RIETI Discussion Paper Series 25-E-120

Intergenerational Decision on Education and Migration within a Family in a Spatial Agglomeration Model

KONDO, Hiroki
Sophia University

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 25-E-120
December 2025

Intergenerational Decision on Education and Migration within a Family in a Spatial Agglomeration Model

Hiroki Kondo
Department of Economics, Sophia University

Abstract

This study examines how urban agglomeration is influenced by both family public goods, which has the advantage of proximity within a family, and human capital, which increases productivity with increasing proximity of residents within a city. In some cases, proximity advantages reinforce agglomeration forces, while in others, they work in the opposite direction and weaken them. When proximity advantages exist among family members, urban population density increases beyond what exists without such advantages. This situation discourages further migration of unskilled workers from more distant regions, thereby considerably dividing society. In these regions, families perpetually remain in the regions as unskilled workers, with lower substantial incomes. The analytical framework and findings of this study provide an important basis for evaluating several important policies.

First, the model exhibits multiple human capital agglomeration patterns: a monocentric equilibrium and polycentric urban structure with multiple core cities. Among them, the polycentric equilibrium enhances overall economic welfare and mitigates persistent social disparities across regions and generations. Thus, Japan should promote such an urban structure by expanding the geographical and administrative scope of local governments, as proposed by the *doshusei* reform. Second, the study examines the impacts of social security systems that provide family public goods to the elderly. The fact that this also mitigates social gap by encouraging parents to invest more in their children's education is also demonstrated in the study.

Keywords: Human Capital; Spatial Agglomeration; Intergenerational Interactions; Migration.
JEL classification: R12; R13; R23; I24; O15; H31; H41

I would like to thank the participants at the Irvine-Japan conference at University of California Irvine, the Economic Theory and Policy Workshop at Aoyama Gakuin University in Tokyo, the seminars at Australian National University and Otaru University of Commerce, the conference at the Max Planck Institute for Tax Law and Public Finance in Munich, and participants in the DP seminar at the Research Institute of Economy, Trade and Industry (RIETI) for valuable comments. I acknowledge the financial support from the Abe Fellowship Program by the Social Science Research Council and the American Council of Learned Societies in cooperation with and with funds provided by the Japan Foundation Center for Global Partnership. This study is conducted as a part of the RIETI project "Economic Analysis on the problem of an aging population and a declining birthrate in China and Japan in the COVID-19 pandemic".

1 Introduction

In an economy undergoing globalization, substantial interregional disparities persist in access to career opportunities, higher wages and higher education. This is a result of the agglomeration of highly skilled occupations and human capital in a limited number of mega cities driven by globalization, as analyzed in the field of new economic geography. However, if individuals are free to migrate across regions, such concentration would not be problematic. That is, individuals born in economically stagnant regions, in principle, have the opportunity to migrate to cities to pursue higher education for their career in the future. However, in practice, such individuals are less likely to pursue higher education, as shown in Figure 1(a) for Japan. Even if one cannot obtain higher education, an individual may migrate to the city, find a job, and secure a better higher-education opportunities for their children or grandchildren; thus, the interregional disparities may not matter in the long run. In reality, however, high-skilled workers migrate to high-income regions, whereas low-skilled workers migrate away from high-income regions, as analyzed by Ganong and Shoag (2017) in the US in recent years. Therefore, agglomeration deepen social divisions and seriously undermine fairness.

This paper incorporates the factors that tied individuals to their hometown into agglomeration model of new economic geography, and examines the extent to which the factors make inequality deepen and become persistent. Based on the framework, I evaluate several important policies. Specifically, I consider the pros and cons of expanding the geographical and administrative scope of local governments, as in the proposal of "dōshūsei" (regional government system). I also examine the impacts of social security for the elderly, may differ from those in the absence of agglomeration.

As the factors that tied individuals to their hometown, I focus on public goods within a family, which have the characteristics of public goods within the very local framework of a family and has the advantage of geographical proximity of members within a family unit. Family members, such as couples, parents, and children, determine the division of housework, including childcare and eldercare, and who participates in the labor market. The efficiencies of housework inputs and the benefits from the output significantly depend on the geographical proximity of the family members, even though recent technological progress in transportation, telecommunication, and automation somewhat mitigate the dependency.

The advantage of geographical proximity of human capital within a city and that of family members within a family unit affect family decisions, including the residential pattern, which is whether the parent and the child co-reside. Relatedly, these proximities determine whether the parent in a family gives higher education to the child as analyzed in Jensen (2012). Consequently, these proximities determine agglomeration pattern. The process of how the advantage of geographical proximity of human capital determines urbanization patterns has been thoroughly analyzed in the field of new economic geography.

However, the advantage of geographical proximity of family members within a family unit has not been adequately discussed despite their significant influence on the locational patterns of economic activities. In some cases, these advantages reinforce agglomeration forces, while in others, they work in the opposite direction and weaken them.

When the child becomes a skilled worker, she will live in the city, earn higher wages, and provide family public goods in a large amount, which benefits the parent. Attracted by higher wages and larger family public goods expected to be provided by her child in the city, an individual is willing to live in a region that is as close as possible to the city and invest in higher education. Consequently, cities and nearby regions have hosted an extremely large population. However, as concentration proceeds, the number of cities is increasingly limited. This situation makes the distance to the city that is closest to a region longer on average. In this situation, when her child becomes a skilled worker and moves to the city, for an individual who lives far away from the city, the benefits of family public goods provided by the child are diluted. Therefore, the situation discourages investment in higher education. She may be willing to migrate to the city or a nearby region and then invest in higher education. However, high population density in such area discourages migration of unskilled workers from regions farther away from such area.

Following the scenario above, agglomeration deepens social divisions and seriously undermines fairness. Inside the city and nearby regions, families enjoy high income family public goods in a large amount from generation to generation. Higher substantial income is inherited in the form of not only higher education but also family public goods provided that attract subsequent generations to the city and nearby regions. Therefore, population density in the city is extremely high. That is, the advantage of geographical proximity of family members within a family unit and that of human capital reinforce agglomeration forces in the city and nearby regions. In contrast, in regions far from the city, families perpetually remain in these regions as an unskilled worker with lower substantial income. In such regions, the advantage of geographical proximity of family members within a family unit weaken agglomeration forces. In Japan, the greater the distance from major metropolitan areas such as Tokyo, Osaka, and Nagoya, the less likely people are to invest in higher education, as in Figure 1(b)-(d). This results in enduring disparities in education and income between regions.

The analytical framework and findings of this study provide an important basis for evaluating several important policies.

First, I examine the implications of expanding the geographical and administrative scope of local governments, such as in the proposal for introducing the "dōshūsei" (regional government) system. The model has multiple human capital agglomeration patterns: monocentric and multicentric patterns. Multicentric equilibrium enhances economic welfare, mitigating an intergenerational persistent social gap. In a multicentric equilibrium, the population in each city is smaller. The wage in the city is lower, but the disutility from the larger population is mitigated more drastically; thus, the utility of a skilled worker in the city improves. Moreover, in any region in the economy, the average distance from the city to it is much shorter. Therefore, the total population receiving higher education is larger. In this sense, multicentric equilibrium enhances economic welfare, mitigating an intergenerational persistent social gap between urban and peripheral areas. Thus, Japan should address the excessive concentration in Tokyo and promote a more multicentric urban structure with multiple regional core cities by introducing "doshusei" to expand the geographical and administrative scope of local governments.

Second, I examine the impacts of social security for the elderly. Specifically, it considers the impact of a fully funded pension system such that the government levies a tax on an individual when she is young, deposits it, and refunds it to her in the form of family public goods provision (not cash) in the next period when she becomes old. This policy reduces the dependency of an individual on the child in the provision of family public goods; thus, it may make an individual less willing to give a higher education for her child. In reality, however, it increases the benefits of higher education and enhances economic welfare. The policy makes the child depend more on old parents who have publicly provided family public goods, especially when the child becomes an unskilled worker and remains in the hometown with the parent. In this case, the child reduces the provision of family public goods more drastically. Hence, an individual wants the child to be a skilled worker.

2 Literature

In the model human capital agglomerates in a limited number of cities given the geographically local in scope externalities. This is based on the new economic geography studies that theoretically and empirically analyze the agglomeration of economic activities and human capital. In this situation, the model considers the family decision on residential pattern and education investment.

Progresses in transportation, transaction, and remote communication technologies reduce the necessity for economic activities to be in close proximity. The location choice of a producer is free from the location of the customers. Moreover, firms and industries with significant input-output linkages need not be in close proximity. Moreover, within a firm, divisions, sectors, and workers can choose their locations separately.

In reality, however, some of these units and activities tend to concentrate in one place. Markets, economic activities and resources which have strong linkages with one another agglomerate in a limited number of cities, departing from the ones the linkages of which are less important. This brings about scale economies in the cities. First, manufacturing sectors cluster in a limited number of sites, departing from some of the final goods markets. Thus, regions and cities specialize in manufacturing sectors as empirically analyzed by Ellison and Glaeser (1997). Next, firms separate their headquarters and production plants. The headquarters agglomerate in larger cities, departing from the production plants in smaller cities or foreign countries. Therefore, a limited number of megacities host headquarters of various sectors, together with various types of related business services. Furthermore, human capital including researchers engaged in R&D, managers in different sectors, and business service specialists tends to be in close proximity and agglomerate in a limited number of megacities.

A new economic geography model by Krugman (1991), Krugman and Venables (1995), and Fujita, Krugman and Venables (1999) analyzes the first phase. Krugman (1991) develops a model with two sectors: manufacturing, which has scale economies in production process, and agriculture, which has constant returns to scale production process and intensively uses immobile land and peasants tied to it. Given the existence of such immobile workers, the market for manufacturing goods remains in rural areas even when all the other mobile workers leave. When transaction costs of manufacturing goods are high, the location of manufacturing is gravitated toward the distribution of the markets; thus, the manufacturing sector and the workers therein tend to disperse. However, when transaction costs are low, the location of manufacturing is free from smaller markets remain in rural areas. Thus, the manufacturing sector and the workers, who are also consumers, agglomerate in a limited number of sites to save the transaction costs, though such costs are lower than they used to be. Krugman and Venables (1995) analyze a two-country model with a manufacturing sector, where firms have input-output linkages. In this model, when transaction costs are low, the location of manufacturing becomes free from the international distribution of the population. Manufacturing agglomerates in one country despite making some of the final goods market much more distant. [1]

Duranton and Puga (2005) and Gabe and Abel (2012) analyzes the second phase, developing a model where the firm's organization and the interregional specialization endogenously evolve, as they are interrelated with one another. When the costs of remote management fall, firms will place their headquarters in larger cities that simultaneously host abundant business services, whereas they place their production plants in smaller cities that specialize in a similar sector's production process. Aarland, Davis, Henderson and Ono (2007) and Davis and Henderson (2008) empirically show that, in the U.S., headquarters will outsource business services, such as specialists in law, accounting, consulting, and advertising. This implies that the scale economies stemming from the linkages between headquarters and business services are very prominent. Therefore, headquarters from dif-

[1] Whether agglomeration occurs depends on the relative sizes of three effects: local competition, forward linkage, and backward linkage effects. The latter two work in favor of agglomeration. As per Krugman (1991), agglomeration of manufacturing reduces costs to achieve a certain level of utility for a household in that region (i.e., forward linkage). Conversely, the agglomeration of manufacturing workers provides a large local market for manufacturing in that region (i.e., backward linkages). Similarly, per Krugman and Venables (1995), agglomeration of intermediate goods firms reduces the costs to produce a final good in that country, and agglomeration of final goods firms provides a large local market for a firm providing intermediates. When transaction costs decrease, both linkage effects become less prominent. However, the local competition gets less fierce more dramatically.

ferent sectors and various types of business services concentrate in a limited number of large cities.

Moreover, technological progress has made it possible for the location of human capital to be away from the location of production plants and related markets of goods and services. Hence, given the scale economies of human capital, the locations draw closer to one another. [2] [3]

Arrow (1962), Uzawa (1965), Romer (1986), Lucas (1988), and Grossman and Helpman (1991) develop dynamic equilibrium analysis in the model with knowledge spillovers. Lucas (1988) argues for the importance of human capital externalities in determining persistent cross-country differences in economic growth rate and industrial specialization patterns. They contribute to the productivity of factors of production and the efficiency in the processes of further human capital accumulation. The scope is geographically local, as the externalities take the form of informal face-to-face interactions of knowledge exchanges in city centers. Therefore, countries and regions that succeed in hosting human capital accumulation can enjoy long-term economic growth. [4]

Jaffe (1986), Jaffe, Trajtenberg and Henderson (1993), Keller (2002), and Orlando (2004) empirically studied that knowledge spillovers are geographically localized. Lychagin, Pinkse, Slade and Van Reenen (2016) show that productivity significantly depends on the geographic location of labs and researchers rather than the headquarters of a firm. Moretti (2004) and Greenstone, Hornbeck and Moretti (2010) empirically analyze the mechanism of localized knowledge spillovers. Moretti (2004) finds geographically local in-scope productivity externalities by showing that the productivity of plants significantly

[2] The agglomeration of manufacturing also causes the agglomeration of human capital if manufacturing employs high-skilled workers more intensively than agriculture. In the model of Forslid and Ottaviano (2003), in manufacturing, unskilled workers engage in the assembly process, whereas skilled workers manage firms as entrepreneurs. The payment for a skilled worker is the fixed costs and, thus, causes increasing returns to scale. Similarly, the agglomeration of headquarters and related business services also cause the agglomeration of human capital if they intensively employ human capital.

[3] The new economic geography model by Krugman (1991) and Krugman and Venables (1995) and the human capital agglomeration model share a similar theoretical framework of scale economies. In the former, the variety of final goods enhances an individual's utility, and the variety of intermediate goods enhances a manufacturing firm's productivity. Similarly, the variety of human capital enhances the productivity of high-tech industries and advanced business service industries. The larger the resources in the city including human capital are, the larger number of varieties survive, and, thus, the higher the utilities and productivities of individuals in the city are. These varieties are measured by the elasticity of substitution in utility or the production function, inducing monopolistic competition profits. Burzynski, Deuster and Docquier (2020), Card (2009), and Ottaviano and Peri (2012) measure and model the elasticity of substitution of human capital. In addition, human capital enjoys the rewards that come externally from the market when concentrated.

[4] Grossman and Helpman (1991) also explain the persistent international specialization pattern by theoretically analyzing the international location of high-tech industries where R&D and human capital are important. Given the local in-scope knowledge spillover in R&D, high-tech industries agglomerate in a limited number of countries.

differs by city in the US. Plants in cities with increased concentration of human capital are more productive and more prominently so between industries with significant patent citation linkage than otherwise. Greenstone, Hornbeck and Moretti (2010) find local in-scope productivity externalities by showing that plants in counties that attract a large manufacturing plant become more productive than those in counties with similar attractive economic backgrounds but fail to attract such a plant in the US. This work then elucidates that such externalities occur between firms in industries that share worker flows and adopt similar technologies.

However, relative to the production side analyses focusing on the advantage of geographical proximity of human capital, the family side analyses focusing on the advantage of geographical proximity of family members are not intensive. In reality, the geographical proximity of family members significantly determines their contributions to the productions of some types goods and services which have the characteristics of public goods and services within a family and the utiltieis. Therefore, the geographical proximity of family members also affects human capital accumulation and its geographical concentration and dispersion pattern through the family decision on residential pattern and education investment.

Konrad, Kunemund, Lommerud and Robledo (2002) analyze the location choice of siblings in a family. However, the economic backgrounds of the hometown and destination are not considered. Education investment is also not analyzed. Education investment for the child will motivate her to move to high-income regions in the future, as analyzed in Docquier and Rapoport (2012) and Kerr et al (2016).

Many studies discuss the family decision to invest in human capital, which drives economic growth, inequality, and intergenerational mobility. In Glomm and Ravikumar (1992), as in Lucas (1988), human capital externalities contribute the efficiency in the processes of human capital formation, but the scope is much more limited to the family. Therefore, the larger the parent's stock of human capital, the higher the efficiency of the child's learning and the larger the amount of educational investment the parent gives to the child. These two channels enhance human capital formation, which drives economic growth as the engine yet renders income inequality and intergenerational mobility to be persistent. Galor and Zeria (1993), Banerjee and Newman (1993), Benabou, (1996), Mookherjee and Ray (2003), and Galor and Moav (2004) argue that credit market imperfections yield unequal opportunities to human capital investments, perpetuate more income inequality and decline intergenerational mobility. [5]

[5]Galor and Moav (2004) generalize the analysis by comparing human capital-driven growth to the early stages of industrial development when the lion's share of growth stemmed from physical capital accumulation. They argue that, at the early industrial stage, high inequality was conducive to growth because the rich have a higher propensity to accumulate physical capital, whereas, at later stages associated with an increasing role of human capital, high inequality may be an impediment to growth according to the aforementioned credit constraints argument.

However, few studies have examined a family decision on education in a framework where agglomeration is proceeding. The return of higher education depends on whether the region where a family locates has abundant job opportunities for high skilled workers. It in turn depends on the geographical agglomeration pattern of human capital in an economy.

The model in this paper considers how interregional distribution of human capital in an economy affects investments in higher education in the regions, and shows that urbanization is less prominent but the population densities in urban areas are more extremely high than those in previous researches. The model also shows that in the long run net migration has ceased but the gross migration remains: high-skilled workers migrate to the city, whereas low-skilled workers migrate away from it. Moreover, in the model, there emerges a serious social division between metropolitan and peripheral areas; that is, inequalities in substantial income and opportunity to higher education. Thus, decline in intergenerational mobility through inequalities in income and education opportunity becomes more serious than previous studies by introducing geographical aspects.

Studies have analyzed the impacts of public policies that relate to education on human capital formation, growth, and inequality. Restuccia and Urrutia (2004) argue that, in the US, persistency in intergenerational mobility mainly stems from the inequality of early education; they posit that policies that mitigate credit constraints at the early stages of education are the most effective in reducing persistency. Caucutt and Lochner (2020) calibrate a dynastic human capital investment model with credit constraints and note that while the effects of relaxing the constraint at a single stage are modest, all life-cycle credit constraints dramatically increase.

