})^2+\beta(1-\beta)(1-s_{HK})s_{HK}} + \frac{\beta(1+\delta s)}{(1-\beta)+(1+\delta s)s_{HK}}\right\}^{-1}.\end{aligned} \tag{5.67}$$

Equation (5.67) tells how the relative share of scientists in the q_K sector is affected by population aging. Since $s_{HL} = 1 - s_{HK}$, once the sign of $\frac{\partial s_{HK}}{\partial s}$ is obtained, so it that of $\frac{s_{HK}}{s_{HL}} = \frac{s_{HK}}{1-s_{HK}}$. This is exactly what directed technical change focuses on. With technical details left to the appendix 5.7.2, the following results concerning the direction of technical change are established.

Proposition 1 *In the two-sector growth model described above, assume the two intermediates are perfect complements. Before population aging, if there is a large (small) proportion of scientists working in the R&D sector yielding technology augmenting physical capital, then population aging will result in even more (less) scientists to that R&D sector. More specifically, there exists a value for $s_{HK} \in (0,1)$, denoted s^*_{HK}, such that*

- $\frac{\partial s_{HK}}{\partial s} > 0$ *and* $\frac{\partial (s_{HK}/s_{HL})}{\partial s} > 0$ *if* $s_{HK} \in (s^*_{HK}, 1)$;
- $\frac{\partial s_{HK}}{\partial s} < 0$ *and* $\frac{\partial (s_{HK}/s_{HL})}{\partial s} < 0$ *if* $s_{HK} \in (0, s^*_{HK})$.[2]

The intuition of Proposition 1 is left to subsection 5.4.3, after Proposition 2 is derived, because results in Proposition 2 are needed to explain Proposition 1.

5.4.2 Fraction of schooling agents, capital, output and welfare

The analysis of the effect of population aging upon $\frac{s_H}{s_L}$ is in the appendix 5.7.3. The main result is that $\frac{\partial s_H/s_L}{\partial s} > 0$. As shown in the appendix 5.7.3, the assumption $\varepsilon = 0$ is not used in getting the sign of $\frac{\partial s_H/s_L}{\partial s}$, which means $\frac{\partial s_H/s_L}{\partial s} > 0$ holds for a general ε.

With respect to $\left(\frac{K}{L}\right)^*$ in equation (5.62), the sign of $\frac{\partial (K/L)^*}{\partial s}$ is ambiguous in general. However, some useful insights can be obtained in the very special case $\alpha = \beta$. When $\alpha = \beta$,

[2] In this two-sector model, s_{HKt} cannot be equal to 0 or 1 in any period. If $s_{HKt} = 0$ for some t, then since research machines from q_K sector fully depreciate after use, at $t+1$ there is no research machine for Y_K production and this makes K useless at $t+1$. Anticipating this, L at t will choose zero saving. This in turn implies K and Y_K are zero at all $s > t$. Unless the elasticity of substitution between Y_L and Y_K is infinite, the case $s_{HKt} = 0$ is ruled out for all t.

(5.62) implies that $\left(\frac{K}{L}\right)^*$ decreases in s_{HK}/s_{HL}. Together with proposition 1, it follows that: $\frac{\partial (K/L)^*}{\partial s} > 0$ *if* $s_{HK} \in (0, s^*_{HK})$ *and* $\frac{\partial (K/L)^*}{\partial s} < 0$ *if* $s_{HK} \in (s^*_{HK}, 1)$.

Combining equations (5.63) and (5.65), it follows that

$$\left(\frac{Y}{L}\right)^* = \text{constant} \cdot \frac{(1-s_{HK})^{\frac{(1-\alpha)(1-\beta)}{\alpha}}}{(s_{HK})^{\frac{1-\alpha}{\alpha}}} \left(\frac{(1-\beta)+(1+\delta s)s_{HK}}{1-\beta}\right)^{\frac{(1-\alpha)\beta}{\alpha}} s^{\frac{(\alpha-1)\beta}{\alpha}}. \quad (5.68)$$

From (5.68) it follows immediately that $\frac{\partial (Y/L)^*}{\partial s} < 0$ when $s_{HK} \in (s^*_{HK}, 1)$. However, in the case $s_{HK} \in (0, s^*_{HK})$, the sign of $\frac{\partial (Y/L)^*}{\partial s}$ is ambiguous. Next the effect of population aging upon the steady-state welfare is analyzed. From equation (5.64), $\frac{\partial \omega^*}{\partial s} < 0$ if $s_{HK} > s^*_{HK}$ and the sign of $\frac{\partial \omega^*}{\partial s}$ is ambiguous when $s_{HK} < s^*_{HK}$. The findings above are summarized in the following proposition:

Proposition 2 1. *Population aging results in more agents going to school versus unskilled labor, namely,* $\frac{\partial (s_H/s_L)}{\partial s} > 0$.

2. *Assuming $\alpha = \beta$, if there is a large proportion of scientists in the R&D sector yielding technology augmenting physical capital (if $s_{HK} > s^*_{HK}$), then population aging will cause a decrease in physical capital per unskilled labor, output per unskilled labor and welfare.*

3. *Assuming $\alpha = \beta$, if there is a small proportion of scientists in the R&D sector yielding technology augmenting physical capital (if $s_{HK} < s^*_{HK}$), then population aging will cause a rise in physical capital per unskilled labor (assuming $\alpha = \beta$), and its effects on output per unskilled labor and welfare are ambiguous.*

After population aging, there are more agents choosing to do education and become skilled upon survival into the old age. This result is similar to the main result in Chapter 4 of the thesis. This shows the insight from the one-sector growth model can be generalized to the two-sector growth model. The intuition here is also similar. The present values of payoffs to a scientist working in q_K and q_L sectors are $\frac{\Pi_{Ht+1}(L)}{H_{Lt+1}} \frac{s}{R_{t+1}}$ and $\frac{\Pi_{Ht+1}(K)}{H_{Kt+1}} \frac{s}{R_{t+1}}$, respectively. With population aging (an increase in s), the discounted lifetime payoff is higher and this provides higher incentives to do education.

In Chapter 4 the one-sector growth model, $\frac{\partial (K/L)^*}{\partial s} > 0$, but here in the two-sector model with $s_{HK} \in (s^*_{HK}, 1)$ the opposite result holds. The reason is that in the case where $s_{HK} > s^*_{HK}$, population aging results in a higher s_{HK}/s_{HL}. In the two-sector model with perfect complements between the two intermediates, with more R&D conducted in the q_K sector, it is equivalent to a rise in 'effective' K, hence more L is needed to balance between Y_K and Y_L, which leads to a fall in $\frac{K}{L}$ ratio.[3] In contrast, when $s_{HK} < s^*_{HK}$, the opposite happens. Similar to the one-sector model in Chapter 4, the two-sector growth model is neo-classical in essence, hence a slower physical capital accumulation leads to a decrease in welfare.

5.4.3 Discussion of Proposition 1

In order to understand the intuition of Proposition 1, some review on established results in directed technical change literature is presented below.

In the directed technical change literature, the price and market size effects jointly determine the direction of technical change. According to the price effect, there are stronger incentives to develop technologies when the goods produced by these technologies command higher prices. According to the market size effect, it is more profitable to develop technologies that have a larger market. In general equilibrium models, the prices are endogenously determined, and so is the price effect. Usual substitution effect in microeconomic theory implies that, other things equal, more supply of a certain input will decrease the price of the good using this input in production. In this chapter's model, relatively more supply of K (L) will decrease the relative price of Y_K (Y_L), other things fixed. Therefore the price effect implies that more K will lower the incentive to do R&D in q_K sector, hence a *decrease* in $\frac{s_{HK}}{s_{HL}}$. At the same time, more K leads to a larger Y_K, other things fixed, and this gives the x_K research machine a larger market. The market size effect implies an increase in $\frac{s_{HK}}{s_{HL}}$. The net effect depends on whether the price or market size effect is stronger.

In the case of perfect complements, Y_H and Y_L must be used in a fixed proportion, hence there is no market size effect and only the price effect needs investigation. When $s_{HK} \in (0, s^*_{HK})$, $\frac{\partial(s_{HK}/s_{HL})}{\partial s} < 0$. As will be established soon in Proposition 2, when $s_{HK} \in (0, s^*_{HK})$, an increase in s will increase physical capital relative to unskilled labor and this will lower the relative price of capital-intensive good. Via price effect, there is less R&D incentive in the q_K sector and s_{HK}/s_{HL} should decrease. As a result, the price effect leads to a decrease in s_{HK}/s_{HL} when $s_{HK} \in (0, s^*_{HK})$.

On the other hand, if $s_{HK} \in (s^*_{HK}, 1)$, $\frac{\partial(s_{HK}/s_{HL})}{\partial s} > 0$. As will be shown in Proposition 2, if $s_{HK} \in (s^*_{HK}, 1)$, an increase in s will decrease physical capital relative to unskilled labor.

[3] This intuition is somewhat similar to the example in Acemoglu (2009) page 502, where H-augmenting technological change can be L-biased. This is absent in the one-sector growth model.

Via price effect, s_{HK}/s_{HL} should increase. The intuition here shows that the endogenous price effect plays a vital role in determining the direction of technical change. Treating prices (of both inputs and intermediates) as exogenous would yield results with limitations.

In the above analysis, the price effect is very strong in giving the main results. Formally, after assuming that the two intermediates are perfect complements, using equations (5.56) and (5.60), the ratio of two intermediate prices in terms of s_{HK} and s_{HL} can be expressed as

$$\frac{p_K}{p_L} = \text{constant} \cdot \frac{s_{HK}}{s_{HL}}. \tag{5.69}$$

which implies that $\frac{p_K}{p_L}$ and $\frac{s_{HK}}{s_{HL}}$ are perfectly positively correlated. This is an extreme case for dominant price effect.

5.5 The other extremity: perfect substitute case

In the above analysis the two intermediates are assumed as perfect complements (the elasticity of substitution $\varepsilon = 0$), which is an extreme case for constant elasticity of substitution. Due to perfect complementarity, the relative amount of two intermediates is fixed. As a consequence, the market size effect is ruled out by construction and only the price effect is in place. In this section, attention is drawn to the other extreme case, where the two intermediates are perfect substitutes, namely, when the elasticity of substitution $\varepsilon \to \infty$. Due to perfect substitution, the relative price of two intermediates must be constant, otherwise only one of the two intermediates will be used in the final production, making production of the other intermediate not profitable at all. As a result, in this section the price effect will be ruled out by construction and only the market size effect is in place.

According to Acemoglu (1998, 2009), the price and market size effects are two conflicting forces determining the direction of technical change. It is expected that the impacts of population aging on the direction of technical change in this section will be just the opposite to the previous section. When $\varepsilon \to \infty$, the algebra becomes extremely messy. Because the only focus is the direction of technical change, to avoid unnecessary tediousness, only the sign of $\frac{\partial(s_{HK}/s_{HL})}{\partial s}$ is analyzed.[4]

[4] As mentioned in the previous section, the result $\frac{\partial s_H/s_L}{\partial s} > 0$ is true for a general value of ε and obviously true in this $\varepsilon \to \infty$ case.

With $\varepsilon \to \infty$, the final production function (5.11) becomes

$$Y_t = Y_{Kt} + Y_{Lt}. \tag{5.70}$$

To make sure that both Y_K and Y_L are used during each period, their prices must satisfy

$$p_{Kt} = p_{Lt}, \quad \text{for all } t. \tag{5.71}$$

Using expressions for Y_K and Y_L from equations (5.58) and (5.59), it follows that

$$Y_t = (1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}} L_t \left(p_{Lt}\right)^{\frac{1-\alpha}{\alpha}} + (1-\beta)^{\frac{2}{\beta}-2}\phi^{\frac{\beta-1}{\beta}} K_t \left(p_{kt}\right)^{\frac{1-\beta}{\beta}}. \tag{5.72}$$

Since the focus is on the comparative statistics in the steady state, all time subscripts are dropped. Using the formula for $\frac{K}{L}$ from equation (5.62), equation (5.72) becomes

$$\begin{aligned} Y &= L\left(p_L\right)^{\frac{1}{\alpha}} \left((1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}} \frac{1}{p_L} + (1-\beta)^{\frac{2}{\beta}-2}\phi^{\frac{\beta-1}{\beta}} \frac{K}{L} (p_K)^{\frac{1}{\beta}-1-\frac{1}{\alpha}} \right) \\ &= L(p_L)^{\frac{1}{\alpha}-1}(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}} \left(1 + \frac{\alpha(1-\alpha)}{\beta(1-\beta)} \frac{s_{HK}}{1-s_{HK}}\right). \end{aligned} \tag{5.73}$$

Next the market clearing condition for the final output, expressed in equation (5.26), is used. Substituting in all the endogenous variables gives

$$\begin{aligned} Y_t =& \frac{\delta s}{1+\delta s} w_t L_t \\ &+ MC_{Lt} \left((1-\alpha)^2 p_{Lt} L_t^{\alpha} (q_{Lt})^{\theta_L} \frac{1}{MC_{Lt}} \right)^{\frac{1}{\alpha}} + MC_{Kt} \left((1-\beta)^2 p_{Kt} K_t^{\beta} (q_{Kt})^{\theta_K} \frac{1}{MC_{Kt}} \right)^{\frac{1}{\beta}} \\ &+ \frac{1}{1+\delta s} w_t L_t + \frac{s}{1+\delta s} \frac{1}{R_{t+1}} \frac{\Pi_{Ht+1}(L)}{H_{Lt+1}} \frac{H_{Lt+1}}{s} + \frac{s}{1+\delta s} \frac{1}{R_{t+1}} \frac{\Pi_{Ht+1}(K)}{H_{Kt+1}} \frac{H_{Kt+1}}{s} \\ &+ \frac{R_t \delta}{1+\delta s} w_{t-1} s L_{t-1} + \frac{\delta s}{1+\delta s} \frac{\Pi_{Ht}(L)}{H_{Lt}} H_{Lt} + \frac{\delta s}{1+\delta s} \frac{\Pi_{Ht}(K)}{H_{Kt}} H_{Kt}. \end{aligned} \tag{5.74}$$

Evaluating all variables at their steady-state values gives (time subscripts dropped)

$$\begin{aligned} Y =& \alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}(p_L)^{\frac{1}{\alpha}} L + (1-\alpha)^{\frac{2}{\alpha}}(p_L)^{\frac{1}{\alpha}} L \psi^{\frac{\alpha-1}{\alpha}} + (1-\beta)^{\frac{2}{\beta}}(p_K)^{\frac{1}{\beta}} K \phi^{\frac{\beta-1}{\beta}} \\ &+ \frac{1}{(1+\delta s)R} \left(\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}(p_L)^{\frac{1}{\alpha}} L + \beta(1-\beta)^{\frac{2}{\beta}-1}(p_K)^{\frac{1}{\beta}} K \phi^{\frac{\beta-1}{\beta}} \right) \\ &+ \frac{\delta s}{1+\delta s} R L \alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}(p_L)^{\frac{1}{\alpha}} \\ &+ \frac{\delta s}{1+\delta s} \left(\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}(p_L)^{\frac{1}{\alpha}} L + \beta(1-\beta)^{\frac{2}{\beta}-1}\phi^{\frac{\beta-1}{\beta}}(p_K)^{\frac{1}{\beta}} K \right). \end{aligned} \tag{5.75}$$

Combining equations (5.73) and (5.75), denoting s_{HK} by x, it follows that

$$\begin{aligned}
&\frac{1}{p_L}(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\left(1+\frac{\alpha(1-\alpha)}{\beta(1-\beta)}\frac{x}{1-x}\right)\\
&=\alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}+(1-\alpha)^{\frac{2}{\alpha}}\psi^{\frac{\alpha-1}{\alpha}}+\frac{\alpha}{\beta}(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}(1-\beta)\frac{x}{1-x}\\
&+\frac{1}{1+\delta s}\frac{1}{R}\left(\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{1}{1-x}\right)\\
&+\frac{\delta s}{1+\delta s}R\alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\\
&+\frac{\delta s}{1+\delta s}\left(\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{1}{1-x}\right),
\end{aligned} \tag{5.76}$$

where $p_K = p_L$ and p_K is obtained from equation (5.56) (note that equation (5.60) for p_L only works for the case where $\varepsilon = 0$ and does not apply if $\varepsilon \to \infty$). R can be got from equation (5.40). In this sense, all variables in the equation (5.76) are either primitive parameters or in terms of x and this implicitly defines x as a function of s. Differentiating the above equation with respect to s, after some rearrangement, the following (x' is used to denote $\frac{\partial s_{HK}}{\partial s}$) is obtained

$$(\text{coefficient of } x')\cdot x' = RHS, \tag{5.77}$$

where

$$\begin{aligned}
&(\text{coefficient of } x')\\
&\equiv\frac{(1-\alpha)^{\frac{2}{\alpha}-1}\alpha\psi^{\frac{\alpha-1}{\alpha}}(1-\beta)}{\beta}\frac{1}{(1-x)^2}+\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{\delta(1-\beta)^2 s}{(1+\delta s)(1-x)(1-\alpha)\left(1-\beta+(1+\delta s)x\right)^2}\\
&+\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{(1-\beta)\delta s x}{(1+\delta s)(1-x)^2(1-\alpha)(1-\beta+(1+\delta s)x)}\\
&-\alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\frac{(1-\alpha)(1-\beta)s}{(1+\delta s)(1-\beta)sx^2}\\
&+\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{\delta s}{1+\delta s}\frac{1}{(1-x)^2}\\
&+\frac{(1-\alpha)^{\frac{2}{\alpha}-2}}{p_L}\psi^{\frac{\alpha-1}{\alpha}}\left\{\frac{\beta}{1-\beta+(1+\delta s)x}\frac{s+\delta s^2+(1-\beta)s}{(1-x)s}\left(1+\frac{\alpha(1-\alpha)x}{\beta(1-\beta)(1-x)}\right)-\frac{\alpha(1-\alpha)}{\beta(1-\beta)}\frac{1}{(1-x)^2}\right\}
\end{aligned} \tag{5.78}$$

and

$$
\begin{aligned}
& RHS \\
& \equiv (1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\frac{1}{p_L}\left\{\frac{\beta}{1-\beta+(1+\delta s)x}\frac{x^2-\beta x-(1-\beta)}{(1-x)s}\left(1+\frac{\alpha(1-\alpha)x}{\beta(1-\beta)(1-x)}\right)\right\} \\
& +\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\left\{\frac{\delta^2(1-\beta)sx}{(1+\delta s)^2(1-x)(1-\alpha)(1-\beta+(1+\delta s)x)}\right\} \\
& -\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\left\{-\frac{\delta(1-\beta)^2x+\delta(1-\beta)x^2}{(1+\delta s)(1-x)(1-\alpha)(1-\beta+(1+\delta s)x)^2}\right\} \\
& +\alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\left\{-\frac{(1-\alpha)(1-\beta+(1+\delta s)x)}{(1+\delta s)^2(1-\beta)sx}+\frac{(1-\alpha)((1-\beta)x+x^2)}{(1+\delta s)(1-\beta)sx^2}\right\} \\
& +\alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\left(-\frac{\delta}{(1+\delta s)^2}\frac{1}{1-x}\right). \qquad (5.79)
\end{aligned}
$$

With technical details in the appendix 5.7.4, the following proposition about the impacts of population aging upon the direction of technological change is obtained.