Some social security policies affect or are affected by education policies. Camago and Stein (2022) discuss that a reduction in credit constraint increases human capital formation and increases a society's demand for programs that increase the returns to human capital formation, such as public health policies, especially in poorer societies. Glomm and Kaganovich (2008) study how public education and social security, the two largest government programs in many countries, affect economic growth and income inequality. The effect is nonmonotonic. An increase in government spending on social security, public education, or both reduces income inequality and enhances growth, if the initial size of the spending is small. The model in this paper shows that social security reduces inequality and enhances human capital accumulation but through a different channel.

3 The OLG Model with Family and Geography

3.1 Geography and Production

Consider a long and narrow economy where many regions stand in line. The regions are indexed by $n = 1, ..., \bar{N}$ from west to east. In the economy, there are two types of workers: skilled and unskilled workers. Let H and L represent skilled and unskilled workers, respectively. The skill of a worker is determined by her family's decision on education investment, which will soon be explained. Let P_n denote the population in region n. The total population in the economy is P.

Let w_L denote the wage of an unskilled worker that is constant wherever the worker lives and whatever the population of that region is. Let w_H denote the wage of a skilled worker. Skilled workers' productivity and, thus, wages depend on the concentration of the workers in the region. That is, if the total population of skilled workers is below a critical value, the wage in that region remains equal to the wage of an unskilled worker, w_L. When it exceeds the critical value, w_H exceeds w_L. As it grows, so grows w_H; however, it grows less dramatically. Specifically, I introduce the following function:

$$w_H = \max[w_L, w(Ph_n)], \tag{1}$$

where $w' > 0$ and $w'' < 0$, and Ph_n is the population of skilled workers in that region. Call the region that has w_H higher than w_L the city.

In the model, production side settings have no micro foundation of scale economies studied in the new economic geography model and the human capital model. I simplify the production side settings to focus on within a family side. The new economic geography model and the human capital model show that progresses in transportation and communication technologies cause markets, economic activities and resources which have stronger linkages with one another tend to be in close proximity, departing from those which have less strong linkages, as we saw in section 2. This brings about scale economies in cities. Based on such a micro foundation, the properties of w, the function of the population of skilled workers are valid.

Given the scale economies, there are multiple equilibria, monocentric equilibrium where the economy has only one city and multicentric equilibrium where it has many cities. Unless the number of cities is too many, any multicentric equilibria are stable. I start with monocentric equilibrium where the economy has only one city in the center. The economy has only one city in the center. Re-index the regions as follows. The city is indexed by $n = 0$. From the city to the east, regions are indexed by $n = 0, 1, ..., N$, and to the west, regions are indexed by $n = 0, -1, ..., -N$, so that $\bar{N} = 1 + 2N$. The city has no unskilled worker in equilibrium, as will be confirmed. Thus, $Ph_0 = P_0$. In this model, what happens in the east half of an economy is the same as that in the west half. Thus, focusing only on the east half, $n = 0, 1, ..., N$, is sufficient.

3.2 Intergenerational Interaction within a Family

An individual lives for three periods: childhood, young adulthood, and old adulthood. Generation t is the generation of individuals born at the beginning of period t. An individual of this generation is in her childhood in period t, her young adulthood in period $t+1$, and her old adulthood in period $t+2$.

In period t in her childhood, an individual of generation t has no consumption. Her parent (an individual of generation $t-1$ and young adulthood in this period) decides whether to give her a higher education.

At the beginning of the next period $t+1$, in her young adulthood, an individual of generation t becomes the parent of a child (an individual of generation $t+1$). If she had higher education during childhood, she becomes a skilled worker who has enhanced productivity with probability μ. However, if she had no higher education, she becomes an unskilled worker. In any case, she can freely choose where to live and work when young. Let n_{t+1} denote the region. When she migrates, she will do so with her child. In n_{t+1}, she supplies a unit of labor inelastically and earns a wage. If she is a skilled worker and lives in the city ($n_{t+1}=0$), she earns a wage of w_H. If she lives in a region other than the city, she earns a wage of w_L. However, if she is an unskilled worker, she earns a wage of w_L wherever she works.

When young, an individual also decides whether to give her child a higher education. If she decides in favor of higher education, her child acquires enhanced productivity and becomes a skilled worker with probability μ. In contrast, if she decides not to provide a higher education, her child will be an unskilled worker. Let edu_{t+1} denote the expenditure for a higher education chosen by an individual of generation t for her child (generation $t+1$) in period $t+1$ when she is young and her child is in her childhood. A higher education costs a fixed amount of e. Therefore, $edu_{t+1} \in \{0, e\}$. Let $\mu(edu)$ denote the probability that her child becomes a skilled worker, and $\mu(e)=\mu$ and $\mu(0)=0$.

An individual retires at the beginning of period $t+2$, when old; she loses her parent (an individual of generation $t-1$) and becomes the grandparent of an individual of generation $t+2$. When old, she cannot migrate to other regions.

When young and old, each individual derives utility from private consumption goods and some types of goods that have the characteristics of public goods within the local framework of a family. Call such goods family public goods. The closer a parent (in old adulthood) and a child (in young adulthood) reside to each other, the larger the mutual benefits of family public goods they provide. Due to the existence of family public goods, the parent and child will live close to one another, as later discussed in-depth.

The government implements a fully funded pension system such that it levies a tax on an individual when she is young and provides a benefit in the next period when she is old. However, the pension benefit takes the form of family public goods, not cash.

Let C^y_{t+1} and C^o_{t+2} denote the consumptions of private goods by an individual of generation t in period $t+1$ when she is young and in period $t+2$ when she is old, respectively. Let G^y_{t+1} and G^o_{t+2} denote the provision of family public goods by her in period $t+1$ when she is young and in period $t+2$ when she is old, respectively. Let $\bar{G}$ denote family public goods provided by the government for an individual when she is old.

Assume that this has been constant over time. When the individual is old, her provision of family public goods is $G^o_{t+2} + \bar{G}$, which is the sum of privately and publicly provided goods.

3.3 Formal Setup for the Optimization Problem

The decisions and maximized utility of an individual of generation t in period $t+1$ when she is in her young adulthood depend on the decisions of her parent (an individual of generation $t-1$), where to live and work (n_t) in the last period t when her parent is in her young adulthood, and the provision of family public goods (G^o_{t+1}) in the current period $t+1$ when her parent is in her old adulthood. They also depend on her type, $x_{t+1} = H$ or L, given her parent's decision on education ($edu_t \in \{0, e\}$) in the last period t, and public policy, $\bar{G}$.

For example, consider how the choice of n_{t+1} by an individual of generation t is affected by the combination of $(x_{t+1}, n_t, G^o_{t+1}+\bar{G}, \bar{G})$. If her parent had given her a higher education when she was young ($edu_t = e$), she becomes a skilled worker ($x_{t+1} = H$) with probability μ. In this case, by leaving her hometown and migrating to the city ($n_{t+1} = 0$), she can earn higher wages w_H. However, by living apart from her parent, the utility from the family public goods provided by her parent decreases. Unless the provision of the family public goods, the sum of that her parent provides for herself and that the government provides for her parent ($G^o_{t+1} + \bar{G}$), is extremely large, she will live in the city. However, when she becomes an unskilled worker ($x_{t+1} = L$), she earns the same wage, w_L wherever she lives. Therefore, her decision on where to live and work (n_{t+1}) depends more on where her parent lives and works (n_t) and her parent's provision of the family public goods, $G^o_{t+1} + \bar{G}$. As it grows, so grows her incentive to live closer to the region where her parent lives (n_{t+1} is equal or closer to n_t). [6]

Similarly, $(x_{t+2}, n_{t+1}, G^o_{t+2}+\bar{G}, \bar{G})$, the choices of an individual of generation t and the public policy affect the decisions of the child. Among the child's decisions, where to live and work (n_{t+2}) and the provision of family public goods when the child is young (G^y_{t+2}) crucially affect her utility when she is old. Therefore, upon the decision making, an individual of generation t reacts to her parent's choices and the public policy, $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$, rationally guessing how her child will react to her choices $(x_{t+2}, n_{t+1}, G^o_{t+2} + \bar{G}, \bar{G})$. Her child's choices depend on her choices and the public policy, $(x_{t+2}, n_{t+1}, G^o_{t+2} + \bar{G}, \bar{G})$, exactly in the same manner her choices depend on her parent's choices and the public policy, $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$, unless there is a change in the backgrounds of the economy, such as wages, population distribution among regions, and the public policy.

[6] Assume a separable utility function. Hence a change in G^o_{t+1} has no direct effect on the individual's choices of private goods consumption when young and old. A change in G^o_{t+1} does, however, have indirect effects on these variables. Specifically, it affects the individual's location choice, n_{t+1}, which changes the wage and thus exerts an effect.

Note that among edu_{t+1}, n_{t+1} and G^o_{t+2}, only G^o_{t+2} is taken in period $t+2$ when she becomes old and her child's type (x_{t+2}) is realized, not in period $t+1$. However, this is also planned in period $t+1$. In this period, when she is young, to maximize her expected lifetime utility, an individual of generation t decides the savings, S_{t+1}. In deciding the savings, she considers that, in the next period $t+2$, the savings S_{t+1}, toghether with her child's type x_{t+2}, will affect her decicion on G^o_{t+2}. Therefore, determining S_{t+1} in period $t+1$ is the same as planning the combination of $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$, where $G^{O,H}_{t+2}$ is G^o_{t+2}, which she will choose in the case where her child becomes a skilled worker, and $G^{O,L}_{t+2}$ is G^o_{t+2}, which she will choose in the case where her child becomes an unskilled worker.

Moreover, there may be multiple optimal combinations $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$ that give the maximum expected lifetime utility. Let $Q(G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G} \mid x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ denote the probability that a combination of $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$ is planned when her parent's choice of $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ is given. Together with the choices of edu_{t+1} and n_{t+1}, below is a formal expression of the optimal choices of an individual of generation t, who is young in period $t+1$:

$$\begin{cases} edu_{t+1} = edu(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G}), \\ n_{t+1} = \psi(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G}), \\ G^o_{t+2} \text{ is planned as the combination } (G^{O,H}_{t+2}, G^{O,L}_{t+2}) \\ \qquad\qquad \text{with probability } Q\left(G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G} \mid x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G}\right). \end{cases} \tag{2}$$

Together with the combination $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$ and the savings S_{t+1} that is consistent with it, the private goods consumption (C^y_{t+1}) and the family public goods provision when young (G^y_{t+1}) are determined. Given the multiplicity of the combination of $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$, there are multiple optimal C^y_{t+1} and G^y_{t+1}. Let $G_Y(G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G} \mid x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ denote the optimal family public goods provision when young in the case where she plans a combination of $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$.

The optimal correspondences depend on the population distributions among regions in the future, as well as that in the current period and the parent's choice of $(x_t, n_t, G_{t+1} + \bar{G}, \bar{G})$. The decision-making of an individual of the current generation depends on the population distribution among regions in the current period. It is also affected by the decision-making of her child in the next generation, which in turn depends on the population distribution in the next period. Moreover, the child's behavior is affected by that of the grandchild. Iterating this discussion, an individual's decision-making and maximized expected lifetime utility are the functions or correspondences of the expected transition of population distribution among regions through time. If the dynamics of such population distribution derived by individuals' behaviors expressed by the optimal correspondences coincide with the expected one in the optimal correspondence, it is the equilibrium dynamics. Therefore, by deriving the optimal correspondences, it is possible to analyze the dynamic process of urbanization.

However, the analysis is formidably complicated. Therefore, I focus on the steady state equilibrium, where the population distribution among regions is time-invariant, though individuals move from region to region. Again, the optimal behavior is the function of the sequence of population distribution among regions through time. In the steady state, however, it is invariant. Thus, for notational simplicity, omit the sequence of population distribution from the optimal correspondences.

The optimal correspondences are obtained by solving the following functional equation of the optimization problem.

First, consider the optimal choices of private goods consumption (C^o_{t+2}) and family public goods provision (G^o_{t+2}) of an individual of generation t when old, in period $t+2$, given the savings, S_{t+1}, the region where she lives and works, n_{t+1}, and the type of worker her child becomes x_{t+2} which are the results of her choice including edu_{t+1} she made when young, in period $t+1$. The optimal choices are obtained by solving the following:

$$\begin{aligned}\max_{C^o_{t+2},G^o_{t+2}} \quad & u(C^o_{t+2}) \\ & + \sum_{\kappa}\sum_{\kappa'} Q\left(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid x_{t+2}, n_{t+1}, G^o_{t+2} + \bar{G}, \bar{G}\right) \\ & \qquad \times v\Big(G^o_{t+2} + \bar{G} + \delta\left(n_{t+1} - \psi(x_{t+2}, n_{t+1}, G^o_{t+2} + \bar{G}, \bar{G})\right) \\ & \qquad\qquad \times G_Y\left(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid x_{t+2}, n_{t+1}, G^o_{t+2} + \bar{G}, \bar{G}\right)\Big) \\ \text{s.t.} \quad & C^o_{t+2} + G^o_{t+2} = (1+r)S_{t+1},\end{aligned}$$

where r is the interest rate. Also, u is the function of the utility from the private goods consumption, and v is that from the total family public goods provision. Assume $u' > 0$, $u'' < 0$, $v' > 0$, and $v'' < 0$. Further, $u + v$ is the homothetic function with respect to the private goods consumption and the total family public goods provision. Her child determines the plan of $(G^{O,H}_{t+3}, G^{O,L}_{t+3})$, the provision of family public goods when her child becomes old, together with G^y_{t+2}, the provision of family public goods when her child is young and she is old. If there are multiple combinations $(G^{O,H}_{t+3}, G^{O,L}_{t+3})$ which her child will choose for a given $(x_{t+2}, n_{t+1}, G^o_{t+2} + \bar{G}, \bar{G})$, let $G^{O,H,\kappa}_{t+3}$ denote the κ-th largest $G^{O,H}_{t+3}$, and let $G^{O,L,\kappa'}_{t+3}$ denote the κ'-th largest $G^{O,L}_{t+3}$. When an individual is old, her provision of family public goods including the one through the government, $G^o_{t+2} + \bar{G}$, enhances the utility of her child as well as that of herself. Similarly, the family public goods provided by her child, G^y_{t+2}, enhance her utility. The impact of family public goods provided by the one on the other's utility is $0 < \delta(n_{t+1} - n_{t+2}) \leq 1$. This impact depends on the distance from the region in which she lives, n_{t+1}, to the region in which her child lives and works,

n_{t+2}. As the distance grows, the impact decreases. Assume the function is

$$\delta(n' - n) = \phi^{|n'-n|},$$

where $0 < \phi \leq 1$. Let $f(S_{t+1}, x_{t+2}, n_{t+1}, \bar{G})$ denote the optimal choice of G^o_{t+2} given her own decisions on S_{t+1} and n_{t+1} in the last period $t+1$ when she was young, and x_{t+2} be the type of worker her child becomes, which she can observe in period $t+2$.

Next, consider the optimal choices of where to live (n_{t+1}), whether to give her child a higher education (edu_{t+1}), and the amount of savings (S_{t+1}) and private goods consumption and family public goods provision (C^y_{t+1} and G^y_{t+1}) in period $t+1$ when young, given $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$, which is the result of her parent's choices and public policy. This problem is formally described as follows:

$$\begin{aligned}
\max_{n_{t+1}, edu_{t+1}, S_{t+1}, C^y_{t+1}, G^y_{t+1}} & \left[u(C^y_{t+1}) + v\left(G^y_{t+1} + \delta\left(n_t - n_{t+1}\right)\left(G^o_{t+1} + \bar{G}\right)\right)\right] \\
& + \left(\frac{1}{1+\rho}\right)\Bigg[\mu(edu_{t+1}) \\
& \quad \times \Bigg\{u\left((1+r)S_{t+1} - f(S_{t+1}, H, n_{t+1}, \bar{G})\right) \\
& \quad + \sum_{\kappa}\sum_{\kappa'} Q\left(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \;\middle|\; H, n_{t+1}, f(S_{t+1}, H, n_{t+1}, \bar{G}) + \bar{G}, \bar{G}\right) \\
& \quad \times v\Big(f(S_{t+1}, H, n_{t+1}, \bar{G}) + \bar{G} + \delta\left(n_{t+1} - \psi(H, n_{t+1}, f(S_{t+1}, H, n_{t+1}, \bar{G}) + \bar{G}, \bar{G}\right)\right) \\
& \quad \times G_Y\left(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid H, n_{t+1}, f(S_{t+1}, H, n_{t+1}, \bar{G}) + \bar{G}, \bar{G}\right)\Big)\Bigg\} \\
& + (1 - \mu(edu_{t+1})) \\
& \quad \times \Bigg\{u\left((1+r)S_{t+1} - f(S_{t+1}, L, n_{t+1}, \bar{G})\right) \\
& \quad + \sum_{\kappa}\sum_{\kappa'} Q\left(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \;\middle|\; L, n_{t+1}, f(S_{t+1}, L, n_{t+1}, \bar{G}) + \bar{G}, \bar{G}\right) \\
& \quad \times v\Big(f(S_{t+1}, L, n_{t+1}, \bar{G}) + \bar{G} + \delta\left(n_{t+1} - \psi(L, n_{t+1}, f(S_{t+1}, L, n_{t+1}, \bar{G}) + \bar{G}, \bar{G}\right)\right) \\
& \quad \times G_Y\left(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid L, n_{t+1}, f(S_{t+1}, L, n_{t+1}, \bar{G}) + \bar{G}, \bar{G}\right)\Big)\Bigg\}\Bigg] \\
& - c(P_{n_{t+1}}), \\
\text{s.t. } & C^y_{t+1} + G^y_{t+1} + S_{t+1} = w(x_{t+1}, n_{t+1}) - edu_{t+1} - \left(\frac{1}{1+r}\right)\bar{G},
\end{aligned}$$

where ρ is the subjective intertemporal discount rate, which is assumedly equal to the interest rate, $\rho = r$. Also, $P_{n_{t+1}}$ is the total population in region n_{t+1} and $c(P)$ is the disutility from the congestion in the region. Assume $c' > 0$ and $c'' \leq 0$. And $w(x_{t+1}, n_{t+1})$

is the wage that an individual with skill x_{t+1} earns in n_{t+1}. Thus, $w(H,0) = w_H$, and $w(x_{t+1}, n_{t+1}) = w_L$ for the other combinations of (x_{t+1}, n_{t+1}).