Proposition 3 *In the two-sector growth model, assume the two intermediates are perfect substitutes. Before population aging, if there is a small (large) proportion of scientists working in the R&D sector yielding technology augmenting physical capital, then population aging will conduct more (less) scientists to that R&D sector. The market size effect is dominant in determining the direction of technical change after population aging. There exists* $\underline{s_{HK}}$ *and* $\overline{s_{HK}}$*, satisfying* $0 < \underline{s_{HK}} < \overline{s_{HK}} < 1$*, such that*

- $\frac{\partial s_{HK}}{\partial s} > 0$ *and* $\frac{\partial (s_{HK}/s_{HL})}{\partial s} > 0$ *if* $0 < s_{HK} < \underline{s_{HK}}$*;*
- $\frac{\partial s_{HK}}{\partial s} < 0$ *and* $\frac{\partial (s_{HK}/s_{HL})}{\partial s} < 0$ *if* $\overline{s_{HK}} < s_{HK} < 1$*.*

The impacts of population aging upon directed technical change in Proposition 3 are exactly opposite to the results in Proposition 1. This means that when the relationship between two intermediates changes from perfect complements to perfect substitutes, the impacts of aging are completely reversed.

When two intermediates are perfect substitutes, there is only price effect and there is no market size effect. In contrast, when two intermediate are perfect complements, there is only market size effect and there is no price effect. According to Acemoglu (1998; 2009), price effect and market size effect have opposite impacts on the direction of technical change. This explains the sharp contrast between Propositions 1 and 3.

5.6 Conclusion

Population aging, as characterized by a higher fraction of old people in the economy, has become a worldwide phenomenon with important consequences. This chapter, utilizing a two-sector endogenous growth model, analyzes how population aging affects the direction and bias of technological change across different research sectors.

Several interesting and intuitive results are obtained. First, similar to the one-sector growth in Chapter 4, it is found that population aging through a higher survival rate increases the overall size of schooling agents and hence scientists. A higher probability of surviving into the old age results in a higher present value of doing research and this provides higher incentives for agents to conduct education.

Including both the price and the market size effects, the analysis about population aging and directed technical change is conducted in two special cases. When the two intermediates are perfect complements, the price effect is dominant. If there is a large initial proportion of scientists in the R&D sector yielding technology augmenting physical capital (if $s_{HK} > s^*_{HK}$), population aging will induce relatively even more scientists into that sector. On the other hand, if the proportion of scientists in the R&D sector yielding technology augmenting physical capital is small (if $s_{HK} < s^*_{HK}$), population aging will induce relatively more scientists into the other research sector.

When the two intermediates are perfect substitutes, the market size effect is dominant and the opposite results are obtained. If there is a large proportion of scientists in the R&D sector yielding technology augmenting physical capital (if $s_{HK} > s^{2*}_{HK}$), population aging will induce relatively less scientists into that sector. On the other hand, if the initial proportion of scientists in the R&D sector yielding technology augmenting physical capital is small (if $s_{HK} < s^{1*}_{HK}$), population aging will induce relatively more scientists into this research sector.

There are several possibilities for future research. First, as in Chapter 4 of the thesis, ability heterogeneities can be introduced. A richer model will take this into account but the algebra will be more complicated. Second, to get analytic results, the two intermediates are assumed either perfect complements or perfect substitutes. It is postulated that other degree of substitutability lies in between these two extremes cases. However, the reduced-form solution is still lacking and needs more investigation. Moreover, extending the current analysis to a multi-sector growth model would be interesting.

5.7 Appendix

5.7.1 Derivation of equation (5.12)

To simplify the notation, all t subscripts are dropped here. The final producer solves the following problem:

$$\max_{Y_K, Y_L} \quad p \cdot \left(Y_L^{\frac{\varepsilon-1}{\varepsilon}} + Y_K^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}} - p_L Y_L - p_K Y_K,$$

where p denotes the price of the final good (p is set to unitary in this chapter's model). This problem yields the following

$$\begin{aligned} p_L =& p \cdot \left(Y_L^{\frac{\varepsilon-1}{\varepsilon}} + Y_K^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{1}{\varepsilon-1}} (Y_L)^{-\frac{1}{\varepsilon}}, \\ p_K =& p \cdot \left(Y_L^{\frac{\varepsilon-1}{\varepsilon}} + Y_K^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{1}{\varepsilon-1}} (Y_K)^{-\frac{1}{\varepsilon}}, \end{aligned}$$

from which it follows that

$$\begin{aligned} (Y_L)^{\frac{1}{\varepsilon}} =& p \cdot \left(Y_L^{\frac{\varepsilon-1}{\varepsilon}} + Y_K^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{1}{\varepsilon-1}} (p_L)^{-1}, \\ (Y_K)^{\frac{1}{\varepsilon}} =& p \cdot \left(Y_L^{\frac{\varepsilon-1}{\varepsilon}} + Y_K^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{1}{\varepsilon-1}} (p_K)^{-1}, \end{aligned}$$

and

$$\begin{aligned} (Y_L)^{\frac{\varepsilon-1}{\varepsilon}} =& p^{\varepsilon-1} \left(Y_L^{\frac{\varepsilon-1}{\varepsilon}} + Y_K^{\frac{\varepsilon-1}{\varepsilon}} \right) (p_L)^{1-\varepsilon}, \\ (Y_K)^{\frac{\varepsilon-1}{\varepsilon}} =& p^{\varepsilon-1} \left(Y_L^{\frac{\varepsilon-1}{\varepsilon}} + Y_K^{\frac{\varepsilon-1}{\varepsilon}} \right) (p_K)^{1-\varepsilon}. \end{aligned}$$

Adding up the previous two equations gives

$$p = \left((p_L)^{1-\varepsilon} + (p_K)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}.$$

Evaluating $p = 1$ and adding in time subscripts gives equation (5.12).

5.7.2 Proof of Proposition 1

In order to prove Proposition 2, equation (5.66) is used, reproduced here as

$$\begin{aligned} &\frac{\partial s_{HK}}{\partial s} \left\{ \frac{\beta}{1 - s_{HK}} + \frac{\beta(\beta - 1)}{\alpha(1-\alpha)(s_{HK})^2 + \beta(1-\beta)(1-s_{HK})s_{HK}} + \frac{\beta(1+\delta s)}{(1-\beta) + (1+\delta s)s_{HK}} \right\} \\ &\quad = \frac{\beta(1 - \beta + s_{HK})}{s\left((1-\beta) + (1+\delta s)s_{HK}\right)} > 0. \end{aligned}$$

The left-hand-side of the above equation can be simplified to:

$$\frac{\partial s_{HK}}{\partial s}\beta\frac{\text{numerator}}{(1-s_{HK})[\alpha(1-\alpha)(s_{HK})^2+\beta(1-\beta)(1-s_{HK})s_{HK}][(1-\beta)+(1+\delta s)s_{HK}]},$$

where

$$\text{numerator} \equiv A(s_{HK})^2 + Bs_{HK} + C,$$

where

$$A \equiv -\beta(1-\beta)^2+\alpha(1-\alpha)(1-\beta)+(1+\delta s)(1-\beta)-\beta(1-\beta)+(1+\delta s)+\alpha(1-\alpha)(1+\delta s), \tag{5.A.1}$$

$$B \equiv \beta(1-\beta)^2+(1-\beta)^2-(1+\delta s)(1-\beta)^2, \tag{5.A.2}$$

$$C \equiv -(1-\beta)^2. \tag{5.A.3}$$

One important fact is that the numerator is a polynomial (in s_{HK}) of degree 2, so its first-order derivative is a polynomial of degree 1, hence linear and monotonic in s_{HK} since $s_{HK} > 0$. The numerator is negative when $s_{HK} = 0$ and positive when $s_{HK} = 1$. Since the numerator is continuous in s_{HK}, it must be equal to zero at least once (existence). Next it needs to be shown that it is equal to zero only once (uniqueness) and this is proved by contradiction.

Suppose the numerator is equal to zero twice, when it must cross the s_{HK} at least three times, and the three crossing points can be labeled as x_1, x_2 and x_3 with $x_1 < x_2 < x_3$. The numerator between x_1 and x_3 must go up from x_1 and then *down* to x_2 then *up* to x_3. The numerator would fluctuate between x_1 and x_3 and this contradicts to the fact that the numerator is monotonic when $s_{HK} > 0$. This shows that there can not be two points at which the numerator is equal to zero. A similar argument shows there can not be more than one point at which the numerator is equal to zero. This shows that there exists one and only one point, denoted by s^*_{HK} such that the numerator is equal to zero at s^*_{HK}, negative when $s_{HK} < s^*_{HK}$ and positive when $s_{HK} > s^*_{HK}$. Q.E.D.

5.7.3 Analysis of $\frac{\partial log(s_H/s_L)}{\partial s}$ in Proposition 2

Taking *log* operation of equation (5.65) and differentiating it with respect to s, together with equation (5.67), gives the change of s_H/s_L with respect to s as

$$\begin{aligned}\frac{\partial log(s_H/s_L)}{\partial s} =& \frac{1}{s} + \frac{-\delta \cdot log(f)}{(1+\delta s)^2} \\ &- \frac{\delta s_{HK} + (1+\delta s)\frac{(1-\beta+s_{HK})(1-s_{HK})s_{HK}[\beta(1-\beta)(1-s_{HK})+\alpha(1-\alpha)s_{HK}]}{s[A(s_{HK})^2+Bs_{HK}+C]}}{(1-\beta)+(1+\delta s)s_{HK}},\end{aligned}$$

where A, B and C are given in equations (5.A.1), (5.A.2) and (5.A.3).

The expression for $\frac{\partial log(s_H/s_L)}{\partial s}$ is equivalent to a polynomial (in s_{HK}) of degree 1, hence monotonic in s_{HK}. It is easy to verify that

$$\frac{\partial log(s_H/s_L)}{\partial s}|_{s_{HK}=0} = \frac{1}{s} + \frac{\delta(-log(f))}{(1+\delta s)^2} > 0,$$

and

$$\begin{aligned}\frac{\partial log(s_H/s_L)}{\partial s}|_{s_{HK}=1} &= \frac{1}{s} + \frac{\delta(-log(f))}{(1+\delta s)^2} - \frac{\delta}{(1-\beta)+(1+\delta s)} \\ &> \frac{1}{s} + \frac{\delta(-log(f))}{(1+\delta s)^2} - \frac{\delta}{\delta s} > 0.\end{aligned}$$

Therefore, $\frac{\partial s_H/s_L}{\partial s} > 0$ for all $s_{HK} \in (0,1)$. With population aging, there will be more agents choosing to do education. In other words, population aging encourages overall educational efforts.

5.7.4 Proof of Proposition 3

To analyze the case where two intermediates are prefect substitutes, equations (5.77) to (5.79) are used, reproduced below

$$(\text{coefficient of } x') \cdot x' = RHS, \tag{5.77}$$

where

(coefficient of x')

$$\equiv \frac{(1-\alpha)^{\frac{2}{\alpha}-1}\alpha\psi^{\frac{\alpha-1}{\alpha}}(1-\beta)}{\beta}\frac{1}{(1-x)^2} + \alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{\delta(1-\beta)^2 s}{(1+\delta s)(1-x)(1-\alpha)\left(1-\beta+(1+\delta s)x\right)^2}$$
$$+ \alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{(1-\beta)\delta s x}{(1+\delta s)(1-x)^2(1-\alpha)(1-\beta+(1+\delta s)x)}$$
$$- \alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\frac{(1-\alpha)(1-\beta)s}{(1+\delta s)(1-\beta)sx^2}$$
$$+ \alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{\delta s}{1+\delta s}\frac{1}{(1-x)^2}$$
$$+ \frac{(1-\alpha)^{\frac{2}{\alpha}-2}}{p_L}\psi^{\frac{\alpha-1}{\alpha}}\left\{\frac{\beta}{1-\beta+(1+\delta s)x}\frac{s+\delta s^2+(1-\beta)s}{(1-x)s}\left(1+\frac{\alpha(1-\alpha)x}{\beta(1-\beta)(1-x)}\right) - \frac{\alpha(1-\alpha)}{\beta(1-\beta)}\frac{1}{(1-x)^2}\right\} \tag{5.78}$$

and

$$RHS$$
$$\equiv (1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\frac{1}{p_L}\left\{\frac{\beta}{1-\beta+(1+\delta s)x}\frac{x^2-\beta x-(1-\beta)}{(1-x)s}\left(1+\frac{\alpha(1-\alpha)x}{\beta(1-\beta)(1-x)}\right)\right\}$$
$$+ \alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\left\{\frac{\delta^2(1-\beta)sx(1-\beta+(1+\delta s)x)-(1+\delta s)\delta(1-\beta)x(1-\beta+x)}{(1+\delta s)^2(1-x)(1-\alpha)(1-\beta+(1+\delta s)x)^2}\right\}$$
$$+ \alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\left\{-\frac{(1-\alpha)(1-\beta+(1+\delta s)x)}{(1+\delta s)^2(1-\beta)sx} + \frac{(1-\alpha)((1-\beta)x+x^2)}{(1+\delta s)(1-\beta)sx^2}\right\}$$
$$+ \alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\left(-\frac{\delta}{(1+\delta s)^2}\frac{1}{1-x}\right). \tag{5.79}$$

The RHS in equation (5.79) consists of four parts (lines). The fourth line is obviously negative. The second line can be rearranged as

$$\alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{\delta(1-\beta)\delta sx(1-\beta+(1+\delta s)x)-(1+\delta s)(\delta(1-\beta)^2x+\delta(1-\beta)x^2)}{(1+\delta s)^2(1-x)(1-\alpha)(1-\beta+(1+\delta s)x)^2}$$
$$= \alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{\delta(1-\beta)x}{(1+\delta s)^2(1-x)(1-\alpha)(1-\beta+(1+\delta s)x)^2}\left[-(1-\beta)-(1-\delta s)(1+\delta s)x\right]$$
$$< 0,$$

and the third line can be rearranged as

$$-\alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\frac{(1-\alpha)(1-\beta)+(1-\alpha)(1+\delta s)x-(1-\alpha)(1-\beta)-(1-\alpha)x}{(1+\delta s)^2(1-\beta)sx}$$
$$= \alpha(1-\alpha)^{\frac{2}{\alpha}-1}\psi^{\frac{\alpha-1}{\alpha}}\frac{-\alpha}{(1+\delta s)^2(1-\beta)}$$
$$< 0.$$

and the sign of the first line depends on $x^2-\beta x-(1-\beta)$. It can be seen that $x^2-\beta x-(1-\beta) < 0$ for $x \in [0, 1)$ and $x^2 - \beta x - (1 - \beta) = 0$ for $x = 1$.

Summarizing up the analysis above, it follows that

$$RHS < 0 \qquad \text{for all } x \in [0, 1].$$

Next sign of *coefficient of* x' is analyzed. Equation (5.78) consists of five lines. Its sign is quite complicated to pin down in general, but attention can be drawn to some (small) ranges near $x = 0$ and $x = 1$.

When $x \to 0$, the third term, $-\alpha(1-\alpha)^{\frac{2}{\alpha}-2}\psi^{\frac{\alpha-1}{\alpha}}\frac{(1-\alpha)(1-\beta)s}{(1+\delta s)(1-\beta)sx^2}$, converges to $-\infty$, while all the other four lines are finite. Therefore, it follows that

$$\text{coefficient of } x'|_{x\to 0} < 0.$$

On the other hand, while $x \to 1$, many terms will converge to either ∞ or $-\infty$. The term with the 'quickest converging speed' can be found and that term will dominant the overall sign of *coefficient of* x' when $x \to 1$. Substituting in $p_L = p_K$ from equation (5.56), among all terms with $(1 - x)$, the term with the *largest* power for $(1 - x)$ in the denominator will have the quickest converging speed (converge to either ∞ or $-\infty$). It turns out that this term is in the 1st, 2nd and 4th lines of equation (5.78), and all of them are positive. Therefore, the overall sign of *coefficient of* x' will be positive, when $x \to 1$.