The optimal choices of edu_{t+1}, n_{t+1}, S_{t+1}, C^y_{t+1} and G^y_{t+1} are expressed as the functions of $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$, the combination of her parent's choices and the public policy. As functions $G^{O,H}_{t+2} = f(S_{t+1}, H, n_{t+1}, \bar{G})$ and $G^{O,L}_{t+2} = f(S_{t+1}, L, n_{t+1}, \bar{G})$ indicate, the optimal plan of $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$ depends on the public policy $\bar{G}$ and her own choice of S_{t+1} and n_{t+1}, which in turn depend on $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$. Hence, the optimal $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$ is also expressed as the function of $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ as in (2). [7] Thus, the maximized expected utility can also be expressed as a function of her parent's choices of x_{t+1}, n_t, and G^o_{t+1}, and the policy $\bar{G}$. This function is the value function.

4 Equilibrium

This study focuses on the optimal correspondences in the steady state where the population distribution among regions is unchanged. However, it remains formidably complicated to derive the correspondences directly from recursively constructed functional equations. Therefore, I adopt the following steps to analyze the correspondence. First, guess the optimal behaviors of an individual, discussing the population distribution that will be achieved in the steady state. Based on the discussions, I construct the candidate of the optimal correspondences (Section 4.1). Second, construct the simultaneous equations system the solutions of which may become the critical values of the optimal correspondences (Section 4.2). Third, confirm that the constructed correspondences are the optimal ones (Section 4.3-4.4).

4.1 Optimal Behavior

This section guesses the optimal behaviors of an individual, based on which it constructs the candidate of optimal correspondences.

An individual is willing to give a higher education for her child and live in a region as close as possible to the city even if she is an unskilled worker. It is not because she earns higher wages there, but because of the following reasons. If her child becomes a high skilled worker, her child earns higher wages in the city and provides a significant amount of family public goods. When she resides in a region that is closer to her child, she benefits more from the family public goods her child provides.

Thus, equilibrium has a tradeoff. That is, the closer to the city an individual resides, the higher the utility she gets from the large family public goods her child provides in the

[7] Function $f(S_{t+1}, x_{t+2}, n_{t+1}, \bar{G})$ is a single-valued function. There can be multiple optimal S_{t+1}. Thus, there can be multiple optimal combinations of $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$.

city, and the higher the population density. Thus, in equilibrium, there is no difference in an individual's utility wherever she lives.

Moreover, in equilibrium, there is no difference in the utility between the choice of giving a higher education for her child and living in a region that is not extremely far from the city and the choice of not giving a higher education for her child and living in a region that is farther away from the city and has a much smaller population. In a region that is farther away from the city, the benefits from the family public goods provided by her child in the city in the case where her child becomes a skilled worker are small. Thus, she remains in the region that has a much lower population density, and opts not to give a higher education for her child, rather than moves to a region that is closer to the city and has a higher population density and gives a higher education for her child.

Hence, when she is an unskilled worker, an individual will not necessarily live in a region that is closer to the city or region in which her parent lives, or give her child a higher education. Similarly, when her child becomes an unskilled worker, her child will not necessarily live closer to the city or her.

However, when her child becomes an unskilled worker, an individual can attract her child to a region close to her by providing family public goods in a significant amount. Accordingly, she can enjoy the family public goods her child provides more; thus, she is willing to do so. This situation entails the competition among individuals to attract their children to reside as close as possible to them by providing family public goods in an amount that is as significant as possible. Similarly, the competition among the parents to attract the individuals took place.

As a consequence of such a competition, in a region that is closer to the city, the population density becomes higher more drastically. A region near the city has a population that is much larger than the level that just cancels out the high value of family public goods provided by the child in the case where the child becomes a skilled worker and lives in the city. An individual will not reside in a region near the city without the provision of family public goods G^o_{t+1} in a significant amount by the parent in a region closer to the city.

If an individual's substantial income is high due to her own high wages or a large amount of family public goods her parent provides, she is advantageous in the competition. That is, she can afford to provide family public goods in a larger amount to attract her child to a closer region in the case where her child becomes an unskilled worker. In reality, a skilled worker in the city earns high wages. She will give a higher education for the child and plan to provide a large amount of family public goods to place the child in a region near the city in the case where the child becomes an unskilled worker. Moreover, following her parent, an individual who lives in a region near the city also plans to provide family public goods that are significant (but not so significant as what her parent provided) to place her child in a region that is closer to the region in which she lives but a bit further from the city where her parent lived.

In a family in which a parent gives her child a higher education, when her child becomes a skilled worker, her child will move to the city, earn higher wages and provide a large amount of family public goods. However, when her child becomes an unskilled worker repeatedly from generation to generation, the amount of family public goods in that family becomes smaller and smaller, and the region in which it is located gets farther and farther away from the city. If her parent provides family public goods in small or zero amounts and lives in a region far away from the city, the individual will not give a higher education for her child and will remain in her hometown far away from the city. In order to enjoy the benefits from the family public goods provided by her child in the city in the case where her child becomes a skilled worker, she must live in a region closer to the city where the population density is higher. However, high population density in such a region discourage her to live there without the parent who lives closer to the city and provides a significant amount of family public goods. In addition, the distance to the city discourages her to give a higher education for her child.

Based on the discussions thus far, assume that the optimal behaviors are described by the following correspondences:

An unskilled worker's behavior Group regions $n = 0, 1, ..., N$ into four sets: $\{0\}$, $\{1, ..., m\}$, $\{m, ..., \bar{n}\}$, and $\{\bar{n}+1, ..., N\}$. Then group regions in the second and third sets into m sets, Θ_l where $l = 0, ..., m-1$. This is constructed in the manner that $1 \in \Theta_{m-1}$ and $\bar{n} \in \Theta_0$. If n is the maximum in Θ_l, then $n+1$ is in Θ_{l-1}. Otherwise, $n+1$ is also in Θ_l.

- The parent lives in the city or region n' that is closer to the city than $\bar{n}$. That is, $n' = 0, ..., \bar{n}$. If the parent's provision of family public goods, including the one through the government, $G^o_{t+1} + \bar{G}$, satisfies $0 < G^o_{t+1} + \bar{G}$ and $\frac{\bar{G}_n}{\phi^{n-n'}} \leq G^o_{t+1} + \bar{G} < \frac{\bar{G}_{n-1}}{\phi^{n-n'-1}}$, where $n = n'+1, ..., \bar{n}$ or $\bar{G}_n \leq G^o_{t+1} + \bar{G}$, where $n = n'$, then

$$\begin{cases} edu(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = e, \\ \psi(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = n, \\ Q\left(\bar{G}, \frac{\bar{G}_{n+l}}{\phi^l} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n \to n+l}, \\ Q\left(\bar{G}, \frac{\bar{G}_{n+l+1}}{\phi^{l+1}} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n \to n+l+1}. \end{cases} \tag{3}$$

  If $G^o_{t+1} + \bar{G}$ is the minimum in the range, that is, $G^o_{t+1} + \bar{G} = \frac{\bar{G}_n}{\phi^{n-n'}}$ and n is not the maximum in Θ_l, then $Q_{n \to n+l} \in [0, 1]$, $Q_{n \to n+l+1} \in [0, 1]$, and $Q_{n \to n+l} + Q_{n \to n+l+1} = 1$. If n is the maximum in Θ_l or $G^o_{t+1} + \bar{G}$ is larger, then $Q_{n \to n+l} = 1$ and $Q_{n \to n+l+1} = 0$.

- The parent lives in region n', which is farther away from the city than $\bar{n}+1$. That is, $n' = \bar{n}+1, ..., N$. If $0 < \bar{G}$ but the parent does not provide family public goods privately, then

$$\begin{cases} edu(L, n', \bar{G}, \bar{G}) = 0, \\ \psi(L, n', \bar{G}, \bar{G}) = n', \\ Q\left(\bar{G}, \bar{G} \mid L, n', \bar{G}, \bar{G}\right) = 1. \end{cases} \tag{4}$$

- If $\bar{G} = 0$ and $G^o_{t+1} = 0$, then for all n',

$$\begin{cases} (edu(L, n', 0, 0), \psi(L, n', 0, 0)) = (e, \bar{n}), (0, n), \\ Q\left(0, 0 \mid L, n', 0, 0\right) = 1, \end{cases} \tag{5}$$

  where $n = \bar{n} + 1, ..., N$.

A skilled worker's behavior Group regions $n = 0, ..., \bar{n}$ into m sets, Ω_l where $l = 1, ..., m$. This is constructed in the following manner. The city ($n = 0$) is included in several Ω_l where $l = 1, ..., r$. Region $\bar{n}$ is included in Ω_m. If n is the maximum in Ω_l, then it is included in Ω_{l+1} as the minimum element. Otherwise, n is only in Ω_l. If $\bar{G} = 0$, all of $n = 0, ..., \bar{n}$ are included all Ω_l of $l = 1, ..., m$.
Suppose that the parent lives in the city or region n' that is closer to the city than $\bar{n}$. That is, $n' = 0, ..., \bar{n}$.

- if $\bar{G} \leq (G^o_{t+1} + \bar{G})\phi^{n'}$, then

$$\begin{cases} edu(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = e, \\ \psi(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = 0, \\ Q\left(\bar{G}, \frac{\bar{G}_l}{\phi^l} \mid H, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n' \to 0 \to l}, \end{cases} \tag{6}$$

  where $l = 1, ..., r$. If $G^o_{t+1} + \bar{G}$ is the minimum in the range, that is, $\bar{G} = (G^o_{t+1} + \bar{G})\phi^{n'}$, then for $l = 1, ..., r$, $0 \leq Q_{n' \to 0 \to l} \leq 1$ and $\sum_{l=1}^{r} Q_{n' \to 0 \to l} = 1$. If $G^o_{t+1} + \bar{G}$ is larger, then $l = 1$ and $Q_{n' \to 0 \to 1} = 1$.

- If $\bar{G}\phi^n \leq (G^o_{t+1} + \bar{G})\phi^{n'} < \bar{G}\phi^{n-1}$ where $n \in \Omega_l$ ($n = 1, ..., \bar{n}$ and $l = r, ..., m$), then

$$\begin{cases} edu(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = e, \\ \psi(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = 0, \\ Q\left(\bar{G}, \frac{\bar{G}_l}{\phi^l} \mid H, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n' \to 0 \to l}, \\ Q\left(\bar{G}, \frac{\bar{G}_{l+1}}{\phi^{l+1}} \mid H, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n' \to 0 \to l+1}. \end{cases} \tag{7}$$

  If $G^o_{t+1} + \bar{G}$ is the minimum in the range, that is, $\bar{G}\phi^n = (G^o_{t+1} + \bar{G})\phi^{n'}$ and $n \in \Omega_l \cap \Omega_{l+1}$, where $l = 1, ..., m - 1$, then $Q_{n' \to 0 \to l} \in [0, 1]$, $Q_{n' \to 0 \to l+1} \in [0, 1]$ and $Q_{n' \to 0 \to l} + Q_{n' \to 0 \to l+1} = 1$. Otherwise, $Q_{n' \to 0 \to l} = 1$ and $Q_{n' \to 0 \to l+1} = 0$.

- If $\bar{G}=0$, (6) holds for $l=1,...,m$. If $G^0_{t+1}=0$, then $Q_{n'\to 0\to l}$ is independent from n' and $\sum_{l=1}^{m} Q_{n'\to 0\to l}=1$.

Figure 2 shows an example of the correspondences.

In the correspondences, $\bar{G}_n/\phi^{n-n'}$ is the minimum $G^{O,L}_{t+2}+\bar{G}$ with which the parent in region n' can place an individual in region n when the individual becomes an unskilled worker. The value of $\bar{G}_n/\phi^{n-n'}$ provided in region n' is equal to the value of public goods in the amount of $\bar{G}_n$ provided in region n. When she wants to place her child in region n, an individual in region n' will choose $\bar{G}_n/\phi^{n-n'}$, the minimum $G^o_{t+2}+\bar{G}$ required to do so.

The region n included in Θ_l is the region such that an unskilled worker in that region will place her child in the region that is l or $l+1$ regions far away from the region n where she lives. If n is the maximum in Θ_l, she will place her child only in region $n+l$. Let $Q_{n\to n+l}\in[0,1]$ denote the probability that an individual in region n will place the child in region $n+l$ in the case where the child becomes an unskilled worker. The smaller n is, the bigger the l of Θ_l in which it is included. The closer the region n in which she resides is to the city, the closer the region in which an individual will place her child is to the city. However, it is harder to place her child in a region closer to where she lives given the larger population. Thus, the closer the region n where she resides is to the city, the longer the distance l between herself and her child is.

The region n in Ω_l is the region such that a skilled worker in the city will place her child in region l when her parent lived there. Let $Q_{n'\to 0\to l}\in[0,1]$ and $Q_{n'\to 0\to l+1}\in[0,1]$ denote the probabilities that a skilled worker in the city (region 0) whose parent lived in region n' will place the child in region l and $l+1$, respectively, in the case where the child becomes an unskilled worker. The smaller n' is, the smaller l of Ω_l it is included in. That is, the closer the hometown of a skilled worker who now lives in the city is, the closer the region in which she will place her child from the city.

The reason is as follows. A skilled worker whose parent was also a skilled worker and lived in the city will place her child in a region closer to the city because of the advantage in the competition to attract the child in a region that is as close as possible to the city. For such an individual, the plans to place her child in regions $1,...,r-1$ and r are indifferent. [8] For such an individual, the value of the sum of family public goods privately provided by her parent, G^o_{t+1}, which is zero as will soon be discussed, and the ones publicly provided for her parent, is $\bar{G}$. However, the value for an individual whose parent lived in a region other than the city is $\bar{G}\phi^{n'}$ where $n'\geq 1$. Therefore, a skilled worker whose parent lives farther from the city is less successful in attracting her child to a region closer to the city when her child becomes an unskilled worker.

[8] As the city has a large population, a skilled worker in the city cannot place her child in the city when her child becomes an unskilled worker without providing large $\bar{G}_0$. Thus, she may not be willing to choose it.

The optimal correspondences are defined for all the domains of $(x_{t+1}, n', G^o_{t+1} + \bar{G})$. I list the correspondences only for the part of the domains that will be observed. For example, the parent will not live farther from the city providing extremely large amount of family public goods. That is, the combination of very large n' and G^o_{t+1} will not be observed. She will also not live closer to the city providing zero or small family public goods (very small n' and G^o_{t+1}). However, to derive the optimal correspondences, we must consider the correspondences for all the domains. Appendix A shows the full description of the correspondences. Depending on the backgrounds of the economy, the correspondence when n' and G^o_{t+1} are large is different. [9] However, the important part of the correspondences are not affected.

After proving some Lemmas, the Proposition shows that if the critical values in the correspondences satisfy the simultaneous equations system in the next subsection, then the correspondences are optimal.

4.2 Simultaneous Equations System

Construct the simultaneous equations system the solution of which may become the critical values of optimal correspondences, such as $\bar{G}_n$, $Q_{n \to n+l}$ and $Q_{n' \to 0 \to n}$, considering the plausible relationships between an individual's utilities under any two of her choices.

I introduce the following *maximum conditional expected lifetime utility function.* Consider an individual's utility in the case where education for her child edu_{t+1}, residential region n_{t+1}, and the plan of family public goods she provides when she becomes old $(G^{O,H}_{t+2}, G^{O,L}_{t+2})$ are fixed, which are in reality freely chosen. That is, she can freely choose only the private goods she consumes when young and old, C^y_{t+1} and C^o_{t+2}, and the family public goods she provides when young, G^y_{t+1}. Let $\tilde{U}(S_{t+1}, edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G}; x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ denote an individual's expected lifetime utility conditional on the savings S_{t+1}, and fixed variables, $(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G})$, as well as her parent's choices and the policy, $(x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$. [10] Moreover, her utility depends on her child's decision on where to live and work, n_{t+2}, and the provision of family public goods, G^y_{t+2}. Assume the child determines them according to the correspondences in section 4.1 and Appendix A, which also depends on $(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G})$. Let $U(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G}; x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ denote the *maximum conditional expected lifetime utility*, which is achieved by choosing S_{t+1} that balances intertemporal utilities. Let $S(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G}; x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ denote such savings.

[9] Specifically, in some cases, there exist the amount of $G^o_{t+1} + \bar{G}$ above which an individual born in region n' will give a higher educatioh for her child, although she lives far from the city. The existence does not affect the optimal correspondences that appear here, as Appendix A shows.

[10] This is the utility maximized by an individual properly choosing C^y_{t+1} and G^y_{t+1}. Under the savings S_{t+1}, C^o_{t+2} is determined.

Using this function construct the simultaneous equations system. Appendix A provides a full description of the simultaneous equations. Below is a list of some important equations.

For a skilled worker consider the following conditions. When her parent lives in region $n' \in \Omega_l$ where $l = 1, ..., m$ and $n' = 0, ..., \bar{n}$, the expected lifetime utility when she lives in the city and plans to provide $\bar{G}_l/\phi^l$ to place her child in region l in the case where her child becomes an unskilled worker is higher than that when she lives in region $\bar{n}$, earns a wage of w_L, and plans not to provide family public goods privately:

$$U(e, \bar{n}, \bar{G}, \bar{G}; H, \bar{n}, \bar{G}, \bar{G}) \leq U(e, 0, \bar{G}, \bar{G}_l/\phi^l; H, \bar{n}, \bar{G}, \bar{G}). \tag{8}$$

Moreover, for a skilled worker whose parent lives in region $n' \in \Omega_l \cap \Omega_{l+1}$ where $l = 1, ..., m-1$, placing her child in region l and placing her child in region $l+1$ in the case where her child becomes an unskilled worker are indifferent:

$$U(e, 0, \bar{G}, \bar{G}_l/\phi^l; H, n', \bar{G}, \bar{G}) = U(e, 0, \bar{G}, \bar{G}_{l+1}/\phi^{l+1}; H, n', \bar{G}, \bar{G}). \tag{9}$$

A skilled worker wil not place her child in the city when she becomes an unskilled worker as $\bar{G}_0$ required to do so is extremely large. That is,

$$U(e, 0, \bar{G}, \bar{G}_0; H, 0, \bar{G}, \bar{G}) < U(e, 0, \bar{G}, \bar{G}_1/\phi; H, 0, \bar{G}, \bar{G}). \tag{10}$$

Next, consider the equations for an unskilled worker in regions $n = 1, .., \bar{n}$.