Summarizing all the results above, it follows that

$$(-) \cdot x' = (-)|_{x\to 0}, \quad \text{and}$$
$$(+) \cdot x' = (-)|_{x\to 1},$$

which immediately gives the following key result (x' denotes $\frac{\partial s_{HK}}{\partial s}$):

$$\frac{\partial s_{HK}}{\partial s}|_{s_{HK}\to 0} > 0 \quad \text{and,}$$
$$\frac{\partial s_{HK}}{\partial s}|_{s_{HK}\to 1} < 0.$$

Since both *coefficient of* x' and RHS in equation (5.77) are continuous in x, the results in proposition 3 follow. Since the sign of *coefficient of* x' is very complicated to determine for a general value of x, only the analysis near $x = 0$ and $x = 1$ can be investigated. However, this already shows how the results in the $\varepsilon \to \infty$ case differ from the $\varepsilon = 0$ case.

Chapter 6

Population aging, education and skill premium under trade

6.1 Introduction

The most important demographic shock nowadays is population aging, which has happened to nearly all developed countries and many developing countries as well.

Population aging's impacts upon education have drawn much attention. While Chapters 4 and 5 have investigated the relationship between population aging and education, the models are restricted to closed economy. In this chapter, the first research question is to study the impacts of aging upon education in an international trade model. In a two-country two-good trade model under general equilibrium, under both autarky and trade equilibrium, population aging encourages domestic educational efforts. This result shows the finding of Chapters 4 and 5 about aging and education can be generalized to a trade equilibrium. Intuitively, under both trade and autarky, population aging increases the expected payoff from education and this leads to higher educational efforts.

While the impacts of aging upon domestic education is the same as in Chapters 4 and 5, the impacts of one country's population aging upon education in the other country are different. Population aging in any country will discourage education in the other country. As population aging in Home leads to more skilled labor in Home, Home will produce more of skilled labor intensive goods and Foreign will produce more of unskilled labor intensive goods. As the demand for unskilled labor in Foreign increases, the incentive to do education in Foreign decreases. After population aging in Home, the shift in production and trade pattern moves more education from Foreign to Home, which can be called the education stealing effect.

Besides population aging and education, another focus of this chapter is on population aging and skill premium. Wage inequality in the U.S. and other OECD countries has changed dramatically in the past 50 years and there is a large literature explaining the change in skill premium over recent decades. The papers analyze the increase in skill premium from two different but related aspects, which are (1) productivity and technology (such as Parro, 2013; Krusell et al., 2000 and Jovanovic, 1998) and (2) international trade (such as Acemoglu, 2003; Burstein et al., 2003 and Burstein and Vogel, 2010).

In the above literature, skilled-biased technical change is necessary to explain the rise in the skill premium. The logic is that, because technological progress in the recent decades favors the skilled labor more than the unskilled labor, it increases the skill premium. Moreover, the above literature points out that classical trade models (the Stolper-Samuelson effect) cannot explain the rise in skill premium, but trade models combined with skill-biased technical change can explain the empirical evidence well.

In trade models explaining the rise in skill premium, skill-biased technical change is a necessary component. This chapter is trying to investigate whether classical trade models without assuming skill-biased technical change can explain the rise in skill premium. This chapter shows that the answer is yes. Even with a neutral technical change, population aging in a two-country two-good trade model can lead to a rise in skill premium, consistent with empirical data. This result suggests that population aging could be a very important factor causing the rise in skill premium in recent decades.

With nearly all literature on skill premium focus on technical change and international trade, just a few studies have mentioned that demographic change could be a possible factor. In a very recent paper, Acemoglu and Autor (2011) review the change in skill premium in recent decades. They focus on skill-biased technical change and international trade as explanations for the rise in skill premium. In terms of the reasons for skill-biased technical change, they claim that changes in skill supplies due to demographic trends can induce endogenous changes in technology and increase the demand for skills. According to their argument, there exists an important relationship between population aging and skill premium.

Van Zanden (2009) analyzes the long-term developments of the skill premium in western Europe since 1300. He studies the relationship between skill premium and estimates of population for England (1300-1800) and Italy (1326-1800). He finds that sill premium tends to increase in periods of population growth. He argues that one reason is the growth in the construction industry, and the related demand for the craftsmen, are intimately related to demographic growth. Another reason he argues is that, the wages of unskilled labor reflect labor productivity in agriculture and in the long run it is linked to population growth. The rapid demographic expansion of the sixteenth century led to a sharp decline in rural real wages, but the urban sector could protect itself against these trends.[1]

[1]This argument has some similarity with the innate skills theorem proposed by Nelson and Phelps (1966). In a seminal work, Nelson and Phelps (1966) argue that skilled labor have higher wage not just because of their higher productivity, but also in their abilities to deal with technological changes. The productivity of skilled labor is less adversely affected by the turmoil created by technological transformations of the workplace, and

According to United Nations (2011), in 1950 just 8 percent of the world population were aged 60 years or more, but by 2011 that proportion rose to 11.2 percent and is expected to reach 22 percent in 2050. A more indicative factor of population aging is the increase in median age. In 2011, most of the developed countries had a median age higher than 40 years, with Japan leading the group with 45.0 years. By 2050, 100 countries are expected to have a median age above 40 years and 158 countries by 2100. Population aging is currently pervasive mostly in developed countries, but it is expected to be common in the developing world very soon and is projected to occur even more rapidly in developing countries than it did in the developed countries. From the skill premium literature, the period where population aging has become most serious is associated with remarkable increase in skill premium, both in developed and developing countries. This co-occurrence suggests that there could be some important relationship between population aging and the increase in skill premium.

This chapter analyzes the importance of population aging upon skill premium, by including population aging in a trade model, and studies how population aging affects skill premium. It is found that, in autarky equilibrium, population aging will decrease skill premium. In contrast, in trade equilibrium, population aging in any country will increase skill premium for all countries in the world. The impact of aging upon skill premium in trade equilibrium is consistent with the empirical data.

The rest of the chapter is organized as follows. Section 6.2 describes the model in autarky, including the problems in the household, education and production sectors. Section 6.3 solves the optimization problems of all sectors and derives the autarky equilibrium. Under the autarky equilibrium, section 6.4 studies the impacts of population aging via comparative statistic analysis. Following the analysis under autarky, section 6.5 extends the autarky model to a two-country trade model, derives the international trade equilibrium and studies the impacts of population aging under trade equilibrium. Section 6.6 discusses the significance of results in section 6.5. In the end, section 6.7 concludes the chapter with a summary of the main results and some future research ideas.

---

it is less costly for them to learn new skills in order to implement a new technology. This theory can explain the rise in skill premium in both the first industrial revolution, or the more recent IT revolution. Hornstein et al. (2005) have a review of both empirical and theoretical studies extending Nelson and Phelps (1966).

6.2 Autarky model

The autarky model is a slight variation of Eicher (1996). Eicher (1996) uses a two-period OLG model, including household, education and production sectors, to study human capital accumulation via education and endogenous economic growth. This model differs from Eicher (1996) by adding in population aging via a higher survival probability. In Eicher's two-period OLG model, survival into the old stage is certain for all agents. This model includes the survival probability in order to study population aging.

The autarky model in this chapter is very different from that of Chapter 4 or 5. The main purpose of this chapter is study population aging under trade equilibrium. If the models of Chapters 4 and 5 are modified to allow for international trade, the algebra becomes very complicated and the model is intractable to solve. For tractability, a variation of Eicher (1996) is used and this model is suitable to study population aging, education and skill premium under international trade.

6.2.1 Household sector

In an OLG model, each agent lives for at most two periods, young and old. Survival into the old age stage is uncertain, with survival rate equal to s where $s \in [0,1]$. Following the literature such as Ludwig and Vogel (2010) and Irmen (2009), population aging is via an increase in the survival rate s. During each period the number of new born agents is assumed to be constant and normalized to one.

Same as Eicher (1996), all agents are born as unskilled labor, and they can choose to be unskilled labor or do education. If they go to school, when they become old they can either work as teachers in the education sector or as engineers in the production sector.

An unskilled labor works and gets his wage when young. He spends part of his wage on good Y_H and good Y_L, and saves the rest for his old-age consumption.

A representative unskilled labor solves the problem expressed by:

$$\max_{c_{1t}^{UH}, c_{1t}^{UL}, c_{2t+1}^{UH}, c_{2t+1}^{UL}} U_t^U = logC_{1t}^U + \delta \cdot s \cdot logC_{2t+1}^U, \quad \delta \in (0,1); \tag{6.1}$$

$$\text{where} \qquad logC_{1t}^U \equiv logc_{1t}^{UH} + \alpha \cdot logc_{1t}^{UL}; \tag{6.2}$$

$$\text{and} \qquad logC_{2t+1}^U \equiv logc_{2t+1}^{UH} + \alpha \cdot logc_{2t+1}^{UL}; \tag{6.3}$$

$$s.t. \quad p_{Ht} \cdot c_{1t}^{UH} + p_{Lt} \cdot c_{1t}^{UL} + s_t = w_t^U; \tag{6.4}$$

$$p_{Ht+1} \cdot c_{2t+1}^{UH} + p_{Lt+1} \cdot c_{2t+1}^{UL} = \frac{1 + r_{t+1}}{s} s_t. \tag{6.5}$$

The welfare function consists of upper level utility function (6.1) and sub-utility functions (6.2) and (6.3). C_{1t}^U and C_{2t+1}^U are consumption indexes, while c_{1t}^{UH} is young unskilled's consumption of good Y_H during t, and c_{1t}^{UL} is young unskilled's consumption of good Y_L during t, similarly c_{2t+1}^{UH} is old unskilled's consumption of good Y_H during $t+1$, and c_{2t+1}^{UL} is old unskilled's consumption of good Y_L during $t+1$. p_{Ht} and p_{Lt} are prices of good Y_H and Y_L during t, similarly for p_{Ht+1} and p_{Lt+1}. s_t is the saving during t, w_t^U the wage rate during t, r_{t+1} the interest rate from t to $t+1$ and s the survival rate.

Following Blanchard (1985), in this model there exists a perfect annuity market against survival risks of unskilled labor, and the savings of those unskilled labor who fail to survive into the old age are evenly distributed to those unskilled labor who indeed survive (which allows us to use the representative agent method within each generation). This mechanism is reflected by equation (6.5). To understand it, suppose that a continuum of mass 1 of unskilled labor save s_t when young, due to the law of large numbers (which applies here due the continuum of unskilled labor), the mass of survivors is s. Total saving plus interest earned at the pure market rate is equal to $(1 + r_{t+1})s_t$, and this is divided evenly among those mass of s survived unskilled labor, with each survivor getting an amount equal to $\frac{(1+r_{t+1})}{s} s_t$, which is the right hand side of equation (6.5).

A schooling agent (student), during youth, must borrow to finance his young-age consumption and his tuition fee. Upon survival into the old age, he works either as a teacher or an engineer and gets his wage rate. He then pays back his young-age borrowing and consumes the amount left. A representative schooling agent (student) solves the problem expressed by:

$$\max_{c_{1t}^{SH}, c_{1t}^{SL}, c_{2t+1}^{SH}, c_{2t+1}^{SL}} U_t^S = logC_{1t}^S + \delta \cdot s \cdot logC_{2t+1}^S, \quad \delta \in (0,1); \tag{6.6}$$

$$\text{where} \qquad logC_{1t}^S \equiv logc_{1t}^{SH} + \alpha \cdot logc_{1t}^{SL}; \tag{6.7}$$

$$\text{and} \qquad logC_{2t+1}^S \equiv logc_{2t+1}^{SH} + \alpha \cdot logc_{2t+1}^{SL}; \tag{6.8}$$

$$s.t. \quad p_{Ht} \cdot c_{1t}^{SH} + p_{Lt} \cdot c_{1t}^{SL} + z_t = b_t; \tag{6.9}$$

$$p_{Ht+1} \cdot c_{2t+1}^{SH} + p_{Lt+1} \cdot c_{2t+1}^{SL} = w_{t+1}^S - \frac{1 + r_{t+1}}{s} b_t. \tag{6.10}$$

In this specification, z_t is the tuition fee for each student, b_t is the borrowing when young, and w_{t+1}^S is the wage rate for a skilled labor when old (teachers and engineers must have the same wage under equilibrium and this common wage is denoted by w_{t+1}^S). It is assumed that there exists a perfect annuity market against survival risks of students, and the total borrowing of students during t are evenly distributed to those skilled labor surviving into the old age, and this is why s appears in the denominator of equation (6.10).

In this model, a young agent chooses to stay unskilled or go to school. When choosing career path, an agent will compare the lifetime welfare of two career paths, and choose the one yielding the higher lifetime welfare. In equilibrium where there are both unskilled labor and schooling agents, these two career paths must yield the same lifetime welfare.

6.2.2 Education sector

The education sector in this model serves two purposes. First, in this sector students are trained and they become skilled labor when old. Second, in the education process, new technolgy is invented and this implies technological progress. During period t, teachers work with students to generate new technology vintages v_{t+1}. During period $t+1$, v_t and v_{t+1} are used for production. Students during t, upon survival into the old stage, will become skilled labor during $t+1$.

Following Eicher (1996; 1999) the technology progress is modelled as

$$v_{t+1} - v_t = \mu \cdot v_t \cdot min\left(\gamma \cdot P_t, S_t\right), \qquad \gamma > 1, \tag{6.11}$$

where P_t is the number of teachers and S_t is the number of students. Equation (6.11) implies that students and teachers are perfect complements with a fixed ratio γ. Rearranging the above formula yields

$$\frac{v_{t+1}}{v_t} = 1 + \mu \cdot min\left(\gamma \cdot P_t, S_t\right). \tag{6.12}$$

The above equation shows that $\mu \cdot min\left(\gamma \cdot P_t, S_t\right)$ is the growth rate of technology vintages. In the steady state, where the numbers of students and teachers are constant over time, technology is growing at a constant rate.

6.2.3 Production sector

This is the same as Eicher (1999). There are two goods produced for final consumption, Y_H and Y_L. Goods Y_H and Y_L differ in skilled-labor intensity and technology vintage complexity.

During t, good Y_H requires skilled labor 'engineers' H_t, unskilled labor L_t^H, and advanced technology v_t. During t, good Y_L requires only unskilled labor L_t^L and relatively old technology v_{t-1}.[2]

Production functions are expressed as:

$$Y_{Ht} = v_t \cdot (H_t)^\rho \cdot \left(L_t^H\right)^{1-\rho}, \qquad 0 < \rho < 1, \tag{6.13}$$

$$Y_{Lt} = v_{t-1} \cdot \theta \cdot L_t^L, \qquad \theta > 0, \tag{6.14}$$

where H_t is the number of engineers, L_t^H the number of unskilled labor working in Y_H sector, and L_t^L is the number of unskilled labor working in Y_L sector. During each period, the production of Y_H requires the most advanced technology, while the production of Y_L only needs technology from the previous period. Therefore, Y_H is called high-tech and Y_L is called low-tech goods. The production of Y_H needs skilled labor and unskilled labor, while the production of Y_L only needs unskilled labor, hence the high-tech good Y_H is skilled labor intensive and the low-tech good Y_L is unskilled labor intensive.[3] Equations (6.14) and (6.15) are both constant returns to scale, combined with price taking behavior, this means both Y_H and Y_L are perfectly competitive and so the profits are zero in both Y_H and Y_L sectors.

Factor market equilibrium in the labor market implies

$$L_t^H + L_t^L = L_t, \tag{6.15}$$

$$S_{t-1} = \frac{P_t + H_t}{s}, \tag{6.16}$$

$$L_t + S_t = 1. \tag{6.17}$$

where equation (6.15) means, in each period young unskilled labor either work in the high-tech good sector (L_t^H) or the low-tech good sector (L_t^L). Equation (6.16) means that only a

[2]During t, v_t is the relatively advanced technology vintage while v_{t-1} is the relatively old technology, from the previous period.

[3]Production function for the high-tech good Y_H is Cobb-Douglas, with both H and L as inputs, while production of Y_L is linear in L without using any H. This is to emphasize the fact that Y_H is skilled labor intensive while Y_L is unskilled labor intensive. In theory, production of Y_L can also be Cobb-Douglas, and as long as Y_H is skilled labor intensive, the essence of model is unchanged and the results are qualitatively the same. Moreover, production functions for Y_H can be assumed linear in H while Y_L linear in L. Under this change, the skill premium is trivially derived as $\theta \cdot \frac{v_t}{v_{t-1}}$, which is positively related with the growth rate of technology vintages. In the equilibrium where the number of student is constant and increasing in the survival rate, population aging will increase technological growth rate and hence the skill premium, so the results are still qualitatively unchanged.

fraction $s \in (0, 1)$ of students during $t-1$ will survive to t and choose to become teachers (P_t) or engineers(H_t). Equation (6.17) means during each period, all young agents (total mass is 1) either stay as unskilled (L_t) or go to school (S_t).

6.2.4 A brief discussion of the autarky model

In this model, new born agents first choose their career paths: staying unskilled or going to school. Upon the career choice, they choose their consumption plans to maximize expected lifetime welfare, taking income, tuition fee, and good prices as given. If they succeed in surviving into the old age (only a fraction s can survive into old), unskilled labor retire, while skilled labor choose either to be a teacher or engineer. Teachers work in the education sector, educate the students, and create more advanced technology vintages. Engineers work in the high-tech good production sector, with some unskilled labor, using the most advanced technology vintage. Other young unskilled labor work in the low-tech sector, with the relatively old technology vintage (the most advanced in the previous period).

This model is general equilibrium, where all input prices, wages and output prices are endogenous. The equilibrium numbers of students and skilled labor, teachers and engineers are also endogenous. This implies that the growth rate of technology vintages, hence the rate of technological progress, is endogenous. The most important exogenous parameter is the survival rate. A higher survival rate means more people will survive to old, and there are more older people per young, which is a key feature of population aging.

6.3 Autarky equilibrium

This section solves for the autarky equilibrium of the model. Impacts of population aging upon the following variables are studied: equilibrium number of students S; equilibrium relative wage $\frac{w^S}{w^U}$; equilibrium relative price of two consumption goods $\frac{p_L}{p_H}$ and equilibrium growth rate of technology vintage $\left(\frac{v_{t+1}-v_t}{v_t}\right)$.[4]

In order to solve the above variables, this section solves the model from the following five aspects: household utility maximization; education sector equilibrium; production sector profit maximization; no career arbitrage between unskilled and skilled labor and bond market clearing condition.

[4]Reduced-form solutions are given by equations from (6.63) to (6.66).

6.3.1 Household utility maximization

This subsection begins with the household sector. Unskilled labor need to solve the problem expressed by equations (6.1) to (6.5), and their consumption decisions are expressed as

$$
\begin{aligned}
c_{1t}^{UH} =& \frac{1}{(1+\alpha)(1+\delta s)} \frac{w_t^U}{p_{Ht}}, && (6.18)\\
c_{1t}^{UL} =& \frac{\alpha}{(1+\alpha)(1+\delta s)} \frac{w_t^U}{p_{Lt}}, && (6.19)\\
c_{2t+1}^{UH} =& \frac{\delta(1+r_{t+1})}{(1+\alpha)(1+\delta s)} \frac{w_t^U}{p_{Ht+1}}, && (6.20)\\
c_{2t+1}^{UL} =& \frac{\alpha\delta(1+r_{t+1})}{(1+\alpha)(1+\delta s)} \frac{w_t^U}{p_{Lt+1}}, && (6.21)
\end{aligned}
$$

where schooling agents need to solve the problem expressed by equations (6.6) to (6.10), and their consumption decisions are expressed as

$$
\begin{aligned}
c_{1t}^{SH} =& \frac{s}{(1+\alpha)(1+\delta s)(1+r_{t+1})} \left(w_{t+1}^S - \frac{1+r_{t+1}}{s} z_t \right) \frac{1}{p_{Ht}}, && (6.22)\\
c_{1t}^{SL} =& \frac{\alpha s}{(1+\alpha)(1+\delta s)(1+r_{t+1})} \left(w_{t+1}^S - \frac{1+r_{t+1}}{s} z_t \right) \frac{1}{p_{Lt}}, && (6.23)\\
c_{2t+1}^{SH} =& \frac{\delta s}{(1+\alpha)(1+\delta s)} \left(w_{t+1}^S - \frac{1+r_{t+1}}{s} z_t \right) \frac{1}{p_{Ht+1}}, && (6.24)\\
c_{2t+1}^{SL} =& \frac{\alpha\delta s}{(1+\alpha)(1+\delta s)} \left(w_{t+1}^S - \frac{1+r_{t+1}}{s} z_t \right) \frac{1}{p_{Lt+1}}. && (6.25)
\end{aligned}
$$

6.3.2 Education sector equilibrium

As mentioned earlier in section 6.2.2, students and teachers are perfect complements with a fixed ratio γ. In equilibrium, the number of students, S_t, and the number of teachers, P_t, satisfy

$$\gamma \cdot P_t = S_t, \tag{6.26}$$

and using equation (6.26), the technology evolution equation (6.11) becomes

$$v_{t+1} - v_t = \mu v_t S_t. \tag{6.27}$$

During any period t under equilibrium, teachers and engineers have the same wage, and this common wage is denoted w_t^S. It is assumed that teachers' wages are fully funded by students via tuition fees. Given the student-teacher ratio $\gamma > 1$, the tuition fee of each student is given by

$$z_t = \frac{w_t^S}{\gamma}. \tag{6.28}$$

6.3.3 Production sector profit maximization

Under competitive equilibrium, taking the Y_H price and inputs prices as given, the Y_H sector solves the problem expressed by[5]

$$\max_{H_t, L_t^H} \quad p_{Ht} \cdot v_t \cdot (H_t)^\rho \left(L_t^H\right)^{1-\rho} - w_t^S H_t - w_t^U L_t^H, \tag{6.29}$$

and the above problem yields the wage rates for skilled and unskilled labor via the inverse demand functions as

$$w_t^S = p_{Ht} \cdot v_t \cdot \rho \left(\frac{H_t}{L_t^H}\right)^{\rho-1}, \tag{6.30}$$

$$w_t^U = p_{Ht} \cdot v_t \cdot (1-\rho) \left(\frac{H_t}{L_t^H}\right)^{\rho}. \tag{6.31}$$

Similarly, the Y_L sector solves the problem expressed by

$$\max_{L_t^L} \quad p_{Lt} \cdot v_{t-1} \cdot \theta \cdot L_t^L - w_t^U L_t^L, \tag{6.32}$$

and the above problem yields the wage rate for unskilled labor as

$$w_t^U = p_{Lt} \cdot v_{t-1} \cdot \theta. \tag{6.33}$$

In equilibrium, unskilled labor must have the same wage working in the Y_H or Y_L sector, hence in equation (6.31) and (6.33) the same notation w_t^U is used for the common wage rate.

6.3.4 No career arbitrage

In this model, since a young agent chooses between unskilled labor and students, career arbitrage implies these two career choices must yield the same lifetime welfare and this is called no career aritrage condition. According to Eicher (1996, 1999), since agents share identical utility functions, and face identical inter-temporal rates of transformation, unskilled labor and students must have the same total expenditure, both during young and old. In order to obtain the total expenditure when young and old, the savings of unskilled and borrowings of students are solved first, which then are used to solve for total expenditures.

Substituting consumption choices (6.18) and (6.19) into budget constraint (6.4), the saving of unskilled labor during young is

$$s_t = \frac{\delta s}{1+\delta s} w_t^U, \tag{6.34}$$

[5]Due to constant returns to scale in the production function and pricing taking behavior, the Y_H sector has zero profit hence there is no need to specify the ownership of Y_H sector. The same applies for Y_L sector.

and substituting consumption choices (6.22) and (6.23) into budget constraint (6.9), the borrowing of students during young is

$$b_t = \frac{s}{(1+\delta s)(1+r_{t+1})}\left(w_{t+1}^S + \delta(1+r_{t+1})z_t\right). \tag{6.35}$$

Using the saving function of unskilled labor and borrowing function of students, the total expenditures of unskilled labor and students during young and old are denoted by

$$y_{1t}^U \equiv w_t^U - s_t, \tag{6.36}$$

$$y_{2t+1}^U \equiv \frac{1+r_{t+1}}{s}s_t, \tag{6.37}$$

$$y_{1t}^S \equiv b_t - z_t, \tag{6.38}$$

$$y_{2t+1}^S \equiv w_{t+1}^S - \frac{1+r_{t+1}}{s}b_t, \tag{6.39}$$

and since unskilled labor and students have the same total expenditure, both during young and old, the following must hold

$$y_{1t}^U = y_{1t}^S, \tag{6.40}$$

$$y_{2t+1}^U = y_{2t+1}^S. \tag{6.41}$$

Equations (6.40) and (6.41), together with the total expenditure expressions (6.34)-(6.39), yield the relationship between wages of skilled labor and unskilled labor as

$$w_t^U = \frac{s}{1+r_{t+1}}w_{t+1}^S - z_t. \tag{6.42}$$

This equation will be used later in section 6.3.6 to solve for the autarky equilibrium.

6.3.5 Bond market clearing

In this model, unskilled labor save when young for old consumption and students borrow when young to finance their education and young consumption. Borrowings of students and savings of unskilled labor are regulated by a bond market. During each period, bond market clears so the total borrowing is equal to total saving, namely

$$L_t \cdot s_t = S_t \cdot b_t. \tag{6.43}$$

Substituting saving expression (6.34) and borrowing expression (6.35) into the above condition, the bond market clearing interest rate can be expressed as a function of the number of students and unskilled labor:

$$1 + r_{t+1} = \frac{w^S_{t+1}}{\delta s \left(\frac{L_t}{S_t} \cdot w^U_t - z_t \right)}. \tag{6.44}$$

Till now, the five conditions (household utility maximiation, education sector equilibrium, production profit maximization optimization, no career arbitrage and bond market clearing) are solved and represented by equations from (6.18) to (6.44). In the next subsection, these equations will be used to solve for the autarky equilibrium.

6.3.6 Solving for the autarky equilibrium

This subsection uses the results from sections 6.3.1 to 6.3.5 to solve for the autarky equilibrium. The variables of particular interest are the equilibrium student size S, relative wage of skilled to unskilled $\frac{w^s}{w^U}$, relative price of low-tech to high-tech good $\frac{p_L}{p_H}$ and the growth rate of technology vintages $\frac{v_{t+1}-v_t}{v_t}$. This subsection will be mainly formula deviation, which can be skipped when reading, and the important results are completely summarized in Proposition 1.

Combining career arbitrage condition (6.42) and bond market clearing condition (6.44) together, and using equations (6.17) and (6.26), the numbers of students and teachers can be expressed as functions of the relative wage by

$$S_t = \frac{\delta s^2}{1+\delta s^2} \frac{1}{\frac{w^S_t}{\gamma w^U_t} + 1}, \tag{6.45}$$

$$P_t = \frac{\delta s^2}{1+\delta s^2} \frac{1}{\frac{w^S_t}{w^U_t} + \gamma}, \tag{6.46}$$

and substituting equation (6.45) into equilibrium interest rate equation (6.44), the interest rate can be expressed in terms of wage rates as

$$1 + r_{t+1} = \frac{1}{s} \frac{w^S_{t+1}}{w^U_t + \frac{w^S_t}{\gamma}}. \tag{6.47}$$

From equations (6.18) and (6.19), the relative price of good Y_L to good Y_H is expressed as

$$\frac{p_{Lt}}{p_{Ht}} = \frac{\alpha c^{UH}_{1t}}{c^{UL}_{1t}}. \tag{6.48}$$

Similarly, from equations (6.20) and (6.21) applying to an agent born at $t-1$, the relative price of good Y_L to good Y_H is expressed as

$$\frac{p_{LT}}{p_{Ht}} = \frac{\alpha c^{UH}_{2t}}{c^{UL}_{2t}}. \tag{6.49}$$

Combining the above two equations, a basic algebraic property[6] implies

$$\frac{p_{Lt}}{p_{Ht}} = \alpha \cdot \frac{1 \cdot c_{1t}^{UH} + s \cdot c_{2t}^{UH}}{1 \cdot c_{1t}^{UL} + s \cdot c_{2t}^{UL}}. \tag{6.50}$$

Due to equations (6.40) and (6.41) and since unskilled labor and students share the same utility structure, if they are born at the same period, their consumption plans are the same, namely $c_{1t}^{UH} = c_{1t}^{SH}$, $c_{1t}^{UL} = c_{1t}^{SL}$, $c_{2t+1}^{UH} = c_{2t+1}^{SH}$ and $c_{2t+1}^{UL} = c_{2t+1}^{SL}$. In period t, the number of new born agents is unity, and the number of old agents (survived from period $t-1$) is s, hence $1 \cdot c_{1t}^{UH} + s \cdot c_{2t}^{UH}$ is the total consumption of good Y_H during t and $1 \cdot c_{1t}^{UL} + s \cdot c_{2t}^{UL}$ is the total assumption of good Y_L during t.

In autarky, total consumption of each good is equal to domestic production, namely, $1 \cdot c_{1t}^{UH} + s \cdot c_{2t}^{UH} = Y_{Ht}$ and $1 \cdot c_{1t}^{UL} + s \cdot c_{2t}^{UL} = Y_{Lt}$. The above equation (6.50) can be expressed as

$$\frac{p_{Lt}}{p_{Ht}} = \frac{\alpha Y_{Ht}}{Y_{Lt}}. \tag{6.51}$$

Using the production functions for Y_{Ht} and Y_{Lt}, (6.13) and (6.14), yields

$$\frac{Y_{Ht}}{Y_{Lt}} = \frac{v_t}{v_{t-1}} \frac{1}{\theta} \frac{L_t^H}{L_t^L} \left(\frac{H_t}{L_t^H} \right)^{\rho}. \tag{6.52}$$

Putting (6.52) into (6.51) yields

$$\frac{p_{Lt}}{p_{Ht}} = \frac{v_t}{v_{t-1}} \frac{\alpha}{\theta} \frac{L_t^H}{L_t^L} \left(\frac{H_t}{L_t^H} \right)^{\rho}. \tag{6.53}$$

From technology evolution equation (6.27) applied to period from $t-1$ to t, technology vintage growth is expressed as

$$\frac{v_t}{v_{t-1}} = 1 + \mu S_{t-1}. \tag{6.54}$$

Market clearing for the unskilled implied the unskilled must achieve the same wage in Y_H and Y_L sector. Combining (6.31) and (6.33) yields another expression for the relative price as

$$\frac{p_{Lt}}{p_{Ht}} = \frac{v_t}{v_{t-1}} \frac{1-\rho}{\theta} \left(\frac{H_t}{L_t^H} \right)^{\rho}. \tag{6.55}$$

Combining (6.53) and (6.55) yields

$$\frac{L_t^H}{L_t^L} = \frac{1-\rho}{\alpha}. \tag{6.56}$$

[6] If $\frac{w}{x} = \frac{y}{z}$, then $\frac{w}{x} = \frac{w+ay}{x+az}$ for any constant $a \neq -\frac{x}{z}$.

Since during each period the number of new born agents is unity, $L_t^H + L_t^L = L_t = 1 - S_t$. L_t^H and L_t^L can be expressed as the following

$$L_t^H = \frac{1-\rho}{1-\rho+\alpha} L_t = \frac{1-\rho}{1-\rho+\alpha}(1-S_t), \tag{6.57}$$

$$L_t^L = \frac{\alpha}{1-\rho+\alpha} L_t = \frac{\alpha}{1-\rho+\alpha}(1-S_t). \tag{6.58}$$

From the teacher-student number relation (6.26), and equation (6.16), H_t can expressed as

$$H_t = s \cdot S_{t-1} - P_t = s \cdot S_{t-1} - \frac{S_t}{\gamma}. \tag{6.59}$$

Substituting equations (6.56) to (6.59)into the relative price formula (6.55) yields the relative price as a function of student number as

$$\frac{p_{Lt}}{p_{Ht}} = (1+\mu S_{t-1}) \frac{1-\rho}{\theta} \left(\frac{s \cdot S_{t-1} - \frac{S_t}{\gamma}}{\frac{1-\rho}{1-\rho+\alpha}(1-S_t)} \right)^{\rho} \tag{6.60}$$

$$= \frac{(1+\mu S_{t-1})(1-\rho)}{\theta} \left(\frac{(1-\rho+\alpha)\left(s \cdot S_{t-1} - \frac{S_t}{\gamma}\right)}{(1-\rho)(1-S_t)} \right)^{\rho}. \tag{6.61}$$

Combining equations (6.30) and (6.31), the relative wage rate can be expressed as a function of the student number as

$$\frac{w_t^S}{w_t^U} = \frac{\rho(1-S_t)}{(1-\rho+\alpha)\left(s \cdot S_{t-1} - \frac{S_t}{\gamma}\right)}. \tag{6.62}$$

As shown in Eicher (1996) page 136, in equilibrium, $S_{t-1} = S_t \equiv S$ (hereafter all variables without time subscripts mean the equilibrium values of the associated variables). Substituting equation (6.62) into (6.45), the equilibrium student number can be solved. Substituting the equilibrium student number into equation (6.62) gives the equilibrium relative wage, into equation (6.61) gives the relative price of low-tech good to high-tech good, into (6.54) gives the equilibrium growth rate of technology vintage. The results are summarized in the following proposition.

Proposition 1 *In the closed economy described in section 6.2, the equilibrium student number, relative wage of skilled to unskilled labor, relative price of low-tech good Y_L to high-tech good Y_H, and growth rate of technology vintage, are all constant through time. They are* *expressed as*

$$S = \frac{\frac{\delta s^2}{1+\delta s^2} + \frac{\rho}{(1-\rho+\alpha)(1-\gamma\cdot s)}}{1 + \frac{\rho}{(1-\rho+\alpha)(1-\gamma\cdot s)}}, \tag{6.63}$$

$$\frac{w^S}{w^U} = \frac{\rho\gamma}{(1-\rho+\alpha)(\gamma\cdot s - 1)}\frac{1-S}{S}, \tag{6.64}$$

$$\frac{p_L}{p_H} = \left(\frac{1-\rho}{\theta}\left(\frac{1-\rho-\alpha}{1-\rho}\right)^{\rho}\left(s-\frac{1}{\gamma}\right)^{\rho}\right)(1+\mu S)\left(\frac{S}{1-S}\right)^{\rho}, \tag{6.65}$$

$$\left(\frac{v_{t+1}-v_t}{v_t}\right) = \mu S. \tag{6.66}$$

In Proposition 1 all variables of interests are expressed as functions of survival rage s. Population aging is defined via an increase in s, so the results in Proposition 1 allows to study how population aging affects the above variables. Having solved the autarky equilibrium as in the above proposition, the next section investigates the impacts of population aging.