First, an unskilled worker residing in region n and residing in region $n+1$ are indifferent when her parent provides family public goods in the amount of $\bar{G}_n/\phi^{n-n'}$:

$$U(e, n, \bar{G}, \bar{G}_{n+l+1}/\phi^{l+1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}) = U(e, n+1, \bar{G}, \bar{G}_{n+l+1}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}), \tag{11}$$

where $n = 1, ..., \bar{n}-1$, $n \geq n'$ and $n \in \Theta_l$ but not the maximum in Θ_l where $l = 0, ..., m-1$. If n is the maximum in Θ_l and, thus, $n + 1 \in \Theta_{l-1}$, the condition is

$$U(e, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}) = U(e, n + 1, \bar{G}, \bar{G}_{n+l}/\phi^{l-1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}). \tag{12}$$

The amount of family public goods provided by her parent with which for an unskilled worker residing in the city and residing in region one are indifferent, $\bar{G}_0$, satisfies (11) with $n' = 0$, $n = 0$ and l that satisfies $1 \in \Theta_l$. However, the population in the city is much larger, and thus $\bar{G}_0$ is so large that an individual in the city will not choose it to place her child there in the case where her child becomes an unskilled worker as in (10).

Second, placing her child in region $n+l$ and placing her in region $n+l+1$ are indifferent when her parent provides family public goods in the amount of $\bar{G}_n/\phi^{n-n'}$:

$$U(e, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}) = U(e, n, \bar{G}, \bar{G}_{n+l+1}/\phi^{l+1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}), \tag{13}$$

where $n \geq n'$ and $n = 1, ..., \bar{n}$ except n that is the maximum in Θ_l where $l = 0, ..., m-1$. [11] When an individual lives in $\bar{n}$ and wants to place her child in that region, the minimum $G^o_{t+2} + \bar{G}$ required to do so is $\bar{G}$ (the private provision of family public goods is zero).

$$\bar{G}_{\bar{n}} = \bar{G}. \tag{14}$$

Region $\bar{n}$ is the farthest from the city in regions where educational investment for her child is rewarded:

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}) \geq U(0, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}), \tag{15}$$

$$U(e, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}) \leq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}). \tag{16}$$

Moreover, when her parent lives in region $\bar{n}$, an individual will not leave that region or abandon a higher education for her child.

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}) \geq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}). \tag{17}$$

An individual whose parent lives in region $\bar{n}+1$ also has no incentive to leave that region. However, she opts not to invest in a higher education:

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}) \leq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}). \tag{18}$$

If the family public goods are not publicly provided ($\bar{G} = 0$), instead of (17) and (18), it holds that:

$$U(e, \bar{n}, 0, 0; L, n', 0, 0) = U(0, \bar{n}+1, 0, 0; L, n', 0, 0). \tag{19}$$

Lastly, populations in regions $n = 0, ..., \bar{n}$ satisfy

$$\frac{2\lambda}{1-\lambda} \sum_{n=1}^{\bar{n}} P_n = P_0, \tag{20}$$

and that in regions farther from the city than $\bar{n}$ is

$$P_n = \frac{1}{2} \frac{P - P_0/\lambda}{N - \bar{n}} \quad \text{for} \quad n = \bar{n}+1, ..., N.$$

As $\bar{N} = 1 + 2N$, this is rewritten as

$$P_n = \frac{P - P_0/\lambda}{\bar{N} - (1 + 2\bar{n})} \quad \text{for} \quad n = \bar{n}+1, ..., N. \tag{21}$$

[11] Given the procedure in constructing Θ_l, the number of the equations is $\bar{n} - m$ but $n + l$ and $n + l + 1$ in equations cover all the regions from m to $\bar{n}$.

The city has no unskilled workers and thus $Ph_0 = P_0$. Therefore, from (1), the wage of high skilled workers is

$$w(P_0) = w_H. \tag{22}$$

I will show that an individual's best reaction to her parent's choice is expressed as the correspondences in subsection 4.1 if the critical values in the correspondences satisfy the simultaneous equations system.

Note that each equation alone does not ensure the optimality of an individual's reaction. For example, equation (13) merely states that when the parent chooses $\bar{G}_n/\phi^{n-n'}$, the maximum conditional expected lifetime utility where she plans to place her child in region $n+l$ by choosing $\bar{G}_{n+l}/\phi^l$ is equal to the one where she plans to place her child in region $n+l+1$ by choosing $\bar{G}_{n+l+1}/\phi^{l+1}$ if her child becomes an unskilled worker. However, these may not necessarily be larger than any other possible maximum conditional expected lifetime utilities with $G^o_{t+2} + \bar{G}$ other than $\bar{G}_{n+l}/\phi^l$ and $\bar{G}_{n+l+1}/\phi^{l+1}$. Similarly, equations (11) and (12) state that when the parent chooses $\bar{G}_n/\phi^{n-n'}$, the maximum conditional expected lifetime utility when she resides in region n is equal to that when she resides in region $n+1$. However, equations (11) and (12) do not indicate that these are larger than any other maximum conditional expected lifetime utilities when she resides in a region other than n and $n+1$.

However, if each equation is satisfied together with all the other equations in the simultaneous equations system, an individual's reaction in that equation is optimal, which I will show in the following subsections. Thus, the best reactions are expressed as correspondences in subsection 4.1 and Appendix A, if the critical values are the solutions of the simultaneous equations.

I will show the optimality of the correspondences in subsections 4.1 and 4.2 in the following steps. First, discuss the optimal behavior of an individual when she gives her child a higher education and cannot choose where to live and work. Second, I consider the case where an individual can choose where to live and work when young. Finally, I consider the optimal decision on education for her child when young.

Before the analysis, below is a summary of some important properties of the maximum conditional expected lifetime utility function.

Lemma 1

(i) Let $S^n_{n+l,n+l+1}$ denote the savings that satisfies:

$$\begin{aligned}&\tilde{U}(S^n_{n+l,n+l+1}, e, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n, G^o_{t+1} + \bar{G}, \bar{G})\\ =&\tilde{U}(S^n_{n+l,n+l+1}, e, n, \bar{G}, \bar{G}_{n+l+1}/\phi^{l+1}; L, n, G^o_{t+1} + \bar{G}, \bar{G}).\end{aligned} \tag{23}$$

If (23) holds, then

$$\begin{aligned}&\tilde{U}(S^n_{n+l,n+l+1}, e, n+1, \bar{G}, \bar{G}_{n+l}/\phi^{l-1}; L, n, G^o_{t+1} + \bar{G}, \bar{G})\\ <&\tilde{U}(S^n_{n+l,n+l+1}, e, n+1, \bar{G}, \bar{G}_{n+l+1}/\phi^{l}; L, n, G^o_{t+1} + \bar{G}, \bar{G}).\end{aligned} \tag{24}$$

(ii) *If* $G^{O,L}_{t+2} < G^{O,L'}_{t+2}$, *then*

$$\frac{\partial U(e,n,\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',G^{o}_{t+1}+\bar{G},\bar{G})}{\partial(w+\phi^{(n-n')}G^{o}_{t+1})} < \frac{\partial U(e,n,\bar{G},G^{O,L'}_{t+2}+\bar{G};L,n',G^{o}_{t+1}+\bar{G},\bar{G})}{\partial(w+\phi^{(n-n')}G^{o}_{t+1})}.$$

(iii) *If* $G^{O,L}_{t+2} = G^{O,H}_{t+2}$, *then*

$$\frac{\partial U(0,n,G^{O,H}_{t+2}+\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',G^{o}_{t+1}+\bar{G},\bar{G})}{\partial G^{o}_{t+1}} < \frac{\partial U(e,n,G^{O,H}_{t+2}+\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',G^{o}_{t+1}+\bar{G},\bar{G})}{\partial G^{o}_{t+1}}.$$

(iv) *If* $G^{O,L}_{t+2} = G^{O,H}_{t+2}$ *and if the government's provision of family public goods through the parent increases by* $d\bar{G}$ *and the tax on the individual increases by* $(1/(1+r))d\bar{G}$ *but the government's provision of family public goods when old does not change, then*

$$\frac{\partial U(0,n,G^{O,H}_{t+2}+\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',G^{o}_{t+1}+\bar{G},\bar{G})}{\partial \bar{G}} < \frac{\partial U(e,n,G^{O,H}_{t+2}+\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',G^{o}_{t+1}+\bar{G},\bar{G})}{\partial \bar{G}}.$$

(v) *If*

$$U(e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,G^{o}_{t+1}+\bar{G},\bar{G}) = U(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^{o}_{t+1}+\bar{G},\bar{G}),$$

where $G^{O,L}_{t+2} < G^{O,L'}_{t+2}$, *then*

$$U(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^{o}_{t+1}+\bar{G},\bar{G}) < U(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^{o}_{t+1}+\bar{G},\bar{G}).$$

See Appendix B for the proof.

4.3 Behavior of An Unskilled Worker

4.3.1 Decision on Where to Place the Child

First, consider an unskilled individual's decision on where to place the child given that her decisions on where to live, and whether to give her child a higher education are fixed.

Lemma 2
Suppose the parent of an unskilled individual lives in region n' *that is closer to the city* ($n' \in \{0,...,\bar{n}\}$)*. Suppose, for an unskilled individual, education and where to live are fixed, in that she gives a higher education for her child and lives in region* n.

(i) *Plan of* $G_{t+2}^{O,L}$*, the provision of family public goods in the case where the child becomes an unskilled worker.*

(a) *When the parent's provision of family public goods including the one through the government,* $G_{t+1}^{o} + \bar{G}$*, is equal to* $\bar{G}_n/\phi^{n-n'}$*, where* $n \in \{n', ..., \bar{n}\}$*, an unskilled individual in region* n *will place her child in region* $n + l$*, where* n *is in* Θ_l*, by providing family public goods in the amount* $G_{t+2}^{O,L} + \bar{G} = \bar{G}_{n+l}/\phi^{l}$*. If* n *is not the maximum in* Θ_l*, placing her child in region* $n + l + 1$ *by providing family public goods in the amount* $G_{t+2}^{O,L} + \bar{G} = \bar{G}_{n+l+1}/\phi^{l+1}$ *is also optimal.*

(b) *The larger the parent's provision of family public goods including the one through the government,* $G_{t+1}^{o}+\bar{G}$*, is, the closer the region in which an unskilled individual will place her child by providing larger* $G_{t+2}^{O,L}+\bar{G}$ *in the case where her child becomes an unskilled worker. The critical amount of* $G_{t+1}^{o}+\bar{G}$ *above which an unskilled individual in region* n *switches her choice from placing her child in region* $n + s + 1$ *to placing her child in region* $n + s$ $(s = 0, ..., l, ..., \bar{n} - n - 1)$ *is larger than that above which an unskilled individual in a region that is closer to the city than region* n *(a region in* $\{0, ..., n - 1\}$*) will do so, and smaller than that above which an unskilled individual in a region that is farther from the city than region* n *(a region in* $\{n + 1, ..., \bar{n}\}$*) will do so.*

(ii) *Plan of* $G_{t+2}^{O,H}$*, the provision of family public goods in the case where the child becomes a skilled worker.*

In the case where $n \in \{n', ..., \bar{n}\}$*, whatever the parent's provision of family public goods including the one through the government,* $G_{t+1}^{o}+\bar{G}$*, is, an unskilled individual will not privately provide family public goods in the case where her child becomes a skilled worker* $(G_{t+2}^{O,H} = 0)$*. In the case where* $n \in \{\bar{n} + 1, ..., N\}$*, when* $G_{t+1}^{o} + \bar{G}$ *is very large, an unskilled individual will choose positive* $G_{t+2}^{O,H}$*.*

See Appendix B for the proof. An individual's choice of where to place the child when her parent lives in region n' that is farther away from the city ($n' \in \{\bar{n}+1, ..., N\}$) is discussed in section 4.3.3 together with the decision on the residential region and education investment for her child.

In reality, her family public goods provision when old, $G_{t+2}^{O,L}$, is not fixed but freely planned. The maximum expected lifetime utility in which education for her child and where to live and work are given but the choice of $G_{t+2}^{O,L}$ is flexible is the maximum of the *maximum conditional expected lifetime utilities*, each of which $G_{t+2}^{O,L}$ is fixed. That is, $G_{t+2}^{O,L}$ under which the *maximum conditional expected lifetime utility* is highest is the optimal plan of $G_{t+2}^{O,L}$.

Lemma 2(i) states that if she resides in region n and her parent resides in region n' and provides family public goods in the amount of $G_{t+1}^{o} + \bar{G} = \bar{G}_n/\phi^{n-n'}$ (the value in region

n is $\bar{G}_n$), an individual's the optimal choices of $G^{O,L}_{t+2} + \bar{G}$ are $\bar{G}_{n+l}/\phi^l$ and $\bar{G}_{n+l+1}/\phi^{l+1}$. Again, note that equation (13) of region n merely states that *the maximum conditional expected lifetime utility* when she places her child in region $n + l + 1$ with $\bar{G}_{n+l+1}/\phi^{l+1}$ is equal to the one when she places her child in region $n + l$ with $\bar{G}_{n+l}/\phi^l$. Generally, these are not necessarily higher than any other possible *maximum conditional expected lifetime utilities* with any other $G^{O,L}_{t+2} + \bar{G}$. However, if equation (13) holds, together with the other equations in the simultaneous equations system, these two are optimal choices, and $S(e, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ and $S(e, n, \bar{G}, \bar{G}_{n+l+1}/\phi^{l+1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ are the optimal savings, as Figure 3 shows.

To show this, consider the intuition of Lemma 1(v). Equation (13) of region $n - 1$ version states that if she resides in a region that is closer to the city such as $n-1$ and her parent provides larger amount of family public goods such as $\bar{G}_{n-1}/\phi^{n-1-n'}$ (the value in region $n - 1$ is $\bar{G}_{n-1}$), *the maximum conditional expected lifetime utility* when she places her child in a distant region $n + l$ with $\bar{G}_{n+l}/\phi^{l+1}$ is equal to the one when she places her child in a closer region $n + l - 1$ with $\bar{G}_{n+l-1}/\phi^l$. [12] Lemma 1(v) states that if that individual moves from region $n-1$ to region n, with other things including the value of the family public goods her parent provides being equal, the former becomes strictly higher than the latter. For the individual, the utility from the family public goods that her child provides becomes larger, simply because she approaches the region where her child lives (region $n+l$ or $n+l-1$). [13] In this scenario, however, the benefits from the family public goods in the case where she places her child in a more distant region $n + l$ increase more drastically than that in the case where she places her child in a closer region $n + l - 1$ does. In the former case, the amount of G^y_{t+2} her child provides is larger, as the amount G^o_{t+2} she provides is smaller. Therefore, by approaching her child, the value of the family public goods provided by her child G^y_{t+2} is more dramatically amplified in the case where her child lives in a more distant region $n + l$ than in the case where her child lives in a closer region $n + l - 1$.

In addition to this, smaller substantial income stemming from lower value of family public goods provided by her parent further discourages an individual in equation (13) of region n from placing her child in a closer region $n+l-1$ with larger amount $\bar{G}_{n+l-1}/\phi^{l-1}$. The parent provides family public goods in the amount of $\bar{G}_n/\phi^{n-n'}$. This is smaller than the provision of family public goods by the parent in equation (13) of region $n-1$ version, $\bar{G}_{n-1}/\phi^{n-1-n'}$. In the former case, when an individual places her child in the closer region $n + l - 1$ with larger amount $\bar{G}_{n+l-1}/\phi^{l-1}$ despite the smaller substantial income, the balance between the utility from the consumption goods and that from family public

[12] If n is the maximum in Θ_l (in this case, $n + 1$ is the minimum of Θ_{l-1}), these two are equal, not in region $n + 1$, but in region $n + 2$. Further, in this case, the following discussions are essentially the same.

[13] The amounts of $G^{O,L}_{t+2}$ required to place her child in these regions becomes a bit smaller, which slightly increases her consumption of private goods when old. However, the benefits from family public goods provided by her child increase more drastically.

goods is overly biased to the latter. Therefore, when she plans $G^{O,L}_{t+2} + \bar{G}$ that is larger than $\bar{G}_{n+l-1}/\phi^{l-1}$, the total provision of family public goods is excessive.

From a similar logic, inversely, when she plans $G^{O,L}_{t+2} + \bar{G}$ to be equal to or smaller than $\bar{G}_{n+l+2}/\phi^{l+2}$, the total provision of family public goods is exceedingly small relative to the consumption of private goods.

Thus, if $G^{o}_{t+1} + \bar{G} = \bar{G}_n/\phi^{n-n'}$, the plan of $G^{O,L}_{t+2} + \bar{G}$ in intermediate size such as $\bar{G}_{n+l}/\phi^{l}$ or $\bar{G}_{n+l+1}/\phi^{l+1}$, and the amount of savings $S(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ where $G^{O,L}_{t+2} + \bar{G}$ is $\bar{G}_{n+l}/\phi^{l}$ or $\bar{G}_{n+l+1}/\phi^{l+1}$ adequately balance her intertemporal utilities when young and when old, and the utilities from the consumption of private goods, and the provision of family public goods when an individual gets old and her child becomes an unskilled worker. [14]

Next, consider how an individual's plan of where she will place her child changes as the parent's provision of family public goods $G^{o}_{t+1} + \bar{G}$ increases. Lemma 1(ii) states that when G^{o}_{t+1} provided by her parent increases, every $U(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n, G^{o}_{t+1} + \bar{G}, \bar{G})$ rises, but more importantly, the one with a larger $\bar{G}^{O,L}_{t+2} + \bar{G}$ rises more drastically, as in Figure 4. Therefore, if $G^{o}_{t+1} + \bar{G}$ provided by her parent exceeds a critical value, $U(e, n, \bar{G}, \bar{G}_{n+l-1}/\phi^{l-1}; L, n, G^{o}_{t+1} + \bar{G}, \bar{G})$ exceeds $U(e, n, \bar{G}, \bar{G}_{n+l}/\phi^{l}; L, n, G^{o}_{t+1} + \bar{G}, \bar{G})$, and the optimal savings jump from $S(e, n, \bar{G}, \bar{G}_{n+l}/\phi^{l}; L, n, G^{o}_{t+1} + \bar{G}, \bar{G})$ to $S(e, n, \bar{G}, \bar{G}_{n+l-1}/\phi^{l-1}; L, n, G^{o}_{t+1} + \bar{G}, \bar{G})$.