6.4 The impacts of population aging under autarky

In this model, population aging is defined as an increase in the survival rate s. Proposition 1 presents the autarky equilibrium values of student number, skill premium, relative price of low-tech to high-tech goods and growth rate of technology vintages. To study the impacts of population aging, comparative statistic analysis will be conducted, with respect to the survival probability s.

As shown in the appendix, the equilibrium student number S expressed in (6.63) is increasing in s, namely, population aging will increase overall educational effort. With S increasing in s, it can be seen that relative wage of skilled labor versus unskilled labor(expressed in (6.64)) is decreasing in s, while the relative price of Y_L versus Y_H and the growth rate of technology vintage (expressed in (6.65) and (6.66)) are increasing in s. The results are summarized in the following proposition.

Proposition 2 *Under autarky equilibrium, population aging, defined as an increase in the survival probability s, will:*

- *Increase the equilibrium student number S;*
- *Decrease the equilibrium relative wage of skilled to unskilled labor $\frac{w^S}{w^U}$;*
- *Increase the equilibrium relative price of Y_L to Y_H $\frac{p_L}{p_H}$;*

- *Increase the equilibrium growth rate of technology vintage μS.*

Discussion of Proposition 2

Proposition 2 shows that population aging defined as a higher survival rate encourages educational effort. With population aging via an increase in the survival rate, students are more likely to survive to the old age. Since students only get payoff during the old age, a higher survival probability increases their expected reward and therefore their incentives to do education are also higher. This explains why the equilibrium student number is increasing in the survival rate.

As the number of students increases, the number of unskilled labor will decrease, and this decreases the relative scarcity of skilled labor, which in turn decreases the relative wage of skilled labor. In the two production sectors, low-tech good Y_L requires more unskilled labor while high-tech good Y_H requires more skilled labor. When the relative wage of skilled labor decreases, the relative input price of Y_L to Y_H increases, and this translates into the higher relative output price of Y_L. Lastly, since student number is increasing in the survival rate and technology vintage growth is a positive byproduct of the education sector, the growth rate is positively related with student number, hence also increasing in the survival rate.

More discussion about the impacts of population aging will be done in section 6.6, where the impacts of population aging under autarky and under trade are compared.

6.5 Two-country trade model

In this section, the autarky model in section 6.2 is extended to an international trade model with two countries, Home and Foreign. The purpose is to study the impacts of population aging under trade equilibrium and compare the impacts of aging under autarky and under trade.

6.5.1 Trade model description

The model structure of both Home and Foreign is the same as under autarky. In Home, the household, education and production sectors are still represented by equations (6.1) to (6.17). In Foreign, the household, education and production sectors are represented by equations (6.1) to (6.17) and all variables are with a superscript *. Moreover, the household utility maximization (equations (6.18) to (6.25)), education sector equilibrium (equations (6.26) to (6.28)), production sector profit maximization (equations (6.29) to (6.33)), no career arbitrage (equations (6.34) to (6.42)) and bond market clearing (equations (6.43) to (6.44)) still apply to both Home and Foreign, with all Foreign variables with a superscript *. Equations from (6.45) to (6.49) still apply to both Home and Foreign (all Foreign variables with a superscript *).

The survival rates in Home and Foreign are s and s^* respectively. Home and Foreign are assumed to be identical in every aspect except for the survival rates, denoted by s for Home and s^* for Foreign. It is assumed that $s > s^*$, hence during each period the fraction of older agents is higher in Home than in Foreign. Home can be interpreted as more developed than Foreign, and population aging is more severe in Home. The purpose of the international trade model is to study the impacts of population aging (upon educational efforts and skill premiums and so on), and therefore, countries' differences in other aspects are assumed away and the only focus is on the different survival rates.

Home and Foreign are initially both in autarky, reaching their respective autarky equilibrium, and Propositions 1 and 2 apply to both of them (with Foreign variables all denoted with *). Then, Home and Foreign open to free international trade with each other in goods Y_H and Y_L. If there are any cross-country price differences in Y_H or Y_L, trade will occur. Free trade (no tariffs or other government regulations) is assumed and there is no transportation cost, for simplicity.

The trade model is general equilibrium, with all input prices, wages, and output prices determined endogenously. The prices of Y_L and Y_H are determined endogenously during international trade in the world market. In trade equilibrium, the total world consumption of Y_L (Y_H) is equal to the total world production of Y_L (Y_H) and this will determine prices of Y_L (Y_H). Output prices in turn determines the domestic wage rates in each country. Taking output prices and wage rates as given, agents make consumption plans and career decisions, which determine equilibrium student numbers and the rates of technological progress in each country.

Technology spillover

In the trade model, some simplifying assumptions on how the technology vintages are determined in Foreign are imposed. In Home, the survival rate is higher than that in Foreign.

According to Proposition 2, Home has more students and the rate of technological progress is higher in Home. As a result, Home has more advanced technology than Foreign. Moreover, Proposition 2 implies that the relative price of Y_H to Y_L is cheaper in Home. When Home and Foreign trade, Home will export Y_H and import Y_L.

During period t, technology vintage v_t is embodied in Y_H and exported to Foreign. It is assumed that during t, the Foreign can study the technology v_t in the Foreign education sector. The Foreign education sector can work with its own technology v^*_{Ht} and/or technology from Home v_t. The research outcome of education sector during t can be used for production during $t+1$. Moreover, since Home technology v_t is studied in Foreign during t, it is assumed the Foreign can use v_t for production during $t+1$.

During t, because $v_t > v^*_{Ht}$, the Foreign education sector will study and work on v_t, instead of v^*_{Ht}. Then during $t+1$, Foreign can use technology vintage $v^*_{Ht+1} = (1+\mu S^*_t) \cdot v_t$ for production of Y_H. During $t+1$, Foreign will use v_t to produce Y_L. There is no technology spillover from Foreign to Home, since the Home has more advanced technology and has no benefit from studying Foreign's technology.

In trade equilibrium, the focus is on the following endogenous variables,

1. Equilibrium number of student in each country;
2. Equilibrium relative wage in each country;
3. Equilibrium relative price in each country;
4. Equilibrium growth rate of technology vintage in each country.

The trade model does not give reduced forms solutions of the above variables. However, comparative statistic analysis can be conducted, with respect to the survival rates in both countries. In this sense, the impacts of aging upon the above variables can be studied. The subsection below will (1) solve the trade model and conduct comparative statistic analysis of the above variables with respect to s and s^*; and (2) compare the comparative statistic analysis in trade equilibrium with that in autarky equilibrium.

6.5.2 Solving the trade model

The basic model structure, expressed by equations (6.1) to (6.17), is still the same both in Home and Foreign, with all Foreign variables denoted by *. Accordingly, equations (6.18) to

(6.50) remain valid, for each country.

The procedures to solve for the trade equilibrium are exactly the same as the autarky equilibrium in subsection, and the only difference is in the market clearing condition for the consumption goods. Starting from equation (6.50) which still applies to Home, the Foreign counterpart of equation (6.50) is

$$\frac{p^*_{Lt}}{p^*_{Ht}} = \alpha \cdot \frac{1 \cdot c^{UH*}_{1t} + s^* \cdot c^{UH*}_{2t}}{1 \cdot c^{UL*}_{1t} + s^* \cdot c^{UL*}_{2t}}. \tag{6.67}$$

With free trade and no transportation costs, and assume partial specialization in both country, it follows that

$$p_{Ht} = p^*_{Ht}, \tag{6.68}$$

$$p_{Lt} = p^*_{Lt}. \tag{6.69}$$

Combining equations (6.50), (6.67)-(6.69) gives

$$\frac{p_{Lt}}{p_{Ht}} = \alpha \cdot \frac{1 \cdot c^{UH}_{1t} + s \cdot c^{UH}_{2t} + 1 \cdot c^{UH*}_{1t} + s^* \cdot c^{UH*}_{2t}}{1 \cdot c^{UL}_{1t} + s \cdot c^{UL}_{2t} + 1 \cdot c^{UL*}_{1t} + s^* \cdot c^{UL*}_{2t}}. \tag{6.70}$$

With international trade in good Y_H and Y_L, total domestic consumption of each good is no longer equal to total domestic production. Instead, total world consumption is equal to total world production, and it follows that

$$1 \cdot c^{UH}_{1t} + s \cdot c^{UH}_{2t} + 1 \cdot c^{UH*}_{1t} + s^* \cdot c^{UH*}_{2t} = Y_{Ht} + Y^*_{Ht}, \tag{6.71}$$

$$1 \cdot c^{UL}_{1t} + s \cdot c^{UL}_{2t} + 1 \cdot c^{UL*}_{1t} + s^* \cdot c^{UL*}_{2t} = Y_{Lt} + Y^*_{Lt}. \tag{6.72}$$

Combining equations (6.70)-(6.72) yields

$$\frac{p_{Lt}}{p_{Ht}} = \alpha \cdot \frac{Y_{Ht} + Y^*_{Ht}}{Y_{Lt} + Y^*_{Lt}}, \tag{6.73}$$

and this equation is the international trade counterpart of equation (6.51). Solving the trade model, in a manner similar to the autarky model (steps from equations (6.51) to (6.63)), gives

$$\frac{p_{Lt}}{p_{Ht}} = \alpha \frac{v_t \cdot (H_t)^\rho \cdot (L^H_t)^{1-\rho} + v^*_{Ht} \cdot (H^*_t)^\rho \cdot (L^{H*}_t)^{1-\rho}}{v_{t-1} \cdot \theta \cdot (L^L_t) + v^*_{Lt} \cdot \theta \cdot (L^{L*}_t)} = \frac{v_t}{v_{t-1}} \frac{1-\rho}{\theta} \left(\frac{H_t}{L^H_t}\right)^\rho, \tag{6.74}$$

$$\frac{v_t}{v_{t-1}} \frac{1-\rho}{\theta} \left(\frac{H_t}{L^H_t}\right)^\rho = \frac{v^*_{Ht}}{v^*_{Lt}} \frac{1-\rho}{\theta} \left(\frac{H^*_t}{L^{H*}_t}\right)^\rho, \tag{6.75}$$

$$L^H_t + L^L_t = 1 - S_t, \tag{6.76}$$

$$L^{H*}_t + L^{L*}_t = 1 - S^*_t, \tag{6.77}$$

$$H_t = s \cdot S_{t-1} - \frac{S_t}{\gamma}, \tag{6.78}$$

$$H^*_t = s^* \cdot S^*_{t-1} - \frac{S^*_t}{\gamma}, \tag{6.79}$$

where $v^*_{Ht}(v^*_{Lt})$ denotes the technology vintage used to produce $Y_H(Y_L)$ during t in Foreign.

The first equality of (6.74) is obtained by putting production functions for Y_H and Y_L (equations (6.13), (6.14) and their Foreign counterparts) into equation (6.73). The second equality of (6.74) is obtained by combining (6.31) and (6.33). Combining (6.31), (6.33), their Foreign counterparts, and (6.68) and (6.69), gives equation (6.75). (6.76) is obtained by (6.15) and (6.17), with (6.77) as its Foreign counterpart. (6.78) is rewriting (6.59), with (6.79) as its Foreign counterpart.

The first step is to solve for six variables, L^H_t, L^L_t, L^{H*}_t, L^{L*}_t, H_t and H^*_t as functions of S_t, S_{t-1}, S^*_t and S^*_{t-1}. The second step is to solve the equilibrium values of S_t, S_{t-1}, S^*_t and S^*_{t-1}.

With the above technology spillover assumption in 6.5.2, it follows that $v^*_{Lt} = v_{t-1}$ and $v^*_{Ht} = (1 + \mu S^*_t) \cdot v_{t-1}$, then after some algebra, equations (6.74) and (6.75) become

$$L^H_t + L^{H*}_t = \frac{1-\rho}{\alpha}(L^L_t + L^{L*}_t), \tag{6.80}$$

$$(1 + \mu \cdot S_t)\left(\frac{H_t}{L^H_t}\right)^{\rho} = (1 + \mu \cdot S^*_t)\left(\frac{H^*_t}{L^{H*}_t}\right)^{\rho}. \tag{6.81}$$

Using (6.76)-(6.79), (6.80) and (6.81), L^H_t, L^L_t, L^{H*}_t, L^{L*}_t, H_t and H^*_t can be solved as functions of S_t, S_{t-1}, S^*_t and S^*_{t-1}. For simplicity, only L^H_t and L^{H*}_t (expressions for other four variables are not used below) are presented below, as functions of S_t, S_{t-1}, S^*_t and S^*_{t-1}, as

$$L^H_t = \frac{1-\rho}{1+\alpha-\rho}(2 - S_t - S^*_t)\left(1 + \left(\frac{1+\mu S^*_t}{1+\mu S_t}\right)^{1/\rho} \cdot \frac{s^* \cdot S^*_{t-1} - S^*_t/\gamma}{s \cdot S_{t-1} - S_t/\gamma}\right)^{-1}, \tag{6.82}$$

$$L^{H*}_t = \frac{1-\rho}{1+\alpha-\rho}(2 - S_t - S^*_t)\frac{\left(\frac{1+\mu S^*_t}{1+\mu S_t}\right)^{1/\rho} \cdot \frac{s^* \cdot S^*_{t-1} - S^*_t/\gamma}{s \cdot S_{t-1} - S_t/\gamma}}{1 + \left(\frac{1+\mu S^*_t}{1+\mu S_t}\right)^{1/\rho} \cdot \frac{s^* \cdot S^*_{t-1} - S^*_t/\gamma}{s \cdot S_{t-1} - S_t/\gamma}}. \tag{6.83}$$

Combining (6.82) and (6.78) gives

$$\frac{H_t}{L^H_t} = \frac{1+\alpha-\rho}{1-\rho} \cdot \frac{1}{2 - S_t - S^*_t} \cdot \left((s \cdot S_{t-1} - \frac{S_t}{\gamma}) + \left(\frac{1+\mu \cdot S^*_t}{1+\mu \cdot S_t}\right)^{1/\rho} \cdot (s^* S^*_{t-1} - \frac{S^*_t}{\gamma})\right). \tag{6.84}$$

Using (6.30) and (6.31), the relative wage can be expressed as

$$\frac{w^S_t}{w^U_t} = \frac{\rho}{1-\rho} \cdot \frac{L^H_t}{H_t}. \tag{6.85}$$

Putting equation (6.85) into equation (6.45) gives

$$\frac{1+\delta s^2}{\delta s^2} \cdot S_t = \frac{\gamma}{\frac{\rho}{1-\rho} \cdot \frac{L_t^H}{H_t} + \gamma}. \tag{6.86}$$

Combining (6.84) and (6.86) gives

$$\frac{(1+\alpha-\rho)\gamma}{(2-S_t-S_t^*)\rho} \left(\left(s \cdot S_{t-1} - \frac{S_t}{\gamma} \right) + \left(\frac{1+\mu S_t^*}{1+\mu S_t} \right)^{1/\rho} \left(s^* \cdot S_{t-1}^* - \frac{S_t^*}{\gamma} \right) \right) \left(\frac{\delta s^2}{(1+\delta s^2) S_t} - 1 \right) = 1. \tag{6.87}$$

Repeating the above process for the Foreign, the Foreign counterpart of equation (6.87) can be expressed as

$$\frac{(1+\alpha-\rho)\gamma}{(2-S_t-S_t^*)\rho} \left(\left(s^* \cdot S_{t-1}^* - \frac{S_t^*}{\gamma} \right) + \left(\frac{1+\mu S_t}{1+\mu S_t^*} \right)^{1/\rho} \left(s \cdot S_{t-1} - \frac{S_t}{\gamma} \right) \right) \left(\frac{\delta (s^*)^2}{(1+\delta (s^*)^2) S_t^*} - 1 \right) = 1. \tag{6.88}$$

Similar to autarky equilibrium, in trade equilibrium, $S_t = S_{t-1} \equiv S$ and $S_t^* = S_{t-1}^* \equiv S^*$, and (6.87) and (6.88) become

$$\frac{(1+\alpha-\rho)\gamma}{(2-S-S^*)\rho} \left(\left(s \cdot S - \frac{S}{\gamma} \right) + \left(\frac{1+\mu S^*}{1+\mu S} \right)^{1/\rho} \cdot \left(s^* \cdot S^* - \frac{S^*}{\gamma} \right) \right) \left(\frac{\delta s^2}{(1+\delta s^2) S} - 1 \right) = 1, \tag{6.89}$$

$$\frac{(1+\alpha-\rho)\gamma}{(2-S-S^*)\rho} \left(\left(s^* \cdot S^* - \frac{S^*}{\gamma} \right) + \left(\frac{1+\mu S}{1+\mu S^*} \right)^{1/\rho} \cdot \left(s \cdot S - \frac{S}{\gamma} \right) \right) \left(\frac{\delta (s^*)^2}{(1+\delta (s^*)^2) S^*} - 1 \right) = 1. \tag{6.90}$$

The above two equations determine the equilibrium student numbers, S and S^* as functions of survival rates s and s^* and other exogenous variables.

6.5.3 Impacts of population aging in international trade equilibrium

The objective is to conduct comparative statistic analysis of S and S^*, with respect to s and s^*, to study the impacts of aging upon educational efforts.