An increase in $G^{o}_{t+1} + \bar{G}$ provided by her parent means an increase in the substantial

[14] As defined, the savings $S(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ is obtained to maximize an individual's lifetime utility conditional on that she plans to provide family public goods in the amount $G^{O,L}_{t+2}$. In reality, however, the order of decision makings of the savings and G^{o}_{t+2} is the opposite. First, when young, an individual determines the savings, and, when old, she decides G^{o}_{t+2}, though, when young, she considers that she will decide G^{o}_{t+2} when old, depending on the savings. Therefore, under the savings $S(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ where $G^{O,L}_{t+2}$ is fixed, an individual does not necessarily choose that $G^{O,L}_{t+2}$ when old. If $G^{O,L}_{t+2} + \bar{G}$ is extremely large like $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l-1}/\phi^{l-1}$, $S(e, n, \bar{G}, \bar{G}_{n+l-1}/\phi^{l-1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ is also large. However, the savings are not adequately large to induce her to choose such a large $\bar{G}_{t+l-1}/\phi^{l-1}$ when she is old and her child becomes an unskilled worker, as Figure 3 shows. The larger $G^{O,L}_{t+2}$ is, the larger the savings she prefers. That is, when $G^{O,L}_{t+2} + \bar{G}$ is larger, an increase in the savings improves her utility more dramatically when she is old and her child becomes an unskilled worker. In other words, when $G^{O,L}_{t+2} + \bar{G}$ is larger, the larger savings adequately balances the consumption of private goods and the provisions of family public goods when old. On the other hand, an increase in savings reduces her utility when young. Overall, when $G^{O,L}_{t+2} + \bar{G}$ is given as a large $\bar{G}_{n+l-1}/\phi^{l-1}$, $S(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ also becomes large but not as large. Therefore, given $S(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$, she will choose $G^{O,L}_{t+2} + \bar{G}$ that is smaller than $\bar{G}_{t+l-1}/\phi^{l-1}$ to balance the consumption of private goods and the provision of family public goods when old as Figure 3 shows. Similarly, if $G^{O,L}_{t+2} + \bar{G}$ is extremely small like $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l+2}/\phi^{l+2}$, $S(e, n, \bar{G}, \bar{G}_{n+l+2}/\phi^{l+2}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ is also small but not small enough to induce her to choose $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l+2}/\phi^{l+2}$ when old. Thus, when an amount of $G^{O,L}_{t+2} + \bar{G}$ is the intermediate size, under $S(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ with such $G^{O,L}_{t+2} + \bar{G}$, she will choose such $G^{O,L}_{t+2} + \bar{G}$ when old as in Figure 3.

income of an individual. The larger the provision of family public goods by her parent $G^o_{t+1} + \bar{G}$, the more likely the consumptions of private goods and the provisions of family public goods when young and old are balanced even when the total provision of family public goods that she plans is larger.

However, note that, from Lemma 1(ii) and (v), the size of $G^o_{t+1} + \bar{G}$ provided by her parent with which for an individual in region n, placing her child in region $n + l$ and placing her child in region $n + l - 1$ are indifferent is bigger than that with which, for an individual in region $n - 1$, these two are indifferent, as shown in Figure 4.

Finally, consider Lemma 2(ii), the plan of where to place the child in the case where her child becomes a skilled worker, $G^{O,H}_{t+2}$.

As we will see in section 4.3.3, if an individual lives in a region closer to the city ($n \in \{n', ..., \bar{n}\}$), the benefits of family public goods provided by her child who lives in the city in the case where her child becomes a skilled worker are large. Therefore, she will not attract her child in a region closer to the one where she lives by providing positive $G^{O,H}_{t+2}$.

In contrast, if she lives in a region farther away from the city ($n \in \{\bar{n} + 1, ..., N\}$), the benefits are small. Thus, she will attract her child in a closer region other than the city. However, in order to do so, she must provide a large amount of family public goods. For example, if she wants to place her child in region s, her provision of family public goods may be much larger than $G^{O,H}_{t+2} + \bar{G} = \bar{G}_s/\phi^{s-n}$. Differently from an unskilled worker, a skilled worker earns a higher wage in the city, and thus she is willing to live in the city rather than in region s. Therefore, the individual will give up doing so unless her parent's provision of family public goods is extremely large and thus her substantial income is large. [15]

4.3.2 Decision on Where to Live

Consider the case that an unskilled individual freely chooses where to live and work when young. However, suppose that an unskilled individual's choice of education is sill fixed as $edu = e$.

Lemma 3
Suppose the parent of an unskilled individual lives in region n' that is closer to the city ($n' \in \{1, ..., \bar{n}\}$). Suppose for an unskilled individual, the choice of education for her child is fixed in that she gives a higher education for her child. If the parent's provision of family public goods including the one through the government, $G^o_{t+1} + \bar{G}$, satisfies $\frac{\bar{G}_n}{\phi^{n-n'}} \leq G^o_{t+1} + \bar{G} < \frac{\bar{G}_{n-1}}{\phi^{n-n'-1}}$ with $n = n' + 1, ..., \bar{n}$, or $\bar{G}_{n'} \leq G^o_{t+1} + \bar{G}$ with $n = n'$, then an unskilled individual lives in region n.

[15] In Lemma 2(ii), edu and n are fixed. In reality, an individual can freely choose edu and n. As we will see in Lemma 5, a combination of $(edu, G^{O,H}_{t+2})$ where $edu = e$ and $G^{O,H}_{t+2} > 0$ will not be chosen.

See Appendix B for the proof. An individual's choice of the residential region when her parent lives in region n' that is farther away from the city ($n' \in \{\bar{n}+1, ..., N\}$) is discussed in section 4.3.3 together with the decision on education investment for her child.

Focus on the case where her parent lives region one ($n' = 1$) and consider an individual's optimal choices of where she lives and where she places her child and the maximized expected lifetime utility, reacting to the parent's provision of family public goods $G^o_{t+1} + \bar{G}$. The discussion can be applied in the more general case where the parent lives in the other region. The relationship between an individual's expected lifetime utility and her parent's provision of family public goods, $G^o_{t+1} + \bar{G}$, depends on the distance to the region where her parent lives, and the distance to the city.

First, consider the distance to the hometown region where her parent lives. Obviously, the larger her parent's provision of family public goods G^o_{t+1} is, the higher an individual's expected lifetime utility is. More importantly, the impact of an increase in G^o_{t+1} is more prominent when an individual lives closer to the region one in which her parent lives. Therefore, the graph of an individual's expected lifetime utility when she lives in region n minus that when she lives in region $n+1$ has a positive slope, as in Figure 5.

Second, consider the distance to the city. This subsection continues to suppose that an individual gives a higher education for her child. Thus, with probability μ, her child becomes a skilled worker, lives in the city, earns a higher income, and provides a larger amount of family public goods. This induces an individual to live closer to the city. On the other hand, with probability $1-\mu$ her child becomes an unskilled worker. In this case, her child may be willing to live farther from the city unless she provides family public goods in a large amount. This situation may discourage an individual from living closer to the city. If the following condition is satisfied:

$$U(e, n+1, \bar{G}, \bar{G}; L, n', \bar{G}, \bar{G}) - c(P_{n+1}) \leq U(e, n, \bar{G}, \bar{G}; L, n', \bar{G}, \bar{G}) - c(P_n), \tag{25}$$

then the former is the more prominent in these two factors that work in the opposite direction, and an individual is eager to reside in a region nearer to the city even in the case she is an unskilled worker. This condition (25) holds when the probability that a child given a higher education becomes a skilled worker is sufficiently high, and the marginal utility decreases not so dramatically when the family public goods provision by the child increases.

In equilibrium, the population density becomes higher in a region closer to the city. (25) means that the utility is higher when an individual lives closer to the city even when the parent's provision of family public goods is zero ($G^o_{t+1} + \bar{G} = 0$) without taking into account the difference in the population density among regions. Unskilled workers are willing to live in a region that is as close as possible to the city, which entails higher population density in a region closer to the city.

However, in equilibrium, the population density in a region closer to the city is higher to the extent that it more than just cancels out the higher value of family public goods provided by the child in the city. Therefore, unless her parent provides a larger amount of family public goods, an individual is not willing to live closer to the city and the region one where her parent lives.

The reason is as follows. By providing family public goods in a significant amount when old, an individual can attract her child closer to the region where she lives in the case where her child becomes an unskilled worker and is, thus, willing to do so. A skilled worker in the city is the most advantageous in attracting her child closer to her given the high wage she earns. Therefore, she provides large $G^o_{t+1} + \bar{G}$ such as $\bar{G}_n$ $(n = 1, ..., m)$ to attract her child. Unskilled workers in these regions are also advantageous in attracting the child given the large amount of family public goods provided by the parent such as $\bar{G}_n$ $(n = 1, ..., m)$, though not so as a skilled worker in the city. This situation further attracts the populations in regions closer to the city.

Figure 5 shows the graph of an individual's expected lifetime utility when she lives in region n minus that when she lives in region $n+1$ for $n \in \{1, ..., \bar{n} - 1\}$. From (25), if regions have the same population, all the graphs locate above the horizontal line, thus all the unskilled workers are willing to live in region one where the parents live, as discussed above. If a region closer to the city has a larger population to the extent that it just cancels out the higher value of family public goods provided by the child in the city when the child becomes a skilled worker, all the graphs start from the origin, $(0, 0)$. Therefore, when $G^o_{t+1} + \bar{G} = 0$, an unskilled worker's expected lifetime utility is the same wherever she lives. In reality however, in Figure 5, the graph of a region that is closer to the city is located more downwardly, and, thus, intersects with the horizontal axis, not at $G^o_{t+1} + \bar{G} = 0$ but at $G^o_{t+1} + \bar{G}$, which is so large as to satisfy (11) or (12). Moreover, the critical $G^o_{t+1} + \bar{G}$ is larger when the region n is closer to the region one where her parent lives and the city. As we saw, unless her parent provides family public goods in the amount that is large enough to overcome the prominently large population, an individual is not willing to live closer to the region one where her parent lives.

I considered the case where an individual's parent lives in region one, $n' = 1$. However, an individual's optimal behavior when the parent lives in the other region is fundamentally the same. The amount $\bar{G}_n/\phi^n$ of public goods provided by her parent in region one $(n' = 1)$ and amount $\bar{G}_n/\phi^{n-n'}$ of public goods provided in region $n' \geq 1$ bring the same impact on an individual in region n if $n \geq n' \geq 0$. In region n, their values are the same as $\bar{G}_n$. Thus, in either case, the individual will live in region n if $n \geq n' \geq 0$. [16]

[16] If $n < n'$, even when family public goods is larger than $\bar{G}_n/\phi^{n-n'}$, an individual will live in region n' not region n. She will not live closer to the city than her parent. As discussed, the closer to the city one lives, the greater the population. When her parent provides family public goods in a large amount and lives in a region closer to the city than where an individual initially lives, by moving closer to the city and her parent, the value becomes higher. Thus, an individual will live closer to her parent, overcoming the rise in disutility that stems from the larger population. However, if her parent lives in a region farther away from the city than the individual initially lives, by living closer to the city but farther away from her

From the discussions mainly focused on equations (11) or (12) and (13), the parent's provision of family public goods determines where an individual lives and where she plan to place her child. Some equations in simultaneous equations systems describe the population distribution among regions that stems from such behaviors. The population distribution let (11) or (12) and (13) hold. Thus, the same distribution will be reproduced. Therefore, population distribution among regions inside region $\bar{n}$ is unchanged, as long as there is neither population outflow to nor inflow from the regions outside it. The next subsection confirms this notion.

4.3.3 Decision on Education

Finally, consider the decision on whether an unskilled individual gives a higher education for her child. This subsection discusses the intuitions behind the case where the parent lives closer to the city and provides family public goods in a significant amount and the case where the parent lives farther from the city and provides family public goods in a small or zero amount. [17]

To summarize what the candidate of optimal correspondence in section 4.1 states, when her parent resides closer to the city, give a higher education for the individual, and provides family public goods in a significant amount in the case where the individual becomes an unskilled worker despite the higher education, an individual will also reside closer to the city, give a higher education for the child, and plan to provide family public goods in a significant amount in the case where the child becomes an unskilled worker. Inversely, when her parent resides farther from the city and does not give a higher education or provide family public goods, an individual will also reside in the same region, and will not give a higher education or provide family public goods.

I analyze the decision on a higher education for her child using the function $U(e, n, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ and the newly introduced function $\bar{U}(0, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

The former is an individual's minimum possible level of expected lifetime utility conditional on that she gives a higher education for her child and chooses $G^{O,L}_{t+2} = 0$. When she can freely choose $G^{O,L}_{t+2}$, her utility is obviously higher than or, at least, equal to $U(e, n, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ where $G^{O,L}_{t+2}$ is fixed to zero. With $G^{O,L}_{t+2} = 0$ her child will move to the less populated region such as $\bar{n}$ in the case where the child becomes an unskilled worker. Thus, she can improve her utility by attracting her child with positive $G^{O,L}_{t+2}$.

The latter is an individual's maximum possible level of lifetime utility conditional on

parent, the value becomes lower. Her utility, thus, necessarily falls.

[17] The the case where the parent lives farther from the city whereas provides family public goods in a large amount and the opposite case where the parent lives close to the city and provides family public goods in a very small amount are implausible. However, it is necessary to analyze them to obtain the optimal correspondences, as explained in the Appendix B.

that an individual lives in region n and does not give a higher education for her child, and her child remains in the region where she lives even if the individual does not provide family public goods when old ($G^{O,L}_{t+2} = 0$). In reality, her child will not necessarily remain in the hometown but move to a less populated region. Thus, an individual's maximum expected lifetime utility conditional on that she lives in region n and does not give a higher education is obviously lower than or, at most, equal to $\bar{U}(0, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

First, using these functions, consider an individual's choice of education investment in the case where the government and the parent do not provide family public goods, ($\bar{G} = 0$ and $G^o_{t+1} = 0$). In this case, where her parent lives does not affect an individual's decision.

As an individual resides in a region closer to the city where her child will provide family public goods in a large amount in the case that her child becomes a skilled worker, the gap $U(e, n, 0, 0; L, n', 0, 0) - \bar{U}(0, n; L, n', 0, 0)$ rises, and thus she will give a higher education for her child more prominently. The reason is as follows. As discussed in the last subsection, if (25) holds, as an individual resides in a region closer to the city where her child lives and provides family public goods in a large amount in the case where her child becomes a skilled worker, $U(e, n, 0, 0; L, n', 0, 0)$, the minimum possible level of expected lifetime utility conditional on that she gives a higher education for her child rises, without considering the increase in the population. In contrast, $\bar{U}(0, n; L, n', 0, 0)$, the maximum possible level of lifetime utility conditional on that she does not give a higher education is independent of the distance to the city where her child will not reside, without taking into account the difference in the population.

From these facts and (15) and (16), $\bar{n}$ is the region that is farthest from the city where the benefits from higher education for her child will exceed the costs. [18] That is, given that regions where she can live are limited to $\{0, ..., \bar{n}\}$ and her parent does not provide family public goods, an individual will give a higher education for her child. Moreover, from Lemma 3, her optimal choice of where to live is $n = \bar{n}$. [19] In contrast, given that regions in which she can live are limited in $\{\bar{n} + 1, ...\}$, an individual will not give a higher education, $edu = 0$. The distance to the city discourages her to do so. Moreover, her utility is the same wherever in $\{\bar{n} + 1, ...\}$ she lives, as the regions have the same population as (21) indicates.

From (19), if the government and her parent do not provide family public goods, to live in $\bar{n}$ giving a higher education, and to live in a region farther away from the city than $\bar{n}$, $n \in \{\bar{n} + 1, ...\}$ without giving a higher education are indifferent. Moving from $\bar{n}$ to $\bar{n} + 1$, which is farther from the city, the benefits from a higher education for her child

[18] When $n = \bar{n}$ and $n = \bar{n} + 1$, the expected lifetime utility conditional on that she gives a higher education for her child is not higher than $U(e, n, 0, 0; L, n', 0, 0)$ but equal to it. Also, when $n = \bar{n}$ and $n = \bar{n} + 1$, the expected lifetime utility conditional on that she does not give a higher education for her child is not lower than $\bar{U}(0, n; L, n', 0, 0)$ but equal to it. See Appendix B for the details.

[19] By living more closely to the city, the benefits rise, whereas the population increases. Therefore, when her parent does not provide family public goods, an individual is not willing to do so.

that are realized in the city becomes lower. By giving up a higher education, her utility is improved. Even so, the improved utility in region $\bar{n}+1$ remains lower than that when she lives in region $\bar{n}$ that is closer to the city and gives a higher education. However, the population is smaller in region $\bar{n}+1$ than $\bar{n}$. It just compensates for the drop in the utility caused by living farther away from the city mentioned above.

Next, based on the discussions above, consider the case where the government provides family public goods but the parent does not ($\bar{G} > 0$ and $G^o_{t+1} = 0$). In this case where her parent lives affects an individual's decision. Again, from (15) and (16), $\bar{n}$ is the region farthest from the city where the benefits from higher education for her child will exceed the costs. However, differently from the case above where $\bar{G} = 0$, to live in $\bar{n}$ giving a higher education and to live in a region farther away from the city than $\bar{n}$ without giving a higher eduation are not indifferent. Which is higher depends on where the parent lives. The family public goods provided by the government for her parent attract the individual to her hometown or a nearby region. From Lemma 3, (17) and (18), an individual whose parent lives in $n \in \{0, ...\bar{n}\}$ will live in region $\bar{n}$. Oppositely, an individual whose parent lives in $\in \{\bar{n}+1, ...\}$ does not have an incentive to give a higher education for her child or leave her hometown.

Finally, consider an individual's choice of education investment in the case where the parent lives closer to the city ($n' \in \{0, ..., \bar{n}\}$) and provides family public goods in a significant amount ($G^o_{t+1} > 0$).

When her parent increases the amount of family public goods provision, $U(e, n, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$, the minimum possible level of the expected lifetime utility conditional on that she lives in region n and gives a higher education rises more dramatically than $\bar{U}(0, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$, the maximum possible level of lifetime utility conditional on that she lives in region n and does not give a higher education rises. Her parent's provision of family public goods increases an individual's substantial income, which reduces the substantial cost of education. This makes an individual more in favor of giving a higher education for her child. Therefore, if regions in which she can live are limited in the ones closer to the city ($n \in \{0, ...\bar{n}\}$), an individual will give a higher education for her child more prominently. [20]

In contrast, given that regions in which she can live are limited in $\{\bar{n}+1, ...\}$, she will live in region $n = \bar{n}+1$ that is closest to her parent providing family public goods in a region in $\{0, ..., \bar{n}\}$. However, whether she will give a higher education is ambiguous given the lower substantial cost of education. Nevertheless, in any case, an individual's utility when she lives in $\bar{n}+1$ is strictly lower than her utility when she lives in $\bar{n}$ and gives a higher education, $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. When her parent increases the

[20] Moreover, as in sections 4.3.1 and 4.3.2, she can improve her utility by living more closely to the city and the parent and planning to provide positive amount of family public goods in order to attract her child closer to her in the case where her child becomes an unskilled worker.

provision of family public goods, the former increases but less dramatically than the latter, since in the former case the distance to the parent is longer.

This is the intuitions behind decisions on education investment for her child that the correspondences in section 4.1 indicate. As discussed in the final part of the last subsection, as long as there is no population flow between the group of regions $\{0, 1, ..., \bar{n}\}$ and that of $\{\bar{n} + 1, ..., N\}$, the population distribution among regions remains unchanged. This subsection confirms the case to be true.

4.4 Behavior of A Skilled Worker

4.4.1 Decision on Where to Place the Child When a Skilled Worker Lives in the City

Consider a skilled worker's decision on where to place the child in the case where her child becomes an unskilled worker. Individuals compete to attract their children to a region that is as close as possible to the region where they live.

First, consider the competition between skilled workers and unskilled workers. Regions near the city have large populations. Therefore, an individual cannot place her child in these regions without providing a large amount of $G^{O,L}_{t+2}$. A skilled worker in the city is advantageous in the competition due to the higher wages. They will place her child in a region in $\{1, ..., m\}$ that is closer to the city by providing a large amount of $\bar{G}_n/\phi^n$ where $n \in \{1, ..., m\}$. Inversely, this is the reason city and regions near the city that have rich individuals have large populations. In contrast, unskilled workers will place her child in a region in $\{m, ..., \bar{n}\}$ by providing smaller $\bar{G}_n/\phi^{n-n'}$, where $n \in \{m, ..., \bar{n}\}$.

Second, consider the competition among skilled workers in the city. The value of family public goods provided by her parent is higher for a skilled worker whose parent lives closer to the city (n' is smaller). Her parent does not privately provide family public goods ($G^{O,H}_{t+2} = 0$), but she has the family public goods publicly provided, and the value evaluated in the city, $\bar{G}\phi^{n'}$, is higher in the case where her parent lives closer to the city (n' is smaller). Therefore, a skilled worker whose parent lives closer to the city is advantageous in attracting her child to a region closer to the city in the case where her child becomes an unskilled worker.

Specifically, a skilled worker whose parent also lives in the city has the largest substantial income and, thus, attracts her child to a region closer to the city in $\{1, ..., r\}$ by choosing a large $\bar{G}_n$. [21] Inversely, a skilled worker whose parent lived in a region farther away from the city will plan to place her child in a region not so close to the

[21] Note that a skilled worker will not place her child in the city when her child becomes an unskilled worker, in the case where the city has an extremely large population. In this case, without providing an extremely large amount of family public goods, she cannot attract her child to the city. That amount of family public goods $\bar{G}_0$ satisfies (11) with $n = n' = 0$ and l of $1 \in \Theta_l$, yet it is so large as to satisfy (10).

city, $n \in \{r, ..., m\}$, providing a smaller amount of family public goods when old, as her substantial income is smaller, though only slightly.

4.4.2 Decision on Where to Live

I show that if (8) is satisfied, and if her parent lives in a region that is not very far from the city and provides family public goods in an amount that is not significant, then a skilled worker is willing to live in the city.

To consider where a skilled worker lives, compare the maximum expected lifetime utility of the unskilled worker discussed thus far and the maximum expected lifetime utility when she lives in the city. The reason is as follows. If a skilled worker lives in a region other than the city, her skill does not contribute. She earns wage w_L that is the same as that of an unskilled worker. Therefore, given that she lives in a region other than the city, a skilled worker's choice depends on her parent's choice of n' and $G_{t+1}^{O,H}$, exactly in the same manner an unskilled worker's choice depends on n' and $G_{t+1}^{O,L}$.

From the discussions thus far, if her choice to live is limited in $n \in \{1, ..., \bar{n}\}$, and if her parent lives in $n' \in \{0, ..., \bar{n}\}$ and does not privately provide family public goods ($G_{t+1}^{O,H} = 0$), a skilled worker lives in region $\bar{n}$, just as an unskilled worker will do so if $G_{t+1}^{O,L} = 0$. The maximum expected lifetime utility of an unskilled worker is the highest when her parent lives in region $\bar{n}$ ($n' = \bar{n}$), as the value of family public goods publicly provided to her parent, $\phi^{n'}\bar{G}$, is the highest when she lives in the same region in which her parent lives. That is, $\phi^{n'}\bar{G} \leq \phi^{\bar{n}}\bar{G}$ for $n' \in \{1, ..., \bar{n}\}$. That is, the left-hand side of (8).

In contrast, when a skilled worker lives in the city, she earns higher wages w_H. The minimum possible value of maximum expected lifetime utility of a skilled worker in the city is the maximum expected lifetime utility of a skilled worker whose parent lives in region $\bar{n}$ and does not privately provide family public goods. In equilibrium, region $\bar{n}$ is the farthest region from the city among the ones with investment in higher education. Therefore, if her parent lives in region $\bar{n}$ and does not privately provide family public goods, the value of family public goods evaluated in the city is the lowest, $\phi^{\bar{n}}\bar{G}$. That is, the right-hand side of (8).

Therefore, (8) means that the highest possible value of maximum expected lifetime utility when a skilled worker lives in a region other than the city is lower than the lowest possible value of maximum expected lifetime utility when she lives in the city. That is, the closer the region where her parent lives is to the city, the higher the maximum expected lifetime utility of a skilled worker when she lives in the city, the lower that when she lives in a region other than the city, and thus the larger the gap. Therefore, if (8) holds, when her parent lives in $n' \in \{0, ..., \bar{n}\}$ and does not provide family public goods ($G_{t+1}^{O,H} = 0$), a skilled worker will live in the city.

As discussed thus far, the larger the amount of family public goods the parent provides

G^o_{t+1} is, the closer to her parent an unskilled worker is willing to live, and, thus, the more likely a skilled worker wants to live in a region that is close to her parent rather than the city despite the lower wages there. However, if (8) holds, as long as her parent's provision of family public goods, including the ones though the government, $G^o_{t+1} + \bar{G}$, is not extremely large, a skilled worker will still live in the city. Moreover, as discussed in section 4.3.1, the parent of a skilled worker will not privately provide family public goods ($G^{O,H}_{t+1} = 0$). Therefore, a skilled worker will live in the city.

Summarizing the discussions in this section yields the following Proposition.

Proposition 1
Suppose that when $\bar{G} = 0$ there exist $\bar{n}$, m, r and sets Θ_l ($l = 0, ..., m-1$) under which the solutions of simultaneous equations system shown in Section 4.2 and Appendix A.2 exist and satisfy $Q_n \in [0, 1]$ and $Q_{n\to l} \in [0, 1]$. Suppose also that when $\bar{G} > 0$ there exist $\bar{n}$, m, r, sets Θ_l ($l = 0, ..., m-1$) and sets Ω_l ($l = 1, ..., m$) under which the solutions of simultaneous equations system shown in Section 4.2 and Appendix A.2 exist and satisfy $Q_{n'\to 0\to n} \in [0, 1]$ and $Q_{n\to l} \in [0, 1]$. Then correspondences shown in Section 4.1 and Appendix A.1, which have the solutions as critical values, are the optimal correspondences.

See Appendix B for the proof.

5 Application and Extension

5.1 Multicentric Equilibria

Consider an economy that has multiple cities. Let K denote the number of cities. Each city has $\bar{n}$ regions from the city to the west and from the city to the east, respectively. In these regions individuals give a higher education for her child, and the child will move to the city if she becomes a skilled worker.

The population, the wage, the number of nearby regions $2\bar{n}$ from which skilled workers migrate to the city, and the population distribution among the city and nearby regions are the same among K cities, if they share the same economic backgrounds and if any two cities separate by at least $2\bar{n}$ regions. Equilibrium can be obtained by solving the model constructed thus far, in which P_0 and P_n for $n \in \{1, ..., \bar{n}\}$ are the populations of each city and regions near the city. Therefore, the total populations of cities and nearby regions in an economy are KP_0 and $2K\Sigma^{\bar{n}}_{n=1}P_n$, respectively. Also, population of a region that is farther away from the city than region $\bar{n}$ and where an individual does not give a higher education is not (21) but the following:

$$P_n = \frac{P - KP_0/\lambda}{\bar{N} - K(1 + 2\bar{n})}. \tag{26}$$

Consider whether the population in cities (the population of skilled workers), population with higher education, and the number of regions where higher education is observed in the whole of an economy with multicentric equilibrium are larger than those in an economy with monocentric equilibrium.

Accordingly, assume these are the same. That is, each city has a population of P_0/K, where P_0 is the population in the city in an economy with monocentric equilibrium. The region that is the farthest from the city for the regions where higher education is observed is $\bar{n}/K$ (or the integer that is the closest to it) regions apart from the city.

If the city in the case of monocentric equilibrium has an excessively large population, the wage in the cities in multicentric equilibrium is lower, but not extremely so. Accordingly, the provision of family public goods by a skilled worker when young in the city is also smaller, but not extremely so. Therefore, the value of the provision of family public goods by a skilled worker in region $\bar{n}/K$ in multicentric equilibrium can be higher than that in region $\bar{n}$ in monocentric equilibrium. If so, in multicentric equilibrium, an individual in regions that are apart from the city by $\bar{n}/K$ or farther will also give a higher education for the child. Thus, the total population in cities, the population with higher education, and the number of regions where higher education is observed exceed those in monocentric equilibrium, though the population size in each city remains smaller than that in the city in monocentric equilibrium as in Figure 6.

Many individuals can be better off. As noted, population size in each city in multicentric equilibrium is smaller than that in monocentric equilibrium. Thus, the wage in the city is lower, but not extremely so. The utility loss caused by a fall in wages can be sufficiently compensated by the utility gain caused by less congestion in the city. Hence, the utility of a skilled worker in the city will be improved. A larger population will have higher education, and the total population in cities in multicentric equilibrium is larger than that in monocentric equilibrium, though the population in each city is smaller. Therefore, a larger population is better off.

Moreover, multicentric equilibrium mitigates the inequality problem. First, a larger population will have higher education. Second, in multicentric equilibrium, the advantage of unskilled workers in regions near the city enjoying large amounts of family public goods provided by the parents in the city is less prominent, as the wage in the city is lower.

Figure 7 shows an example of the lifetime expected utilities in monocentric equilibrium and in multicentric equilibrium. Except unskilled workers who luckily have the parent in the city and live in regions near the city, individuals are better off in multicentric equilibrium.

Inversely saying, in monocentric equilibrium, the population size in the city tends to be excessively large. An individual gives a higher education for her child to maximize not her child's utility but her own utility. An individual aims to enjoy family public goods in the larger amount provided by her child when her child succeeds in higher education, expecting that in this case, her child will move to the city as long as the utility in the city is higher than that in the other regions, though it is only slightly so. Therefore, as long as the city is not very far from the region where she lives, an individual will give a higher education for her child. Thus, the city size will be excessively large.

However, it is difficult for an economy to move from monocentric equilibrium to multicentric equilibrium. Once monocentric equilibrium with a big city is achieved, it is challenging to set up a new city in an area that is far from the current big city by inducing an individual in that area to give a higher education for the child. It is impossible if that area does not sufficiently large population in total. Even when the area has a sufficiently large population, the expectation that a new city will be soon formalized must be shared by the individuals in the area. When such an expectation is not shared and, thus, some individuals in the area give a higher education for their children whereas others do not, the child who becomes a skilled worker moves to the current big city that is far from the hometown.

Inversely, once multicentric equilibrium with a proper number of cities occurs, it can stably continue. Suppose that now the size of each city is optimal. When an individual gives a higher education for her child and the child becomes a skilled worker, the child moves to the city that is nearest to the region where she lives as she hopes and expects. The child will not move to the other particular city, since if that city accommodates more population, the utility in the city becomes lower than that in the other cities. Therefore, multiple cities keep accommodating the same optimal population and thus the multicentric equilibrium continues.

In contrast, multicentric equilibrium where the number of cities is overly large and, thus, the size of each city is insignificant is unstable. When a specific city attracts more population, the utility in that city is improved.

5.2 Social Security Policy

This section considers the effects of the social security policy. Its introduction reduces the dependency of an individual on the child in the provision of family public goods. Therefore, an individual may be less willing to give a higher education to enjoy the child's provision of family public goods in a larger size. If so, the population of skilled workers P_0 and the geographical size of urban area $\bar{n}$ will shrink. In reality, however, the net benefits of higher education become larger. Thus, the introduction of the social security policy induces higher education, which expands the urban size in its population and area as in Figure 6. Moreover, it enhances economic welfare as in Figure 7.

To show it, it is sufficient to show that an increase in $\bar{G}$ induces an individual in region $\bar{n}+1$ to switch her choice from not giving a higher education for her child to giving it. More specifically, an increase in $\bar{G}$ raises the left-hand side of (16) more dramatically than it raises the right-hand side. Hence, (16) is harder to hold.

First, consider the substantial cost of higher education. An increase in $\bar{G}$ does not change an individual's lifetime disposable income given the property of the fully funded social security system. However, pension benefits take the form of family public goods when an individual is old, not cash. Therefore, the family public goods provided by the government for the parent benefit the individual. Hence, an increase in $\bar{G}$ raises the substantial disposable income, and it reduces the substantial cost of higher education as in Lemma 1(iv). [22]

Second, the benefits of higher education grow. Specifically, the difference in the total family public goods when old (the sum of $\bar{G}$ and the value of G^y_{t+2} in region $\bar{n}+1$) between the case where her child becomes a skilled worker and the case where her child becomes an unskilled worker expands. An increase in $\bar{G}$ reduces the dependency of an individual on the child's provision of family public goods. At the same time, the child depends more on the individual's provision through the government $\bar{G}$, but more drastically so in the case where the child becomes an unskilled worker and remains in the hometown with the individual. That is, in that case, the child reduces the contribution more drastically. Therefore, an individual finds that the total provision of family public goods increases in both cases but does so more drastically in the case where her child becomes a skilled worker.

The situation is further detailed as follows. First, consider the case where an individual in region $\bar{n}+1$ does not invest in higher education and thus her child becomes an unskilled worker. The child will live together with her in region $\bar{n}+1$, as, by doing so, the child can maximize the benefits from the family public goods provided by the individual through the government, $\bar{G}$. An increase in $\bar{G}$ makes the child depend more on the individual and reduce the provision of family public goods for herself. Moreover, as she lives in the same region, for the individual the impact of the reduction is direct and thus large.

Next, consider the case where an individual in region $\bar{n}+1$ gives a higher education

[22] The substantial costs of a higher education are the difference in the utility from the private goods when young and old (C^y_{t+1} and C^o_{t+2}), and family public goods when young ($G^o_{t+1}+\bar{G}+G^y_{t+1}$) between the cases where $edu = 0$ and $edu = e$. In this sense, Lemma 1 (iii) and (iv) show how the substantial costs of a higher education change when G^o_{t+1} and $\bar{G}$ change, respectively. They show that when G^o_{t+1} and $\bar{G}$ increase, the substantial costs decrease. Note that in the conditional expected lifetime utility function, $G^o_{t+2}+\bar{G}$ is fixed. Thus, the utility from the total provision of family public goods when old is also fixed. When G^o_{t+1} changes, the optimal choices of private goods when young and old (C^y_{t+1} and C^o_{t+2}) and family public goods when young ($G^o_{t+1}+\bar{G}+G^y_{t+1}$) change, whereas the total family public goods provision when old ($G^o_{t+2}+\bar{G}+G^y_{t+2}$) does not change. In contrast, when $\bar{G}$ changes, $G^o_{t+2}+\bar{G}+G^y_{t+2}$ also changes simply because $\bar{G}$ is included. However, Lemma 1(iv) considers the change in the conditional expected lifetime utility caused by the changes in the optimal choices of C^y_{t+1}, C^o_{t+2}, and G^y_{t+1} when $\bar{G}$ changes, which are caused by a change in taxation and the family public goods from the parent through the government when young. This situation does not consider the change in the total provision of family public goods when old, $G^o_{t+2}+\bar{G}+G^y_{t+2}$. Therefore, to consider the change in the net benefits of higher education caused by an increase in $\bar{G}$, beyond Lemma 1(iv), it is essential to consider the difference in the change in utility from family public goods when old ($G^o_{t+2}+\bar{G}+G^y_{t+2}$) between the case where $edu = 0$ and the case where $edu = e$.

to her child and her child becomes a skilled worker. In this case, the child will move to the city; that is, $\bar{n}+1$ regions far from the region where the individual lives. Therefore, the impact of an increase in the individual's provision of family public goods through the government $\bar{G}$ on her child's utility is much smaller. That is, an increase in $\bar{G}$ makes the child depend more on her parent and reduces the provision of family public goods, though much less drastically. Again, as she lives far from the city, for the individual, the impact of the reduction of the child's provision of family public goods is slight.

In summary, for an individual in region $\bar{n}+1$, an increase in the public provision of family public goods $\bar{G}$ increases the total family public goods provision more drastically in the case where her child becomes a skilled worker. [23] Moreover, the substantial costs fall. Thus, the net benefits of higher education increase.

Thus far I discussed the impact of an increase in $\bar{G}$, other things being equal. If an increase in $\bar{G}$ is slight, no change in equilibrium takes place. However, if it is significant, individuals in region $\bar{n}+1$ will shift their decision on investment in higher education from $edu = 0$ to $edu = e$. It increases the population with higher education and that of skilled workers in the city, and it raises the wage. It further induces individuals to give a higher education and, thus, the urban size expands in its population and area as in Figure 6. Moreover, all the individuals can be better off as in Figure 7. A larger population will have higher education. It raises the wage in cities, and enhances the utilities of unskilled workers near the cities.