Dividing the previous equations (6.89)-(6.90) gives

$$\frac{\left(s \cdot S - \frac{S}{\gamma} \right) + \left(\frac{1+\mu S^*}{1+\mu S} \right)^{1/\rho} \cdot \left(s^* \cdot S^* - \frac{S^*}{\gamma} \right)}{\left(s^* \cdot S^* - \frac{S^*}{\gamma} \right) + \left(\frac{1+\mu S}{1+\mu S^*} \right)^{1/\rho} \cdot \left(s \cdot S - \frac{S}{\gamma} \right)} \cdot \frac{\frac{\delta s^2}{(1+\delta s^2) S} - 1}{\frac{\delta (s^*)^2}{(1+\delta (s^*)^2) S^*} - 1} = 1. \tag{6.91}$$

In the appendix 6.8.2, equation (6.91) is used to investigate the impacts of aging upon student number in both Home and Foreign. Leaving details in the appendix 6.8.2, the following proposition is presented and its discussion is left to section 6.6.

Proposition 3 *Under international trade between Home and Foreign, if population aging becomes more severe in Home, namely s increases, the equilibrium student number in Home increases, while the equilibrium student number in Foreign decreases. On the other hand, if population aging becomes more severe in Foreign, namely s^* increases, the equilibrium student number in Home decreases, while the equilibrium student number in Foreign increases. Formally:*

- $\frac{\partial S}{\partial s} > 0$, $\frac{\partial S^*}{\partial s} < 0$, $\frac{\partial S}{\partial s^*} < 0$ *and* $\frac{\partial S^*}{\partial s^*} > 0$.

Due to equation (6.66) and its Foreign counterpart, the equilibrium growth rate of technology vintage is μS in Home and μS^* in Foreign. The above proposition then implies the following proposition and its discussion is left to section 6.6.

Proposition 4 *Under international trade between Home and Foreign,*

1. *equilibrium growth rate of technology vintage in Home increases with s and decreases with s^*;*

2. *equilibrium growth rate of technology vintage in Foreign decreases with s and increases with s^*.*

The analysis below investigates the impacts of aging upon (1) relative wage of skilled and unskilled labor and (2) relative price of Y_L and Y_H.

$\frac{w_t^S}{w_t^U}$ in Home can be obtained from equations (6.84) and (6.85). From equation (6.74), the equilibrium relative price in Home can be expressed as

$$\frac{p_L}{p_H} = \frac{1-\rho}{\theta} \cdot (1+\mu S) \cdot \left(\frac{H}{L^H}\right)^{\rho}. \tag{6.92}$$

where $\frac{H}{L^H}$ is obtained from (6.84) with all variables evaluated at the equilibrium values as

$$\frac{H}{L^H} = \frac{1+\alpha-\rho}{1-\rho} \cdot \frac{1}{2-S-S^*} \cdot \left((s \cdot S - \frac{S}{\gamma}) + \left(\frac{1+\mu \cdot S^*}{1+\mu \cdot S}\right)^{1/\rho} \cdot (s^* S^* - \frac{S^*}{\gamma}) \right). \tag{6.93}$$

Since the equilibrium relative wage and relative price both depend on $\frac{H}{L^H}$, the analysis below studies how $\frac{H}{L^H}$ is affected by s and s^*. Leaving details in the appendix 6.8.3, the results are presented in the following proposition and its discussion is left to section 6.6.

Proposition 5 *Under international trade between Home and Foreign, suppose population aging becomes more severe in Home, namely s increases, then in both the Home and the*

Foreign country, the equilibrium ratio of engineers to unskilled labor working in high-tech sector decreases, and the relative wage of skilled to unskilled labor increases, and the relative price of low-tech good Y_L decreases. If population aging becomes more severe in Foreign, namely s^* increases, then in both the Home and the Foreign country, the equilibrium ratio of engineers to unskilled labor working in high-tech sector decreases, and the relative wage of skilled to unskilled labor increases, and the relative price of low-tech good Y_L decreases. Formally,

- $\frac{\partial H/L^H}{\partial s} < 0$, $\frac{\partial w^S/w^U}{\partial s} > 0$, $\frac{\partial H^*/L^{H*}}{\partial s} < 0$, $\frac{\partial w^{S*}/w^{U*}}{\partial s} > 0$, *and* $\frac{\partial p_L/p_H}{\partial s} < 0$;
- $\frac{\partial H/L^H}{\partial s^*} < 0$, $\frac{\partial w^S/w^U}{\partial s^*} > 0$, $\frac{\partial H^*/L^{H*}}{\partial s^*} < 0$, $\frac{\partial w^{S*}/w^{U*}}{\partial s^*} > 0$, *and* $\frac{\partial p_L/p_H}{\partial s^*} < 0$.[7]

The discussions of Propositions 3 to 5 are in the next section, where the impacts of population aging under autarky and trade are compared.

6.6 Discussions of Propositions 3 to 5

The impacts of population aging under autarky are presented in Proposition 2, while those under trade are presented in Propositions 3 to 5. This section compares the impacts of population aging under autarky and trade equilibrium. As it turns out, the impacts of population aging are different under autarky versus under trade, and introducing international trade can give more insights understanding population aging.

Propositions 2 to 5 jointly give the following results concerning the impacts of population aging, in terms of (i) student number, (2) rate of technological progress, (3) skill premium, and (4) relative price of two consumption goods:

i. Student number: Population aging increases equilibrium student number under autarky; under trade population aging in one country increases student number in its own country, while decreases student number in the other country;

ii. Equilibrium growth rate of technology vintages: Under autarky, population aging increases the equilibrium growth rate of technology vintages; under trade population aging in one country increases technology vintage growth rate in its own country, while decreases that in the other country

[7] Under free trade, the prices of Y_H and Y_L are equal across Home and Foreign, hence superscript * is not added to p_L or p_H.

iii. Relative wage of skilled to unskilled labor (skill premium): Population aging decreases skill premium under autarky; under trade population aging in any country increases skill premium in both countries;

iv. Relative price of Y_L to Y_H p_L/p_H: Population aging increases p_L/p_H under autarky; under trade population aging in any country decreases p_L/p_H in both countries.

i. Student number

In autarky, as the survival rate goes up, the expected future payoff of being a student goes up, hence there are more agents doing education after population aging. In trade equilibrium, if Home survival rate goes up, the expected future payoff of being a student goes up hence the student number in Home goes up, as in autarky. However, in Foreign, since it exports the low-tech good Y_L due to cross-country price differences, more students in Home will reinforce the initial cross-country differences and increase the benefit for Foreign to specialize more in Y_L. Because the production of low-tech good Y_L only requires unskilled labor, more production of Y_L calls for more unskilled labor and result in less students in Foreign. Therefore, in trade, a higher survival rate in Home will discourage education in Foreign ('education-stealing effect'). On the other hand, if the survival rate in Foreign goes up (s^* increases), the incentive to do education in Foreign goes up, and this will decrease the initial cross-country differences between Home and Foreign. As a result, the opposite happens, resulting in higher educational efforts in Foreign and lower educational efforts in Home (Foreign 'steals' education from Home).

ii. Growth rates of technology vintages

In this model, technological progress is through development of new technology vintages. Technology vintage development is a positive byproduct of the education sector, hence if there are more agents doing education, the growth of technology vintages and technological progress is faster. This, together with the impacts of population aging upon student number, explains why population aging in one country will faster its own technological progress, while slows down technological progress in the other country.

iii. Relative wage of skilled to unskilled labor

The impacts of population aging upon skill premium are exactly opposite under autarky versus under trade. While population *decreases* skill premium under autarky, it *increases* skill premium (in both Home and Foreign) under trade. Under autarky, after s goes up, overall educational efforts go up and there are less unskilled labor in Home. This increases the relative scarcity of unskilled labor which in turn increases the relative wage of unskilled labor. Under trade, after s goes up in Home, there are less unskilled labor in Home. However, this does not increase the relative scarcity of unskilled labor, since now Home shifts to produce more of high-tech good Y_H and less of low-tech good Y_L. In fact, a higher s makes Home specialize so much in Y_H that it *increases* the relative scarcity of skilled labor.[8] This can be seen from Proposition 5, where the ratio of skilled labor to unskilled labor in the Y_H sector, $\frac{H}{L^H}$, is decreasing in s, hence skilled labor becomes more scarce after population aging. Due to the more scarcity of skilled labor in Y_H sector, the relative wage of skilled labor, namely skill premium, goes up. At the same time in Foreign, after s goes up, there are more unskilled labor, and since Foreign produces more of low-tech good, more unskilled labor decreases the relative scarcity of unskilled labor in Foreign, which in turn increases the relative wage of skilled labor, namely skill premium, in Foreign. A similar logic applies to the case where s^* goes up.

In closed economy, aging will decrease skill premium. In contrast, in trade equilibrium, population aging in any country will increase skill premium for all countries in the world. The autarky effect of aging upon skill premium is inconsistent with empirical evidence, and it is no surprise since this model has neither capital-complementarity (there is no physical capital in production) or skill-biased technological change.[9], which is considered the reason for skill premium to rise.

Without skill-biased technical change, the impact of trade upon skill premium is not to induce skill-biased technical change, but rather, via factor and production reallocation, as in classic trade models. In this sense, this chapter has shown that standard two-country two-good trade model can explain the increase in skill premium to a much better extend than

[8]In this model, the production of low-tech good Y_L does not need skilled labor at all, so Y_H has extreme skilled-labor intensity.

[9]This can be seen from equations (6.30) and (6.31). Dividing the two equations gives skill premium, which does not relate with technology vintage (v_t) but only depends on factor ratio H/L^H. This means that this model has ruled out skill-biased technical change.

currently believed.

This chapter rules out skill-biased technical change, but the model can still explain the changes in skill premium quite well. In standard skill-biased technical change literature, Acemoglu and Autor (2009) either assumes the supply of skilled and unskilled labor to be exogenous, or endogenously determined via education. However, Acemoglu and Autor (2009) have not explicitly modeled how population aging encourages human capital accumulation. The current literature on aging and education has found that population aging is a key determinant of the increase in human capital accumulation via education. Hence, even with skill-biased technical change, population aging is an essential driver for the change in skill premium.

iv. Relative price of Y_L to Y_H

The prices of Y_L and Y_H depend on input prices, which are the wages of skilled and unskilled labor. Since Y_L is unskilled labor intensive (Y_L only requires unskilled labor), a higher relative wage of skilled labor will drive down the relative price of Y_L. This, together with the above intuition for relative wages, explains why the impacts of population aging differ under autarky and trade equilibrium.

Empirically, the argument that trade should increase the relative price of skill-intensive goods which raises the derived demand for skills, is inconsistent with observed evidence, where the relative price of skill-intensive goods stay constant or declining (Acemoglu, 2003). However, in the seminal work of Acemoglu (2003), he mentions that international trade will increase the relative price of skill-intensive goods (same prediction as the classic trade model), which in turn encourages skill-biased technical change. He somehow solves the conflicts between theory and data by claiming 'the increased productivity of skilled workers both in the U.S. and in other countries may eventually return the relative price of skill-intensive goods to its original level in the U.S. So existing evidence...does not refute trade-based explanations.' This argument can be applied here, which solves the conflict between theory and empirical evidence.

6.7 Conclusion

Population aging, characterized by a remarkable increase in the fraction of old agents, has become the most important demographic shock to many countries. This chapter uses an overlapping generation model with all agents live for at most two period, young and old. Survival to the old age is uncertain. Population aging is modeled via an increase in the survival rate. The impacts of population aging upon educational efforts, technological progress, and skill premium are investigated, in both autarky and two-country trade equilibrium.

This chapter combines (i) population aging, (ii) education, (iii) endogenous economic growth and (iv) international trade all together in a general equilibrium.

Population aging encourages domestic educational efforts. After population aging via a higher survival rate, the expected value of future payoff from education is higher, and this leads to a higher incentive to do education. Moreover, in the two-country international trade equilibrium, population aging in any country will discourage educational effort in the other country. In this sense, population aging 'steals' education from the other country, *education stealing effect*.

Second, this chapter gives some important results related to the skill premium literature. In autarky, after population aging, there are more skilled labor compared to unskilled labor, and this increase in the relative supply of skilled labor decreases skill premium. However, this chapter gets the opposite result in trade equilibrium. In trade equilibrium, population aging in any country will increase skill premium of both countries, via production reallocation. After population aging in any country, there are more skilled labor and this country will produce more of high-tech (skilled labor intensive) goods and less of low-tech (unskilled labor intensive) goods, and this leads to an increase in the relative demand of skilled labor, hence raising skill premium. In the other country, there are less skilled labor (because of education stealing effect) and the relative supply of unskilled labor will rise, which leads to a decline of the relative wage of unskilled labor, hence a rise in skill premium.

There are several possibilities for future research. First, in this model I have used Cobb-Douglas forms in utility functions and production functions of two goods, for tractability. One possible generalization is to use constant elasticity of substitution function forms. This can allow for a general degree of substitution elasticity between good consumptions in utility and input employments in production. Second, in this model, population aging is via an increase in the survival probability. In a broader sense, population aging is a combination of a higher survival rate and a reduction in fertility. Population growth rate can be introduced into the model, and population aging is associated with a lower population growth rate. Third, physical capital can be added to the production functions, and this allows for a form of technology-capital complementarity. If capital substitutes more for unskilled labor than for skilled labor, then the conjecture is that technical change should raise skill premium under autarky as well as trade equilibrium. These extension possibilities are left for future research.

6.8 Appendix

6.8.1 Proof of $\frac{\partial S}{\partial s} > 0$ in Proposition 2

The autarky equilibrium student number is expressed by equation (6.63), reproduced here as

$$S = \frac{\frac{\delta s^2}{1+\delta s^2} + \frac{\rho}{(1-\rho+\alpha)(1-\gamma\cdot s)}}{1 + \frac{\rho}{(1-\rho+\alpha)(1-\gamma\cdot s)}}. \tag{6.63}$$

The above equation expressed as

$$S = \frac{g(s) + h(s)}{1 + h(s)} \tag{6.63'}$$

where

$$g(s) \equiv \frac{\delta s^2}{1+\delta s^2} = \frac{1}{1+\frac{1}{\delta s^2}},$$
$$h(s) \equiv \frac{\rho}{(1-\rho+\alpha)(1-\gamma\cdot s)}.$$

It is easy to see that $g'(s) > 0$ and $h'(s) > 0$ (the denominator of $h(s)$ is decreasing in s) and $0 < g(s) < 1$. Differentiating equation (6.63) with respect to s yields

$$\frac{\partial S}{\partial s} = \frac{g'(s)(1+h(s)) + (1-g(s))h'(s)}{(1+h(s))^2} > 0.$$

6.8.2 Solving the signs of $\frac{\partial S}{\partial s}$, $\frac{\partial S}{\partial s^*}$, $\frac{\partial S^*}{\partial s}$, $\frac{\partial S^*}{\partial s^*}$ in Proposition 3

Equation (6.91) is reproduced below

$$\underbrace{\frac{\left(s\cdot S - \frac{S}{\gamma}\right) + \left(\frac{1+\mu S^*}{1+\mu S}\right)^{1/\rho}\cdot\left(s^*\cdot S^* - \frac{S^*}{\gamma}\right)}{\left(s^*\cdot S^* - \frac{S^*}{\gamma}\right) + \left(\frac{1+\mu S}{1+\mu S^*}\right)^{1/\rho}\cdot\left(s\cdot S - \frac{S}{\gamma}\right)}}_{\text{Part I}} \cdot \underbrace{\frac{\frac{\delta s^2}{(1+\delta s^2)S} - 1}{\frac{\delta (s^*)^2}{(1+\delta (s^*)^2)S^*} - 1}}_{\text{Part II}} = 1. \tag{6.91}$$

It follows that

$$\text{Part I of (6.91)} = \frac{1 + B/A}{A + B^{-1}} = \frac{1 + BA}{1 + BA} \cdot B = B \equiv \left(\frac{1 + \mu S^*}{1 + \mu S} \right)^{1/\rho} \tag{6.94}$$

where $A \equiv \frac{s^* \cdot S^* - \frac{S^*}{\gamma}}{s \cdot S - \frac{S}{\gamma}}$ and (6.94) implies

- $\frac{\partial log(\text{Part I of (6.91)})}{\partial s} = \frac{\partial log(\text{Part I of (6.91)})}{\partial s^*} = 0$;
- $\frac{\partial log(\text{Part I of (6.91)})}{\partial S} < 0$;
- $\frac{\partial log(\text{Part I of (6.91)})}{\partial S^*} > 0$.

Now Part II of (6.91) is analysed. s and S only enter the numerator, while s^* and S^* only enter the denominator of part II of (6.91). It is easy to see that the numerator of part II of (6.91) is increasing in s and decreasing in S, while the denominator of part II of (6.91) is increasing in s^* and decreasing in S^*. Therefore, the following results hold

- $\frac{\partial(\text{Part II of (6.91)})}{\partial s} > 0$;
- $\frac{\partial(\text{Part II of (6.91)})}{\partial s^*} < 0$;
- $\frac{\partial(\text{Part II of (6.91)})}{\partial S} < 0$;
- $\frac{\partial(\text{Part II of (6.91)})}{\partial S^*} > 0$.

Combing the above results, comparative statistic analysis can be conducted. The purpose of comparative statistic analysis is to find whether S and S^* are increasing or decreasing in s and s^*. The usual way to do comparative statistic analysis is to find the sign of the derivatives, but here the expressions are messy so a simpler method is used. Namely, by looking at how part I and part II of (6.91) change with s and s^*, how S and S^* change with s and s^* can be determined and this is done in the next paragraph.