6 Conclusion and Discussion

This study constructs a model with family public goods, with the advantage of geographical proximity within a family unit, and human capital, with the advantage of geographical proximity within a city stemming from local in-scope knowledge spillovers to analyze the dynamic process of accumulation and agglomeration of human capital. An individual's location choice depends on the relative size of these benefits of public goods within a family unit and knowledge spillovers within a city. When the hometown is farther from the city, by migrating to the city, the benefits of family public goods become smaller. However, when an individual is a skilled worker, she will be attracted by a high wage in the city. Considering such propensities of her child's decision on location, an individual decides on education.

[23] Even when the total family public goods when old in the case where her child becomes a skilled worker grows more dramatically than in the case where her child becomes an unskilled worker, it does not necessarily follow that the utility from the total family public goods when old in the case where $edu = e$ increases more dramatically than that in the case where $edu = 0$. In the case where $edu = e$, the total family public goods when old is larger, and thus, the marginal utility is lower. Therefore, the utility does not rise as drastically. However, in the case where the individual lives in region $\bar{n}+1$, the gap in the total family public goods when old between the cases where $edu = e$ and $edu = 0$ is almost equal.

In the steady state, a city and nearby regions host a larger population, discouraging the migration from farther regions to the area. This situation seriously divides society into two: those inside and outside the city and nearby regions.

For those inside, families enjoy education investment and higher substantial income from generation to generation because higher substantial income is inherited in the form of not only higher education but also family public goods that the parent provides to attract the child to a region as close as possible to the city in the case where the child becomes an unskilled worker. Individuals who become unskilled workers despite the higher education from their parents move from the city or regions near the city to farther away regions but remain within the area. Conversely, skilled workers move from these regions to the city.

However, in regions far from the city, families perpetually remain in regions farther away from the city as unskilled workers with lower substantial income. The parent has no incentive to move to a region closer to the city given the high population density. She also has no incentive give a higher education for the child. The city where the child will live and provide larger family public goods in the case where the child becomes a skilled worker is exceedingly far from the region for individuals there to enjoy it.

The model has multiple human capital agglomeration patterns. The population size in the city tends to be excessive in monocentric equilibrium. However, in multicentric equilibrium, the population in each city is smaller than that in monocentric equilibrium, but the total population in cities is larger. In multicentric equilibrium, the wage in the city is lower, but not extremely so. Rather the disutility from the larger population is mitigated more drastically, improving the utility of a skilled worker in the city. Moreover, in any region in the economy, the average distance from the city to it is much shorter, which induces an individual to give a higher education. Therefore, multicentric equilibrium mitigates the intergenerational persistent social gap between urban and peripheral areas.

The study also discusses the effects of the fully-funded pension system, where the pension benefits take the form of family public goods provision. This social security policy does not change the lifetime disposable income given the property of the fully funded system. However, publicly provided family public goods when the parent is old also benefit an individual. It then increases the individual's substantial lifetime income and reduces the substantial costs of higher education. However, such an individual's dependency on the parent's provision of family public goods makes her reduce her provision of family public goods. The child also reduces her contribution but more drastically so when she becomes an unskilled worker. Thus, the difference in the total family public goods when an individual is old between the case where her child becomes a skilled worker and the case where her child becomes an unskilled worker expands. Therefore, the social security policy induces higher education and enhances economic welfare.

The framework in this paper provides a first step to further study of interregional migration and agglomeration process incorporating within a family interactions and decisions. I point out several directions of future research.

First, this study does not explicitly introduce land and housing markets. However, from the results, we can infer the land price and housing rent in the case where the land and housing markets are introduced. These are high in a region with a higher population density and a higher substantial income. The result that high population densities in a city and nearby regions discourage immigration of unskilled workers in the area corresponds to the phenomena in the real world where expensive living costs in the area discourage immigration. Gyourko, Mayer and Sinai (2013) shows that gap in housing price is more extreme than that in income. Ganong and Shoag (2017) shows that high housing prices in high income area deter unskilled workers immigration.

Note that housing is complementary to family public goods, which is complementary to land. In the city, land prices and rents are very high. Thus, it is harder for an individual to provide family public goods in large amounts. This situation may mitigate intergenerational immobility.

Second, in this study, the parent gives a higher education for the child to enjoy the larger size of family public goods provided by the child who becomes wealthier in the city. However, there are many other motives for educational investment for a child. Ehrlich and Lui (1991) consider the intergenerational trade. Many studies discuss altruistic motives. For example, Glomm and Ravikumar (1992) consider the parental altruism directed from the child's human capital, and Banerjee and Newman (1993) consider the warm glow motive of intergenerational transfer. However, as long as the advantage of geographical proximity of family members exists, the distance still matters. An individual gets utilities from education investment, the child's human capital, and the child's provision of family public goods. The benefits from the child's family public goods fall as an individual lives farther away from the city, though the benefits from education investment and the human capital are independent of the distance from the city. Thus, an individual who lives far away from the city will hesitate to give a higher education, unless the weight of parental altruism on the utility is adequately large. Therefore, if the model introduces parental altruism, education investment and, thus, the skilled worker increases on average, but the interregional distributions of skilled and unskilled workers, education investment, and the interregional migration pattern patterns do not change fundamentally. Together with intergenerational transfers motivated by altruism and human capital externalities, geography plays a crucial role in determining intergenerational immobility.

Third, this study focuses on the steady-state equilibrium where unskilled workers do not move to cities. They do not move to a region between a city and the hometown. However, to trace interregional migration patterns historically, it is important to analyze transitional dynamics. In the process of economic development, unskilled workers as well as skilled workers migrate from rural to urban areas as described by Lewis (1954) and Harris and Todaro (1970). In a developed economy as well, unskilled and skilled workers migration to cities continues driven by agglomeration forces as transportation and transaction costs fall, as analyzed in new economic geography models. Moreover, this study shows that there exist multiple steady-state urbanization patterns and that it is challenging to move from one urbanization pattern to the other. Therefore, to investigate which pattern will be achieved, it is vital to analyze transitional dynamics to elucidate which equilibrium will be achieved and the stage of urbanization from which it is still possible to avoid the monocentric but achieve the preferable multicentric equilibrium. However, it is formidably difficult to analyze the dynamics including the transition. Thus, it is useful to rely on the simulation analysis.

Finally, note the geographical scope that the model fits when we apply public policies analyzed in this paper. The model in this paper relates to the literature on income convergence across states derived from the neoclassical economic growth theory by Barro and Sala-i-Martin (1992) in that migration plays a crucial role. In a low-income state, the wage is low and the return of capital is high. Thus, investment in physical capital is intensive in low-income states but labor migrates to high-income states, which entails per capita income convergence across the states. Urban agglomeration does not contradict the convergence across states in a country or across countries in, e.g., EU. Across states which have several or at least one city, the gaps in industrial structures, occupations, and earnings shrink and migrations fall. Kaplan and Schulhofer-Wohl (2017) shows that interstate migration in the U.S. is falling as the gaps across the states decline. However, the gap between city and rural areas in a state, and the gap between states or countries that succeed in hosting cities and the ones that failed are problematic, especially when migrations across the areas are falling despite that. In this situation, it is important to discuss the allocation of policies including education policy and social security policy across governmental tiers.

Appendix A: Full Description of Optimal Correspondences and Simultaneous Equations System

1 Optimal Correspondences

1.1 An Unskilled Worker's Behavior

Group regions $n = 0, 1, ..., N$ into four sets: $\{0\}$, $\{1, ..., m\}$, $\{m, ..., \bar{n}\}$, and $\{\bar{n}+1, ..., N\}$. Then group regions in the second and third sets into m sets, Θ_l where $l = 0, ..., m-1$. This is constructed in the manner that $1 \in \Theta_{m-1}$ and $\bar{n} \in \Theta_0$. If n is the maximum in Θ_l, then $n+1$ is in Θ_{l-1}. Otherwise, $n+1$ is also in Θ_l.

Case L(i)

The parent lives in the city or region n' that is closer to the city than $\bar{n}$. That is, $n' = 0, ..., \bar{n}$.

- If the parent's provision of family public goods, including the one through the government, $G^o_{t+1} + \bar{G}$, satisfies $0 < G^o_{t+1} + \bar{G}$ and $\frac{\bar{G}_n}{\phi^{n-n'}} \le G^o_{t+1} + \bar{G} < \frac{\bar{G}_{n-1}}{\phi^{n-n'-1}}$, where $n = n'+1, ..., \bar{n}$ or $\bar{G}_n \le G^o_{t+1} + \bar{G}$, where $n = n'$, then

$$\begin{cases} edu(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = e, \\ \psi(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = n, \\ Q\left(\bar{G}, \frac{\bar{G}_{n+l}}{\phi^l} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n \to n+l}, \\ Q\left(\bar{G}, \frac{\bar{G}_{n+l+1}}{\phi^{l+1}} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n \to n+l+1}. \end{cases} \tag{3}$$

  If $G^o_{t+1} + \bar{G}$ is the minimum in the range, that is, $G^o_{t+1} + \bar{G} = \frac{\bar{G}_n}{\phi^{n-n'}}-$ and n is not the maximum in Θ_l, then $Q_{n \to n+l} \in [0,1]$, $Q_{n \to n+l+1} \in [0,1]$, and $Q_{n \to n+l} + Q_{n \to n+l+1} = 1$. If n is the maximum in Θ_l or $G^o_{t+1} + \bar{G}$ is larger, then $Q_{n \to n+l} = 1$ and $Q_{n \to n+l+1} = 0$.

- If $G^o_{t+1} + \bar{G}$ satisfies $0 < G^o_{t+1} + \bar{G} = \frac{\bar{G}}{\phi^{\bar{n}-n'}}$ and if (17) holds with bind,

$$\begin{cases} \left(edu(L, n', G^o_{t+1} + \bar{G}, \bar{G}), \psi(L, n', G^o_{t+1} + \bar{G}, \bar{G})\right) = (e, \bar{n}), (0, \bar{n}+1), \\ Q\left(\bar{G}, \bar{G} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = 1. \end{cases}$$

  If (17) holds with inequality, the optimal correspondence is (3) with $n = \bar{n}$, $l = 0$ and $Q_{\bar{n} \to \bar{n}} = 1$.

- If $G^o_{t+1} + \bar{G}$ is small enough to satisfy $0 < G^o_{t+1} + \bar{G} < \frac{\bar{G}}{\phi^{\bar{n}-n'}}$, and if G^o_{t+1} and n' additionally satisfy the following,

$$U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) \leq U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}), \tag{27}$$

$$U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) \leq U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}), \tag{28}$$

then the optimal correspondence can still be (3), where $n = \bar{n}$, $Q_{\bar{n}\to\bar{n}}$=1 and $Q_{\bar{n}\to\bar{n}+1} = 0$. If (27) and (28) do not hold,

$$\begin{cases} edu(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = 0, \\ \psi(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = \bar{n} \;\text{ or }\; \bar{n}+1, \\ Q\left(\bar{G}, \bar{G} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = 1. \end{cases} \tag{29}$$

If (27) is satisfied but (28) is not, then an individual lives in region $\bar{n}+1$.

If (28) is satisfied but (27) is not, then an individual lives in region $\bar{n}$.

If (27) or (28) is not satisfied, and if G^o_{t+1} and n' satisfy

$$U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) \geq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}), \tag{30}$$

then she lives in region $\bar{n}$. Otherwise, she lives in region $\bar{n}+1$.

Case L(ii)

The parent lives in region n', which is farther away from the city than $\bar{n}+1$. That is, $n' = \bar{n}+1, ..., N$. Let $G^*_{n'}$ denote $G^o_{t+1} + \bar{G}$ that satisfies:

$$\begin{aligned} & U(0, n', \bar{G}, \bar{G}; L, n', G^*_{n'}, \bar{G}) \\ = MAX\big[\; & U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^*_{n'}, \bar{G}), \\ & ... \\ & U(e, n', \bar{G}, \bar{G}; L, n', G^*_{n'}, \bar{G}) \quad \big]. \end{aligned} \tag{31}$$

The larger n' is, the larger $G^*_{n'}$ is, and the less plausible $G^*_{n'}$ exists.

- If $G^*_{n'}$ exists and if the parent's provision of family public goods including the one through the government, $G^o_{t+1} + \bar{G}$, is large enough to satisfy $G^*_{n'} \leq G^o_{t+1} + \bar{G}$, then

$$\begin{cases} edu(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = e, \\ \psi(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = n, \\ Q\left(\bar{G}, \bar{G} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = 1, \end{cases} \tag{32}$$

where n is between $\bar{n}$ and n'.

- If $G^*_{n'}$ exists and if $G^o_{t+1} + \bar{G}$ is positive but it is smaller than $G^*_{n'}$ $(0 < G^o_{t+1} + \bar{G} < G^*_{n'})$, then

$$\begin{cases} edu(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = 0, \\ \psi(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = n', \\ Q\left(\bar{G}, \bar{G} \mid L, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = 1. \end{cases} \tag{4}$$

- If $G^*_{n'}$ does not exist, then for any $0 < G^o_{t+1} + \bar{G}$ the optimal correspondence is (4).

Case L(iii)

If $\bar{G} = 0$ and $G^o_{t+1} = 0$, then for all n',

$$\begin{cases} (edu(L, n', 0, 0), \psi(L, n', 0, 0)) = (e, \bar{n}), (0, n), \\ Q\left(0, 0 \mid L, n', 0, 0\right) = 1, \end{cases} \tag{33}$$

where $n = \bar{n} + 1, ..., N$.

1.2 A Skilled Worker's Behavior

Group regions $n = 0, ..., \bar{n}$ into m sets, Ω_l where $l = 1, ..., m$. This is constructed in the following manner. The city $(n = 0)$ is included in several Ω_l where $l = 1, ..., r$. Region $\bar{n}$ is included in Ω_m. If n is the maximum in Ω_l, then it is included in Ω_{l+1} as the minimum element. Otherwise, n is only in Ω_l. If $\bar{G} = 0$, all of $n = 0, ..., \bar{n}$ are included all Ω_l of $l = 1, ..., m$.

Case H(i)

The parent lives in the city or region n' that is closer to the city than $\bar{n}$. That is, $n' = 0, ..., \bar{n}$. Let $G^{\#}_{n'}$ denote $G^o_{t+1} + \bar{G}$ that satisfies:

$$\begin{aligned} & U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^{\#}_{n'}, \bar{G}) \\ = MAX\Big[\; & U(e, n', \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', G^{\#}_{n'}, \bar{G}), \\ & \quad ... \\ & U(e, \bar{n}, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', G^{\#}_{n'}, \bar{G}) \;\Big]. \end{aligned} \tag{34}$$

- If $G^{\#}_{n'}$ exists, it is larger than $\bar{G}$. More specifically, $G^{\#}_{\bar{n}} \geq \bar{G}$ and $G^{\#}_{n'} > \bar{G}$ for $n' = 0, ..., \bar{n} - 1$.
- If the parent's provision of family public goods including the one through the government, $G^o_{t+1} + \bar{G}$, is large enough to satisfy $G^{\#}_{n'} < G^o_{t+1} + \bar{G}$, then an individual does not live in the city and behaves as (3) in Case L(i), where L is replaced by H.

- Suppose the case where $G^o_{t+1} + \bar{G}$ is smaller than $G^{\#}_{n'}$, or $G^{\#}_{n'}$ does not exist.
  - if $\bar{G} \leq (G^o_{t+1} + \bar{G})\phi^{n'}$, then

$$\begin{cases} edu(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = e, \\ \psi(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = 0, \\ Q\left(\bar{G}, \frac{\bar{G}_l}{\phi^l} \mid H, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n' \to 0 \to l}, \end{cases} \tag{6}$$

    where $l = 1, ..., r$. If $G^o_{t+1} + \bar{G}$ is the minimum in the range, that is, $\bar{G} = (G^o_{t+1} + \bar{G})\phi^{n'}$, then for $l = 1, ..., r$, $0 \leq Q_{n' \to 0 \to l} \leq 1$ and $\sum_{l=1}^{r} Q_{n' \to 0 \to l} = 1$. If $G^o_{t+1} + \bar{G}$ is larger,then $l = 1$ and $Q_{n' \to 0 \to 1} = 1$.
  - If $\bar{G}\phi^n \leq (G^o_{t+1} + \bar{G})\phi^{n'} < \bar{G}\phi^{n-1}$ where $n \in \Omega_l$ ($n = 1, ..., \bar{n}$ and $l = r, ..., m$), then

$$\begin{cases} edu(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = e, \\ \psi(H, n', G^o_{t+1} + \bar{G}, \bar{G}) = 0, \\ Q\left(\bar{G}, \frac{\bar{G}_l}{\phi^l} \mid H, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n' \to 0 \to l}, \\ Q\left(\bar{G}, \frac{\bar{G}_{l+1}}{\phi^{l+1}} \mid H, n', G^o_{t+1} + \bar{G}, \bar{G}\right) = Q_{n' \to 0 \to l+1}. \end{cases} \tag{7}$$

    If $G^o_{t+1} + \bar{G}$ is the minimum in the range, that is, $\bar{G}\phi^n = (G^o_{t+1} + \bar{G})\phi^{n'}$ and $n \in \Omega_l \cap \Omega_{l+1}$, where $l = 1, ..., m-1$, then $Q_{n' \to 0 \to l} \in [0, 1]$, $Q_{n' \to 0 \to l+1} \in [0, 1]$ and $Q_{n' \to 0 \to l} + Q_{n' \to 0 \to l+1} = 1$. Otherwise, $Q_{n' \to 0 \to l} = 1$ and $Q_{n' \to 0 \to l+1} = 0$.
  - If $\bar{G} = 0$, (6) holds for $l = 1, ..., m$. If $G^0_{t+1} = 0$, then $Q_{n' \to 0 \to l}$ is independent from n' and is equal to Q_l, the solution of the simultaneous equation system in Appendix A. And $\sum_{l=1}^{m} Q_{n' \to 0 \to l} = 1$.