If s increases, part II of (6.91) increases, while part I is unchanged. For (6.91) to hold again, either S increases, or S^* decreases. Similarly, if s^* increases, part II of (6.91) decreases, while part I is unchanged. For (6.91) to hold again, either S decreases, or S^* increases. To sum up, S is increasing in s, decreasing in s^*, while S^* is decreasing in s and increasing in s^*.

6.8.3 Derivations of the comparative statistics results in Proposition 5

From equation (6.85), w^S/w^U is negatively related with H/L^H. From equation (6.92), p_L/p_H is positively related with S and H/L^H. The analysis below studies how H/L^H is affected by s^* (and s).

Combining equations (6.89) and (6.93) gives the following

$$\frac{(1-\rho)\gamma}{\rho} \cdot \frac{H}{L^H} \cdot \left(\frac{\delta s^2}{(1+\delta s^2)S} - 1 \right) = 1.$$

If s^* increases, then due to proposition 2, S decreases, and the above equation implies $\frac{H}{L^H}$ decreases. This in turns implies that w^S/w^U increases and p_L/p_H decreases.

Using similar method, and combining equation (6.90) and the Foreign counterpart of (6.93), gives how H^*/L^{H*}, w^{S*}/w^{U*} and p_L/p_H change after an increase in s. After s increases, H^*/L^{H*} decreases, which implies w^{S*}/w^{U*} increases and p_L/p_H decreases.

The analysis below solves how a change in s will affect the equilibrium $\frac{H}{L^H}$ ratio and relative wage in Home. Suppose s increases, then due to the above analysis, the relative price p_L/p_H decreases. From equation (6.92), p_L/p_H is positively related with S and H/L^H. Since S is increasing in s, for p_L/p_H to decrease in s, H/L^H must be decreasing in s. Therefore, after s increases, H/L^H decreases, and from equation (6.85), this implies that the relative wage w^S/w^U increases.

Similarly, in the Foreign, using the same method, it can be shown that after s^* increases, H^*/L^{H*} decreases and the relative wage w^{S*}/w^{U*} increases.

6.8.4 The necessity of technology spillover during each period with complete knowledge

Section 6.6.3 works with a special form of technology spillover process. During period t, Home country (aging and advanced country) exports the high good Y_H which embodies the technology vintage v_t. It is assumed that v_t can be fully studied by Foreign during t, then during $t-1$, v_t becomes available in Foreign to produce Y_L. The assumption that v_t can be fully studied by Foreign can be called complete knowledge. The analysis here shows that the assumption of complete knowledge is necessary to have any equilibrium in which the fractions of students in both countries are constant through time.

The analysis starts from equation (6.74). Note that in any equilibrium where the fractions of students are constant through time in both countries, the following variables are also constant through time: S_t, H_t, L_t^H, L_t^L, S_t^*, H_t^*, L_t^{H*}, and L_t^{L*}. Time subscripts are dropped to denote equilibrium values. During equilibrium, the technology vintage evolution in Home becomes:

$$v_t = (1 + \mu S) \cdot v_{t-1}.$$

During t, Home will export Y_H to Foreign and technology vintage v_t is embodied in Y_H. During technology spillover, v_t can be partially or fully studied in Foreign and used in the next period t. Section 6.6.3 assumes that Foreign has complete knowledge about v_t so v_t can be completely studied by Foreign. Now this model works with a general case, where Foreign may not have complete knowledge about v_t so that only part of v_t can be studied. During t, there are three technology vintages in Foreign:

1. v_t: embodied in the imported Y_H from Home;

2. v_{Ht}^*: the relative advanced technology vintage during t in Foreign, used to produce Y_H;

3. v_{Lt}^*: the relative old technology vintage during t in Foreign, used to produce Y_L.

During the next period $t+1$, there are two technology vintages in Foreign used to produce goods. The first is v_{Ht+1}^*, relatively advanced, used to produce Y_H. The second is v_{Lt+1}^*, relatively old, used to produce Y_L. v_{Ht+1}^* is obtained via R&D during t, while v_{Lt+1}^* is carried from the available technology vintages during t. During t, both v_t and v_{Ht}^* are more advanced than v_{Lt}^*, and this implies: (1) R&D during t is conducted on v_t and v_{Ht}^* and the outcome is v_{t+1}^*; (2) some combination of v_t and v_{Ht}^* is carried onto $t+1$ and the outcome is v_{Lt+1}^*. A formal representation of the above idea is the following:

$$v_{Ht+1}^* = (1 + \mu S^*)(\phi \cdot v_t + \cdot(1 - \phi) \cdot v_{Ht}^*), \quad 0 \leq \phi \leq 1; \tag{6.A1}$$

$$v_{Lt+1}^* = \pi \cdot v_t + (1 - \pi) \cdot v_{Ht}^*, \qquad 0 \leq \pi \leq 1. \tag{6.A2}$$

Equation (6.A1) is saying, during t, the Foreign is doing R&D on v_t and v_{Ht}^*, with relative weights equal to ϕ and $(1 - \phi)$. Similarly, as implied by equation (6.A2), some combination of v_t and v_{Ht}^* is carried onto $t + 1$ to produce Y_L and the relative weights are π and $(1 - \pi)$.

Using equations (6.A1) and (6.A2), the $(t + 1)$ equilibrium version of equation (6.74) becomes:

$$\alpha \frac{v_{t+1} \cdot H^{\rho} \cdot (L^{H})^{1-\rho} + (1+\mu S^{*})(\phi \cdot v_t + (1-\phi) \cdot v^{*}_{Ht}) \cdot (H^{*})^{\rho} \cdot (L^{H*})^{1-\rho}}{v_t \cdot \theta \cdot L^{L} + (\pi \cdot v_t + (1-\pi) \cdot v^{*}_{Ht}) \cdot \theta \cdot L^{L*}}$$
$$= (1+\mu S)\frac{1-\rho}{\theta}\left(\frac{H}{L^H}\right)^{\rho}, \quad \text{for all } t \text{ during equilibrium.}$$

Simplifying the notations for the time-invariant equilibrium variables, the above equation can be represented by the following:

$$\frac{\text{constant} + \text{constant} \cdot \left(\phi + (1-\phi)\frac{v^{*}_{Ht}}{v_t}\right)}{\text{constant} + \text{constant} \cdot \left(\pi + (1-\pi)\frac{v^{*}_{Ht}}{v_t}\right)} = \text{constant}, \tag{6.A3}$$

holding for all t during equilibrium.

In order to have the above equation to hold for all t during equilibrium, the following must be true:

$$\phi = \pi = 1;$$
$$\text{or} \quad \frac{v^{*}_{Ht}}{v_t} = \text{constant}, \quad \text{for all } t \text{ during equilibrium} \tag{6.A4}$$

Next the ratio $\frac{v^{*}_{Ht}}{v_t}$ is solved using (6.A1). The periods t, $t-1$ and $t-2$ versions of equation (6.A1) are

$$v^{*}_{Ht} = (1+\mu S^{*})(\phi \cdot v_{t-1} + \cdot(1-\phi) \cdot v^{*}_{Ht-1}),$$
$$v^{*}_{Ht-1} = (1+\mu S^{*})(\phi \cdot v_{t-2} + \cdot(1-\phi) \cdot v^{*}_{Ht-2}),$$
$$v^{*}_{Ht-2} = (1+\mu S^{*})(\phi \cdot v_{t-3} + \cdot(1-\phi) \cdot v^{*}_{Ht-3});$$

where the above system implies

$$v^{*}_{Ht} = (1+\mu S^{*})\phi v_{t-1} + (1+\mu S^{*})^2\phi(1-\phi)v_{t-2}$$
$$+ (1+\mu S^{*})^3\phi(1-\phi)^2 v_{t-3} + (1-\phi)^3(1+\mu S^{*})^3 v^{*}_{Ht-3}.$$

Denote the first period where trade equilibrium occurs by T, and extend the above equation to T, yields

$$v^{*}_{Ht} = (1+\mu S^{*})\phi v_{t-1} + (1+\mu S^{*})^2\phi(1-\phi)v_{t-2} + \cdots + (1+\mu S^{*})^{t-T}\phi(1-\phi)^{t-T-1}v_T$$
$$+ \left[(1-\phi)(1+\mu S^{*})\right]^{t-T} v^{*}_{HT}.$$

Dividing the above expression by $v_t = (1+\mu S)^{t-T} v_T$, the $\frac{v^*_{Ht}}{v_t}$ ratio can be solved as

$$\frac{v^*_{Ht}}{v_t} = \underbrace{\frac{1+\mu S^*}{1+\mu S}\phi + \left(\frac{1+\mu S^*}{1+\mu S}\right)^2 \phi(1-\phi) + \cdots + \left(\frac{1+\mu S^*}{1+\mu S}\right)^{t-T} \phi(1-\phi)^{t-T-1}}_{\text{geometric sequence with common factor } = \frac{1+\mu S^*}{1+\mu S}(1-\phi)}$$
$$+ \left[(1-\phi)\left(\frac{1+\mu S^*}{1+\mu S}\right)\right]^{t-T} \frac{v^*_{HT}}{v_T}.$$

For the above expression to be constant through time, two conditions must satisfied: (1) the geometric sequence in the first line above has common factor equal to 0, which requires $\phi = 1$; and (2) the second line above is equal to 0, which also requires $\phi = 1$. If $\phi = 1$, equation (6.A1) becomes

$$v^*_{Ht+1} = (1+\mu S^*) \cdot v_t. \tag{6.A1'}$$

The above equation implies that in Foreign during t, the education sector can fully study the most advanced technology vintage of Home, namely v_t, and improve upon on that. In other words, in the R&D process, Foreign must has complete knowledge about the technology vintage embedded in the imported goods from Home during each period.

Chapter 7

Conclusion

7.1 Overview

Cognitive skills, consisting of both fluid intelligence (more important for technology innovation) and crystalized intelligence (more important for technology adoption), have vital effect on technological progress and economic growth. Though this is noticed by Heckman (1995) and recently emphasized by Rohwedder et al. (2010), no economic studies have formally modeled this phenomena, and important insights are missing. One goal of the thesis has been to analyze how individual aging, via its impacts upon cognitive skills, affects technological progress.

Individual aging changes agents' abilities. As an agent becomes old, he is better at learning current knowledge, but worse at innovating new things. Using a multi-country framework, Chapter 3 examines the impacts of individual aging on technological progress, and it is shown how the impacts of individual aging interact with a country's pre-aging technology level.

Several interesting and insightful results are found. First, chapter 3 predicts that countries will diverge, rather than converge, after the shock of individual aging. Since individual aging lowers agents' innovation ability, this tends to slow down technological progress. However, expecting this negative impacts, countries close to the world frontier will dismiss old innovators and hire young agents. On the other hand, countries faraway from the world frontier still retain old agents. As a result, the average (or aggregate) innovation ability is higher in advanced countries, which makes them grow faster. Individual aging magnifies the technology differences between advanced countries and lagged countries, causing a divergence.

Secondly, chapter 3 predicts that advanced countries can even enjoy a faster technological progress after individual aging. Following a severe individual aging, countries close to the world frontier dismiss all old innovators and replace them by young innovators. This means individual aging, while lowering old agents' innovation ability, *raises* the country's average innovation ability. In this way, individual aging is beneficial to advanced countries.

While individual aging could benefit advanced countries via its impacts on innovation ability, poorer countries can also benefit from the changes in adoption ability. Individual aging lowers innovation ability (which is bad for poorer countries due to employment rigidity), but it raises adoption ability. Technology adoption is also an important part of technological progress, especially for countries faraway from the world frontier. After taking individual aging's adoption-ability-raising effect into account, the net impact of individual aging on technological progress is ambiguous. Therefore, individual aging could faster technological progress of countries falling behind the world frontier, and so is beneficial to *all* countries (but still relatively more to advanced countries).

Besides individual aging, the thesis has also analyzed population aging, which is one of the most serious challenges facing many countries in the world. Population aging is characterized by a significant shift towards the older ages in the demographic distribution. Following a decline in mortality, together with a reduction in fertility, the number of old agents per unit of young has increased remarkably and will continue into the future.

Many economics studies have analyzed the impacts of demographic changes, especially how education investment and technological progress will be affected. Chapter 4 has also tackled this problem. Utilizing an overlapping-generations model with one-sector endogenous growth setup, population aging is via an increase in the survival rate and how population aging influences educational efforts, factor accumulation, output growth and welfare is analyzed.

Chapter 4 gives some interesting interests. First, consistent with the empirical findings and similar to some other papers in the literature, I find a positive relationship between survival rate and educational effort. After population aging, there are more agents choosing to go to school and do education. As a result, there are more skilled labor doing research during each period. This is due to the fact that with a higher survival rate, the discounted payoff of doing education rises and so more people go to school.

Secondly, it is found that population aging encourages physical capital accumulation. While there are more schooling agents who cannot save but must borrow to finance their youth consumption, and each schooling agent could borrow more than before population aging, population aging will increase the saving of each unskilled labor, and the last factor dominates the previous two. Therefore, population aging encourages physical capital accumulation. At the same time, because my model has a large resemblance with the neo-classical model, output growth is also enhanced by population aging, through the faster physical capital accumulation.

Thirdly, Chapter 4 shows that the impact of population aging upon welfare is ambiguous in general. Population aging, by lowering the amount of unskilled labor during each period, tends to increase wage rate while decrease the rental rate. I present necessary and sufficient conditions for either population aging is welfare-increasing or welfare-decreasing.

While the above one-sector model focuses on overall changes in technological progress, the relative impacts of population aging across each individual research sector are lacking. Since population aging affects relative amount of different input factors, which are complemented by different technologies, it could have important consequences upon the direction of technological progress in different research sectors. This is related to the directed technical change (DTC) literature. By extending the one-sector into a two-sector growth model, Chapter 5 has investigated this problem with several valuable insights.

Similar to the one-sector growth setup in Chapter 4, in Chapter 5 it is found that population aging (through an increase in survival rate and hence an increased longevity) would increase the overall size of schooling agents and hence scientists. A higher probability of surviving into the old age and getting R&D payoff leads to a higher present value of doing research and provides higher incentives for agents to do costly education.

Including both the price and the market size effects, Chapter 5 conducts the analysis in two special cases. When the two intermediates are perfect complements, the price effect is dominant. If there is a large proportion of scientists in the R&D sector yielding technology augmenting physical capital, then population aging will induce relatively even more scientists into that sector. On the other hand, if the proportion of scientists in the R&D sector yielding technology augmenting physical capital is small, population aging will induce relatively more scientists into the other research sector. This case is very similar to Irmen (2009), whose model does not feature the market size effect.

When the two intermediates are perfect substitutes, the market size effect is dominant and the opposite results are obtained. If there is a large proportion of scientists in the R&D sector yielding technology augmenting physical capital, then population aging will induce relatively less scientists into that sector. On the other hand, if the proportion of scientists in the R&D sector yielding technology augmenting physical capital is small, population aging will induce relatively more scientists into this research sector.

Last but not least, Chapter 6 studies the impacts of population aging upon education, technological progress and especially skill premium in an international trade model. While Chapters 4 and 5 analyze the relationship between population aging and educational efforts, they are restricted to closed economy equilibrium, and international trade is excluded. Therefore, in Chapter 6, it is natural to investigate the impact of population aging upon educational efforts under an international trade equilibrium.

In recent decades, skill premium has risen considerably in many countries and a large literature is investigating this problem. The current literature explains the rise in skill premium either from skill-biased technical change or international trade. However, there is few literature studying how population aging affects skill premium. Chapter 6 of the thesis investigates how population aging affects skill premium, in both autarky and international trade equilibrium.

Chapter 6 yields some interesting results. First, it is found that population aging encourages domestic educational efforts, and the intuition is the same as Chapters 4 and 5. However, in an international trade framework, population aging of one country will discourage educational effort of the other country, because the other country will produce more of unskilled labor intensive goods, which requires only unskilled labor.

Moreover, Chapter 6 finds that population aging will decrease skill premium in autarky equilibrium. In contrast, in an international trade equilibrium, population aging will raise skill premium in both countries, and this is consistent with empirical data on skill premium.

7.2 Future work

The thesis provides a comprehensive analysis of aging, including both individual aging and population aging, upon technological progress. However, some of the model features are kept simple to derive analytical results, and the following possible extensions are left for future work.

Endogenous individual aging and skill changes

In Chapter 3, it is assumed that the probabilities of abilities change (namely, p^A and p^I) are exogenous and vary across countries. A more realistic model is to consider them endogenously determined by agents' educational level and what jobs they do when young. Evidence from psychology suggests education and job characteristics are important in determining the magnitude of individual aging's impacts on abilities. A more fruitful model would explicitly include influencing (economic) factors of p^A and p^I in the analysis, and focus on optimal fiscal policy (education subsidy) and trade policy (which direct workers to certain job positions). Adding these complicates the current model very much and they are left for further research.

Individual aging in an international trade model

In Chapter 3, technological progress depends crucially on the ability of managers (adopters and innovators) hired by the principal, which means retaining rules of agents play an essential role. Relative incentives of hiring different agents, the retaining rules, depend on the relative importance of technology adoption and innovation, which further depends on a country's distance to world technology frontier. In the absence of countries' interactions (which constitute the world technology frontier), there is no technology adoption and innovation is the only channel left, whose consequence is the triviality of retaining rules (old low skill innovators are always fired and retaining rules about adopters become irrelevant). So even if there is no international trade, the implicit interactions among countries indirectly determine the technological progress of a country. The future research plans to have a two-country model with two final goods, and international trade in goods is allowed which results in changes in production patterns compared to autarky.

General multi-sector growth model

In Chapter 5 there are only two sectors and the degree of substitutability is either zero or infinite, which gives clear reduced-form solutions. A richer model would include a general number of multi-sector growth model with general degrees of substitutability across different sectors. This analysis would also serve as a test for the generality of the current results in Chapter 5.

Continuous time model

While the overlapping-generation models in discrete time are used in Chapters 4 to 6, most of the models in the directed technical change literature are continuous in time. The continuous-time version of overlapping-generation models, developed early by Yaari (1965) and later by Blanchard (1985), has also been popular in the analysis of population aging (Boucekkine et al., 2002; Echevarria and Iza, 2006; and Heijdra and Romp, 2009). The continuous-time model might give a more realistic description of the demographic structure (Lee, 2003). Moreover, compared to the discrete-time setup, at each instant there are more generations alive and the potential interactions across different generations can be analyzed.

References

Acemoglu, Daron. 1998. “Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality.” *Quarterly Journal of Economics*, 113: 1055-1089.

Acemoglu, Daron. 2002. “Directed Technical Change.” *Review of Economic Studies*, 69: 781-810.

Acemoglu, Daron. 2003. “Patterns of Skill Premia.” *Review of Economic Studies*, 70: 199-230.

Acemoglu, Daron. 2007. “Equilibrium Bias of Technology.” *Econometrica*, 75(5): 1371-1409.

Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth.* Princeton University Press.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn and David Hemous. 2012. “The Environment and Directed Technical Change.” *American Economic Review*, 102(1): 131-66.

Acemoglu, Daron, Philippe Aghion and Fabrizio Zilibotti. 2006. “Distance to Frontier, Selection, and Economic Growth.” *Journal of the European Economic Association*, 4(1): 37-74.

Acemoglu, Daron and David Autor. 2010. “Skills, tasks and technologies: implications for employment and earnings.” In Ashenfelter, O. and D. Card (eds.) *Handbook of Labor Economics*, Volume 4, Amsterdam: Elsevier.

Acemoglu, Daron and David Autor. 2011. “Skills, tasks and technologies: implications for employment and earnings.” Handbook of Labor Economics, Elsevier, Volume 4, Part B, 2011, Pages 1043-1171.

Acemoglu, Daron and Dan Cao. 2010. “Innovation by Entrants and Incumbents.” NBER Work Paper 16411.

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91(5): 1369-1401.

Acemoglu, Daron and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarter Journal of Economics*, 119: 1049-1090.

Aghion, Philippe and Peter Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica*, 60(2): 323-351.

Aghion, Philippe and Peter Howitt. 1998. *Endogenous Growth Theory.* Cambridge, Mass. MIT Press.

Anderson, Edward. 2005. "Openness and inequality in developing countries: a review of theory and recent evidence." *World Development*, 33: 1045-1063.

Arrow, Kenneth J. 1962. "The Economic Implications of Learning by Doing." *Review of Economic Studies*, 29: 155-173.

Atkinson, Anthony B. 2007 "The distribution of earnings in OECD countries." *International Labor Review*, 146: 41-60.

Auerbach, Alan J. and Kotlikoff, Laurence J. 1987. *Dynamic Fiscal Policy.* Cambridge University Press.

Auerbach, Alan J., Laurence J. Kotlikoff, Robert P. Hagemann and Giuseppe Nicoletti. 1989. "The Economic Dynamics of an Ageing Population: The Case of Four OECD Countries." OECD Economics Department Working Papers 62, OECD Publishing.

Autor, David H., Lawrence F. Katz and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *The Review of Economics and Statistics*, 90(2): 300-323.

Autor, David H., Lawrence F. Katz and Alan B. Krueger. 1998. "Computing inequality: have computers changed the labor market?" Quarterly journal of Economics, 113(4): 1169-1213.

Barro, Robert. 2001. "Human Capital and Growth." *American Economic Review*, 91(2): 12-17.

Barro, Robert and Xavier Sala-I-Martin. 1995. *Economic Growth.* McGraw-Hill, New York, London/Montreal.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Lifecycle of Earnings." *Journal of Political Economy*, 75(4, part 1): 352-65.

Blanchard, Olivier. 1985. "Debt, Deficits and Finite Horizons." *Journal of Political Economy*, 93(2): 223-247.

Boucekkine, Raouf, David de la Croix and Omar Licandro. 2002. "Vintage Human Capital, Demographic Trends, and Endogenous Growth." *Journal of Economic Theory*, 104(2): 340-375.

Burstein, Ariel, Javier Cravino and Jonathan Vogel. 2013. "Importing skill-biased technology." *American Economic Journal: Macroeconomics*, 5(2): 32-71.

Burstein, Ariel and Jonathan Vogel. 2010. "Globalization, technology, and the skill premium: a quantitative analysis." NBER Working Papers 16459, National Bureau of Economic Research, Inc.

Casterline, John B. 2011. "Fertility Prospects in the Arab Region." Population Division Expert Paper No.2011/6.

Cattell, Raymond B. 1971. Abilities: Their Structure, Growth, and Action. New York: Houghton Mifflin.

Chamberlin, Edward. 1933. The Theory on Monopolistic Competition. Cambridge, Mass.: Harvard University Press.

Clark, Robert L. and Joseph J. Spengler. The Economics of Individual and Population Aging. Cambridge University Press, 1980. Cambridge Books Online. Cambridge University Press.

Cutler, David, M., James M. Poterba, Louise M. Sheiner and Lawrence H. Summers. 1990. "An Aging Society: Opportunity or Challenge?" *Brookings Papers on Economic Activity*, 1: 1-73.

Davis, Donald R. and Prachi Mishra. 2007. "Stolper-Samuelson is dead: and other crimes of both theory and data." NBER Chapters, in: Globalization and Poverty, pages 87-108 National Bureau of Economic Research, Inc.

de la Croix, David and Omar Licandro. 1999. "Life Expectancy and Endogenous Growth." *Economics Letters*, 65(2): 255-263.

Diamond, Peter. 1965. "National Debt in a Neoclassical Growth Model." *American Economic Review*, 55(No. 5, part 1): 1126-1150.

Dixit, Avinash K. and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*, 67: 297-308.

Drandakis, E. and Edmund Phelps. 1965. "A Model of Induced Invention, Growth and Distribution." *Economic Journal*, 76: 823-840.

Ehrlich, Isaac and Francis T. Lui. 1991. "Intergenerational Trade, Longevity, and Economic Growth." *Journal of Political Economy*, 99(5): 1029-1059.

Eicher, Theo S. 1996. "Interaction between Endogenous Human Capital and Technological Change." *Review of Economic Studies*, 63: 127-145.

Eicher, Theo S. 1999. “Trade, Development and Converging Growth Rates: Dynamic Gains from Trade Reconsidered.” *Journal of International Economics*, 48: 179-198.

Feyrer, James. 2007. “Demographics and Productivity.” *Review of Economics and Statistics*, 89(1): 100-109.

Fougère, Maxime and Marcel Mérette. 1999. “Population Ageing and Economic Growth in Seven OECD Countries.” *Economic Modelling*, 16(3): 411-427.

Galor, Oded and David N. Weil. 1996. “The Gender Gap, Fertility, and Growth.” *American Economic Review*, 86(3): 374-387.

Gertler, Mark. 1999. “Government Debt and Social Security in a Life-Cycle Economy.” *Carnegie-Rochester Conference Series on Public Policy*, Vol.50: 61-110.

Goldberg, Pinelopi Koujianou and Nina Pavcnik. 2007. “Distributional effects of globalization in developing countries.” *Journal of Economic Literature*, 45: 39-82.

Gosling, Amanda, Stephen Machin and Costas Meghir. 2000. “The changing distribution of male wages in the UK.” *Review of Economic Studies*, 67: 635-666.

Grafenhofer, Dominik, Christian Jaag, Christian Keuschnigg and Mirela Keuschnigg. 2005. “Probabilistic Aging.” Manuscript. Et. Gallen, Switerland: University of St. Gallen IFF.

Griliches, Zvi. 1969. “Capital-skill complementarity.” Review of Economics and Statistics, 51(4): 465-468.

Grossman, Gene M. and Elhanan Helpman. 1991(a). “Quality Ladders in the Theory of Growth.” *Review of Economic Studies*, 43-61.

Grossman, Gene M. and Elhanan Helpman. 1991(b). *Innovation and Growth in the Global Economy.* Cambridge, Mass.: MIT Press.

Habakkuk, H. J. 1962. *American and British Technology in the Nineteenth Century: Search for Labor-Saving Inventions.* Cambridge: Cambridge University Press.

Hall, Robert and Charles Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker Than Others?" *Quarterly Journal of Economics*, 114(1): 83-116.

Hanushek, Eric A. and Dennis D. Kimko. 2000. "Schooling, Labor-Force Quality, and the Growth of Nations." *American Economic Review*, 90(5): 1184-1208.

Hanushek, Eric A. and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature*, 46(3): 607-668.

Haub, Carl. 2011. "Future Fertility Prospects for India." Population Division Expert Paper No.2011/4.

Heckman, James J. 1995. "Lessons from the Bell Curve." *Journal of Political Economy*, 103(5): 1091-1120.

Heijdra, Ben J. and Ward E. Romp. 2009. "Human Capital Formation and Macroeconomic Performance in an Ageing Small Open Economy." *Journal of Economic Dynamics and Control*, 33(3): 725-744.

Herrnstein, Richard J. and Charles Murray. The Bell curve: Intelligence and Class Structure in Amarican Life. New York: Free Press, 1994.

Hicks, John. 1932. *The Theory of Wages.* London: Macmillan.

Hornstein, Andreas, Per Krusell and Giovanni L. Violante. 2005. "The effects of technical change on labor market inequalities." Handbook of Economic Growth, in: Philippe Aghion and Steven Durlauf (ed.), Handbook of Economic Growth, edition 1, volume 1, chapter 20, pages 1275-1370 Elsevier.

Irmen, Andreas. 2009. “Population Aging and the Direction of Technical Change.” CESifo Working Paper No.2888.

Hock, Heinrich S. and David N. Weil. 2006. “The Dynamics of the Age Structure, Dependency, and Consumption.” Brown University, Department of Economics Working Papers 2006-08.

Hunt, Earl B. 1995. Will We Be Smart Enough?: A Cognitive Analysis of the Coming Workforce. Russell Sage Foundation Publications.

Hviding, Ketil and Marcel Merette. 1998. “Macroeconomic Effects of Pension Reforms in The Context of Ageing Populations: Overlapping Generations Model Simulations for Seven OECD Countries.” OECD Working Paper No. 201, Paris.

Jamison, Eliot A., Dean T. Jamison and Eric A. Hanushek. 2007. “The Effects of Education Quality on Income Growth and Mortality Decline.” *Economics of Education Review*, 26(6): 771-788.

Jones, Charles. 1995. “R&D-Based Models of Economic Growth.” *Journal of Political Economics*, 103: 759-784.

Jones, Gavin W. 2011. “Recent Fertility Trends, Policy Responses and Fertility Prospects in Low Fertility Countries of East and Southeast Asian.” Population Division Expert Paper No.2011/5.

Jovanovic, Boyan. 1998. “Vintage capital and inequality.” *Review of Economic Dynamics*, 1: 497-530.

Kalemli-Ozcan, Sebnem, Harl E. Ryder and David N. Weil. 2000. “Mortality Decline, Human Capital Investment, and Economic Growth.” *Journal of Development Economy*, 62: 1-23.

Katz, Lawrence and Kevin Murphy. 1992. "Changes in relative wages, 1963-1987: supply and demand factors." *The Quarterly Journal of Economics*, 107: 35-78.

Kennedy, Charles. 1964. "Induced Bias in Innovation and the Theory of Distribution." *Economic Journal*, 74: 541-547.

Koren, Miklos and Marton Csillag. 2011. "Machines and machinists: capital-skill complementarity from an international trade perspective." CEPR Discussion Paper No. DP8317.

Krusell, Per, Lee E. Ohanian, Jose Victor Rios-Rull and Giovanni L. Violante. 2000. "Capital-skill complementarity and inequality: a macroeconomic analysis." Econometrica, 68: 1029-1053.

Lee, Ronald D. 2003. "The Demographic Transition: Three Centuries of Fundamental Change." *Journal of Economic Perspectives*, 17(4): 167-190.

Lee, Ronald and Lawrence Carter. 1992. "Modeling and forecasting the time Series of U.S. Mortality." *Journal of the American Statistical Association*, 87(419): 659-671.

Ludwig, Alexander and Edgar Vogel. 2010. "Mortality, Fertility, Education and Capital Accumulation in a Simple OLG Economy." *Journal of Population Economics*, 23(2): 703-735.

Machin, Stephen and John Van Reenen. 1998. "Technology and changes in skill structure: evidence from seven OECD countries." *The Quarterly Journal of Economics*, 113: 1215-1244.

Machin, Stephen and John Van Reenen. 2007. "Changes in wage inequality." Open Access publications from London School of Economics and Political Science CEPSP18, London School of Economics and Political Science.

Miles, David. 1999. "Modelling The Impact of Demographic Change Upon the Economy." *The Economic Journal*, 109: 1-36.

Naito, Takumi and Laixun Zhao. 2009. "Aging, Transitional Dynamics, and Gains from Trade." *Journal of Economic Dynamics and Control*, 33: 1531-1542.

Nelson, Richard R. and Edmund Phelps. 1966. "Investment in Humans, Technology Diffusion, and Economic Growth." *American Economic Review*, 56(2): 69-75.

Nerlove, Marc. 1974. "Household and Economy: Toward a New Theory of Population and Economic Growth." *Journal of Political Economy*, 82(3): S200-S218.

Newell, Richard, Adam Jaffee and Robert Stavins. 1999. "The Induced Innovation Hypothesis and Energy-Saving Technological Change." *Quarterly Journal of Economics*, 114: 907-940.

Omran, Abdel R. 1971. "The Epidemiologic Transition: A Theory of the Epidemiology of Population Change." The Milbank Memorial Fund Quarterly, 49(4): 509-538.

Oeppen, Jim and James Vaupel. 2002. "Broken Limits to life Expectancy." *Science*, 296 (5570): 1029-1030.

Parro, Fernand. 2013. "Capital-skill complementarity and the skill premium in a quantitative model of trade." *American Economic Journal: Macroeconomics*, 5(2): 72-117.

Patton, George C. 2011. "Global Patterns of Mortality in Young People." Population Division Expert Paper No.2011/11.

Popp, David. 2002. "Induced Innovation and Energy Prices." *American Economic Review*, 92: 160-180.

Prettner, Klaus and Alexia Prskawetz. 2010. "Demographic Change in Models of Endogenous Economic Growth. A Survey." Vienna Institute of Demographiy Working Papers 8/2010.

Rebelo, Sergio. 1991. "Long-Run Policy Analysis and Long-Run Growth." *Journal of Political Economy*, 99(3): 500-521.

Rohwedder, Susann and Robert J. Willis. 2010. "Mental Retirement." *Journal of Economic Perspectives*, 24(1): 119-138.

Romer, Paul M. 1986. "Increasing Returns and Long-run Growth." *Journal of Political Economy*, 94: 1002-1037.

Romer, Paul. 1987. "Growth Based on Increasing Returns Due to Specialization." *American Economic Review*, 77: 56-62.

Romer, Paul. 1990. "Endogenous Technological Change." Journal of Political Economy, 98: S71-S102.

Samuelson, Paul. 1965. "A Theory of Induced Innovation along Kennedy-Weisäcker Lines." *Review of Economics and Statistics*, 47: 343-356.

Salthouse, Timothy A. 1990. Theoretical Perspectives on Cognitive Aging. Hillsdale, NJ:Erlbaum.

Sayan, Serdar. 2005. "Heckscher-Ohlin Revisited: Implications of Differential Population Dynamics for Trade within an Overlapping Generations Framework." *Journal of Economic Dynamics and Control*, 29(9): 1471-1493.

Scherer, Frederick M. 1984. Innovation and Growth: Schumpeterian perspectives. Cambridge, Mass.: MIT Press.

Schmookler, Jacob. 1966. *Invention and Economic Growth.* Cambridge, Mass.: Harvard University Press.

Solow, Robert. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, 70: 65-94.

Tuljapurkar, Shripad, Nan Li and Carl Boe. 2000. "A Universal pattern of Mortality Decline in the G7 Countries." *Nature*, June 15, 405: 789-792.

United Nations. 2008. Demographic Yearbook

United Nations. 2011(a). Department of Economic and Social Affairs, Population Division (2011). *Mortality Estimates from Major Sample Surveys: Towards the Design of a Database for the Monitoring of Mortality Levels and Trends.* Technical Paper No.2011/2.

United Nations. 2011(b). Department of Economic and Social Affairs, Population Division (2011). *World Population Prospects: The 2010 Revision, Highlights and Advance Tables.* Working Paper No. ESA/P/WP.220.

United Nations. 2012. Department of Economic and Social Affairs, Population division. *Changing Levels and Trends in Mortality: the Role of Patterns of Death by Cause.* United Nations Publication, ST/ESA/SER.A/318.

Van Zande, Jan Luiten. 2009. "The skill premium and the 'great divergence'." *European Review of Economic History*, 13: 121-153.

Weil, David N. 2006. "Population Aging." NBER Working Paper No. 12147.

Yaari, Menahem E. 1965. "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer." *Review of Economic Studies*, 32: 137-150.

Zhang, Junsen, Jie Zhang and Ronald Lee. 2001. "Mortality Decline and Long-Run Economic Growth." *Journal of Public Economics*, 80: 485-507.