Case H(ii)

The parent lives in region n', which is farther away from the city than $\bar{n} + 1$. That is, $n' = \bar{n} + 1, ..., N$. Let $G^{\#}_{n'}$ denote $G^o_{t+1} + \bar{G}$ that satisfies the following:

$$\begin{aligned} & U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^{\#}_{n'}, \bar{G}) \\ = MAX\big[& \; U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^{\#}_{n'}, \bar{G}), \\ & \; ... \\ & \; U(e, n', \bar{G}, \bar{G}; L, n', G^{\#}_{n'}, \bar{G}) \;\; \big]. \end{aligned} \tag{35}$$

Let $G^{\flat}_{n'}$ denote $G^o_{t+1} + \bar{G}$ that satisfies the following with $G^o_{t+1} \geq 0$:

$$U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^{\flat}_{n'}, \bar{G}) = U(0, n', \bar{G}, \bar{G}; L, n', G^{\flat}_{n'}, \bar{G}). \tag{36}$$

The larger n' is, the smaller $G^{\flat}_{n'}$ is.

- Suppose the case that $G^*_{n'} < G^\flat_{n'}$.
  - If the parent's provision of family public goods, including the one through the government, $G^o_{t+1} + \bar{G}$, satisfies $\bar{G} \leq G^o_{t+1} + \bar{G} \leq G^\#_{n'}$, then an individual lives in the city and behaves as (6) or (7) in Case H(i) indicates.
  - If $G^\#_{n'} \leq G^o_{t+1} + \bar{G}$, then an individual does not live in the city but behaves as (32) in Case L(ii), where L is replaced by H.
- Suppose the case that $G^\flat_{n'} \leq G^*_{n'}$.
  - If $\bar{G} \leq G^o_{t+1} + \bar{G} < G^\flat_{n'}$, then an individual lives in the city and behaves as (6) or (7) in Case H(i).
  - If $G^\flat_{n'} \leq G^o_{t+1} + \bar{G} < G^*_{n'}$, then an individual does not live in the city but behaves as (4) in Case L(ii), where L is replaced by H.
  - If $G^*_{n'} \leq G^o_{t+1} + \bar{G}$, then an individual does not live in the city but behaves as (32) in Case L(ii), where L is replaced by H.
- Suppose the case that $G^*_{n'}$ does not exist.
  - If $\bar{G} \leq G^o_{t+1} + \bar{G} < G^\flat_{n'}$, then she lives in the city and behaves as (6) or (7) in Case H(i).
  - If $G^\flat_{n'} \leq G^o_{t+1} + \bar{G}$, then an individual behaves as (4) in Case L(ii), where L is replaced by H.

2 Simultaneous Equations System

First, I formally set up the conditional expected lifetime utility function; that is, the expected lifetime utility conditional on the fixed $(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}, G^{O,L}_{t+2})$, as well as her parent's choices and the policy, $(x_{t+1}, n_t, G^o_{t+1}, \bar{G})$. Assume that the child determines them according to the functions constructed in Section 4.1 and Appendix A.1.

Let $U^{O,L}(S_{t+1}, n_{t+1}, G^{O,L}_{t+2} + \bar{G}, \bar{G})$ denote the expected utility when an individual is old and her child becomes an unskilled worker. This is the function of the savings she decided on when she was young, and n_{t+1}, $G^{O,L}_{t+2}$ and $\bar{G}$ are given. Similarly let $U^{O,H}(S_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, \bar{G})$ denote the expected utility in the case where her child becomes a skilled worker. That is,

$$\begin{aligned} U^{O,L}(S_{t+1}, n_{t+1}, G^{O,L}_{t+2} + \bar{G}, \bar{G}) =& u((1+r)S_{t+1} - G^{O,L}_{t+2}) \\ &+ \sum_{\kappa}\sum_{\kappa'} Q(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid L, n_{t+1}, G^{O,L}_{t+2} + \bar{G}, \bar{G}) \\ &\times v\Big(G^{O,L}_{t+2} + \bar{G} + \delta\big(n_{t+1} - \psi(L, n_{t+1}, G^{O,L}_{t+2} + \bar{G}, \bar{G})\big) \\ &\quad \times G_Y\big(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid L, n_{t+1}, G^{O,L}_{t+2} + \bar{G}, \bar{G}\big)\Big), \end{aligned}$$

and

$$\begin{aligned} U^{O,H}(S_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, \bar{G}) =& u((1+r)S_{t+1} - G^{O,H}_{t+2}) \\ &+ \sum_{\kappa}\sum_{\kappa'} Q(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid H, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, \bar{G}) \\ &\times v\Big(G^{O,H}_{t+2} + \bar{G} + \delta\big(n_{t+1} - \psi(H, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, \bar{G})\big) \\ &\quad \times G_Y\big(G^{O,H,\kappa}_{t+3} + \bar{G}, G^{O,L,\kappa'}_{t+3} + \bar{G} \mid H, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, \bar{G}\big)\Big), \end{aligned}$$

where $\psi(x) = \phi^{|x|}$ and $0 < \phi \leq 1$. Let $U^Y(S_{t+1}, edu_{t+1}, n_{t+1}; x_{t+1}, n_t, G^O_{t+1} + \bar{G}, \bar{G})$ denote the maximized utility when young, $u(C^y_{t+1}) + v\left(G^y_{t+1} + \delta(n_{t+1} - n_t)(G^o_{t+1} + \bar{G})\right)$ subject to $C^y_{t+1} + G^y_{t+1} + S_{t+1} = w(x_{t+1}, n_{t+1}) - edu_{t+1} - (1/(1+r))\bar{G}$. In these equations, C^y_{t+1} and G^y_{t+1} are chosen to maximize the utility when young, but n_{t+1}, n_t, G^o_{t+1}, and S_{t+1} are fixed. By using them, the maximum conditional expected lifetime utility is expressed as follows:

$$
\begin{aligned}
&U(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G}; x_{t+1}, n_t, G^{o}_{t+1} + \bar{G}, \bar{G}) = \\
\max_{S_{t+1}} \; &\tilde{U}(S_{t+1}, edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G}; x_{t+1}, n_t, G^{o}_{t+1} + \bar{G}, \bar{G}) \\
=&U^{Y}(S_{t+1}, edu_{t+1}, n_{t+1}; x_{t+1}, n_t, G^{o}_{t+1} + \bar{G}, \bar{G}) \\
&+ \left(\frac{1}{1+\rho}\right) \Bigg[\lambda(edu_{t+1}) \times U^{O,H}(S_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, \bar{G}) \\
&\qquad + (1 - \lambda(edu_{t+1})) \times U^{O,L}(S_{t+1}, n_{t+1}, G^{O,L}_{t+2} + \bar{G}, \bar{G}) \Big\} \Bigg]
\end{aligned}
$$

I employ this function to construct the simultaneous equations system by taking the following steps.

2.1 Equations Related to a Skilled Worker's Choices

Condition for a skilled worker to live in the city

When the parent lives in region n', where $n' \in \Omega_l$, $l = 1, ..., m$, and $n' = 0, ..., \bar{n}$, and does not provide family public goods privately, the expected lifetime utility when she lives in the city and plans to provide $\bar{G}_l/\phi^l$ to place her child in region l in the case where her child becomes an unskilled worker is higher than that when she lives in region $\bar{n}$, earns a wage of w_L and plans not to provide family public goods privately.

$$U(e, \bar{n}, \bar{G}, \bar{G}; H, \bar{n}, \bar{G}, \bar{G}) \leq U(e, 0, \bar{G}, \bar{G}_l/\phi^l; H, \bar{n}, \bar{G}, \bar{G}). \tag{8}$$

A wage of high skilled workers satisfies:

$$w_H = w(P_0), \tag{22}$$

and it exceeds the substantial income of the richest unskilled worker:

$$w_L + \bar{G}_1 < w_H. \tag{37}$$

Equal utility condition

For a skilled worker whose parent lives in region $n' \in \Omega_l \cap \Omega_{l+1}$ where $l = 1, ..., m-1$, placing her child in region l and placing her child in region $l+1$ in the case where her child becomes an unskilled worker are indifferent:

$$U(e, 0, \bar{G}, \bar{G}_l/\phi^l; H, n', \bar{G}, \bar{G}) = U(e, 0, \bar{G}, \bar{G}_{l+1}/\phi^{l+1}; H, n', \bar{G}, \bar{G}). \tag{9}$$

The number of the equations is $m-1$. A skilled worker wil not place her child in the city when she becomes an unskilled worker as $\bar{G}_0$ required to do so is extremely large. That is,

$$U(e, 0, \bar{G}, \bar{G}_0; H, 0, \bar{G}, \bar{G}) < U(e, 0, \bar{G}, \bar{G}_1/\phi; H, 0, \bar{G}, \bar{G}). \tag{10}$$

Balanced population inflow and outflow condition when $\bar{G} > 0$

Inflow to and outflow from region $n = 1, ..., r-1$ are balanced if

$$\lambda P_n + (1-\lambda)(Q_{n\to n+l}P_n + Q_{n\to n+l+1}P_n) = (1-\lambda)\lambda Q_{0\to 0\to n}\frac{1}{2}P_0. \tag{38}$$

For region r, it is [24]

$$\lambda P_r + (1-\lambda)(Q_{r\to r+l}P_r + Q_{r\to r+l+1}P_r) = (1-\lambda)\lambda\left(Q_{0\to 0\to r}\frac{1}{2}P_0 + Q_{1\to 0\to r}P_1\right). \tag{39}$$

For region $n = r+1, ..., m-1$, it is

$$\lambda P_n + (1-\lambda)(Q_{n\to n+l}P_n + Q_{n\to n+l+1}P_n) = (1-\lambda)\lambda\sum_{n'\in\Omega_n} Q_{n'\to 0\to n}P_{n'}, \tag{40}$$

and for region m, it is [25]

$$\begin{aligned}&\lambda P_m + (1-\lambda)(Q_{m\to m+l}P_m + Q_{m\to m+l+1}P_m)\\ =&(1-\lambda)\left(\lambda\sum_{n'\in\Omega_m} Q_{n'\to 0\to m}P_{n'} + Q_{1\to m}P_1 + Q_{2\to m}P_2\right).\end{aligned} \tag{41}$$

If $n = 1$ is the maximum in Θ_{m-1}, then $Q_{1\to m} = 1$ and $0 \le Q_{2\to m} \le 1$. Otherwise, $0 \le Q_{1\to m} \le 1$ and $Q_{2\to m} = 0$. Moreover, in these equations,

$$\sum_{n=1}^{r} Q_{0\to 0\to n} = 1. \tag{42}$$

For $n' = 1, ..., \bar{n}$, if n' is only in Ω_n where $n = r, ..., m$, then

$$Q_{n'\to 0\to n} = 1, \tag{43}$$

[24] In the city, population λP_0 has the parent in the city. They give a higher education for the child, but, among them, $(1-\lambda)\lambda P_0$ become an unskilled worker, and $Q_{0\to 0\to n}\frac{1}{2}(1-\lambda)\lambda P_0$ will be placed in region n.

[25] In the city, population $\lambda(P_{n'} + P_{-n'})$ has the parent in region n' or $-n'$. They give a higher education for the child, but, among them, $(1-\lambda)\lambda(P_{n'} + P_{-n'})$ become an unskilled worker, and $Q_{n'\to 0\to n}\frac{1}{2}(1-\lambda)\lambda(P_{n'} + P_{-n'})$ will be placed in region n if $n' \in \Omega_n$. However, given the symmetry between the east and west half of the economy, $P_{n'} = P_{-n'}$. Therefore, $Q_{n'\to 0\to n}\frac{1}{2}(1-\lambda)\lambda(P_{n'} + P_{-n'}) = Q_{n'\to 0\to n}(1-\lambda)\lambda P_{n'}$.

and if $n' \in \Omega_n \cap \Omega_{n+1}$ where $n = r, ..., m-1$, then

$$Q_{n' \to 0 \to n} + Q_{n' \to 0 \to n+1} = 1. \tag{44}$$

The total number of equations (38), (39), (40) and (41) is m. The numbers of equations (42), (43), and (44) are, respectively 1, $\bar{n} - (m - r)$ and $m - r$.

Balanced population inflow and outflow condition when $\bar{G} = 0$

Inflow to and outflow from region $n = 1, ..., m-1$ are balanced if

$$\lambda P_n + (1-\lambda)(Q_{n \to n+l} P_n + Q_{n \to n+l+1} P_n) = (1-\lambda) Q_n \frac{1}{2} P_0, \tag{45}$$

and for region m, it is

$$\begin{aligned} &\lambda P_m + (1-\lambda)(Q_{m \to m+l} P_m + Q_{m \to m+l+1} P_m) \\ =&(1-\lambda)\left(Q_m \frac{1}{2} P_0 + Q_{1 \to m} P_1\right). \end{aligned} \tag{46}$$

In these equations,

$$\sum_{l=1}^{m} Q_l = 1. \tag{47}$$

2.2 Equations Related to an Unskilled Worker's Choices

Equal utility condition

First, for an unskilled worker residing in region n and residing in region $n+1$ are indifferent when her parent provides family public goods in the amount of $\bar{G}_n/\phi^{n-n'}$:

$$U(e, n, \bar{G}, \bar{G}_{n+l+1}/\phi^{l+1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}) = U(e, n+1, \bar{G}, \bar{G}_{n+l+1}/\phi^{l}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}), \tag{11}$$

where $n = 1, ..., \bar{n}-1$, $n \geq n'$, and $n \in \Theta_l$ but not the maximum in Θ_l. Thus, $n+1$ is also in Θ_l where $l = 0, ..., m-1$. If n is the maximum in Θ_l and thus $n+1 \in \Theta_{l-1}$, the condition is:

$$U(e, n, \bar{G}, \bar{G}_{n+l}/\phi^{l}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}) = U(e, n+1, \bar{G}, \bar{G}_{n+l}/\phi^{l-1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}). \tag{12}$$

The total number of equations (11) is $\bar{n} - m$ ($\bar{n} - 1$ minus $m - 1$), and the total number of equations (12) is $m - 1$. The amount of family public goods provided by her parent with which for an unskilled worker residing in the city and residing in region one are indifferent, $\bar{G}_0$, satisfies (11) with $n' = 0$, $n = 0$ and l that satisfies $1 \in \Theta_l$. However, the population in the city is much larger, and thus $\bar{G}_0$ is so large that an individual in the city will not choose it to place her child there in the case where her child becomes an unskilled worker as in (10).

Second, placing her child in region $n+l$ and placing her child in region $n+l+1$ are indifferent when her parent provides family public goods in the amount of $\bar{G}_n/\phi^{n-n'}$:

$$U(e, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}) = U(e, n, \bar{G}, \bar{G}_{n+l+1}/\phi^{l+1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}), \tag{13}$$

where $n \geq n'$ and $n = 1, ..., \bar{n}$ except n that is the maximum in Θ_l where $l = 0, ..., m-1$. Thus, the total number of the equations is $\bar{n} - m$.

When an individual lives in $\bar{n}$ and wants to place her child in that region, the minimum G^o_{t+2} required to do is $\bar{G}$ (the private provision of family public goods is zero):

$$\bar{G}_{\bar{n}} = \bar{G}. \tag{14}$$

Balanced population inflow and outflow condition

Consider the condition where inflow to and outflow from region $n = m+1, ..., \bar{n}$ are balanced. If $n - l' \in \Theta_{l'}$ and $n - (l'+1) \in \Theta_{l'}$, where $l' = 0, ..., m-1$,

$$P_n = (1-\lambda)(Q_{n-(l'+1)\to n} P_{n-(l'+1)} + Q_{n-l'\to n} P_{n-l'}). \tag{48}$$

However, if $n - l'$ is the maximum in Θ'_l, and, thus, $n - l' + 1$ is in $\Theta_{l'-1}$,

$$P_n = (1-\lambda)(Q_{n-(l'+1)\to n} P_{n-(l'+1)} + Q_{n-l'\to n} P_{n-l'} + Q_{n-(l'-1)\to n} P_{n-(l'-1)}). \tag{49}$$

The total number of equation (48) is $\bar{n} - 2m$, and that of equation (49) is m.

In these equations, if n is not the maximum in Θ_l, where $l = 0, ..., m-1$, then

$$Q_{n\to n+l} + Q_{n\to n+l+1} = 1, \tag{50}$$

and if n is the maximum in Θ_l where $l = 0, ..., m-1$, then

$$Q_{n\to n+l} = 1. \tag{51}$$

The total number of (50) is $\bar{n} - m$ and that of (51) is m.

Population in region $n = \bar{n}+1, ..., N$ is all the same as:

$$P_n = \frac{P - P_0/\lambda}{\bar{N} - (1+2\bar{n})} \quad \text{for} \quad n = \bar{n}+1, ..., N. \tag{21}$$

2.3 Boundary Conditions

Region $\bar{n}$ is the farthest from the city in regions where the educational investment is rewarded:

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}) \geq U(0, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}), \tag{15}$$

$$U(e, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}) \leq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}). \tag{16}$$

An individual whose parent lives in region $\bar{n}$ has no incentive to abandon education investment for her child and to move to a region farther from the city:

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}) \geq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}). \tag{17}$$

An individual whose parent lives in region $\bar{n}+1$ has no incentive to opt to invest a higher education or to move to a region closer to the city:

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}) \leq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}). \tag{18}$$

If the family public goods are not publicly provided ($\bar{G} = 0$), instead of (17) and (18), it holds that

$$U(e, \bar{n}, 0, 0; L, n', 0, 0) = U(0, \bar{n}+1, 0, 0; L, n', 0, 0). \tag{19}$$

2.4 Solution Procedure when $\bar{G} > 0$

Given $\bar{n}$, m, r, sets Θ_l ($l = 0, ..., m-1$), and sets Ω_l ($l = 1, ..., m$), I construct the simultaneous equations system, comprising $5\bar{n}+3$ equations and seven inequalities in the case where $\bar{G} > 0$.

That is, $\bar{n}-1$ of equal utility conditions for an individual to place the child in any of given regions ($m-1$ of (9) and $\bar{n}-m$ of (13)), $\bar{n}$ of equal utility conditions for an unskilled worker to reside in either region n or $n+1$ ($\bar{n}-m$ of (11), $m-1$ of (12) and (11) with $n' = 0$, $n = 0$, and l of $1 \in \Theta_l$), $\bar{n}+1$ of conditions of balanced population inflow and outflow (m of (38)-(41), $\bar{n}-2m$ of (48), m of (49), and (21)), $2\bar{n}+1$ equations about $Q_{n' \to 0 \to n}$ and $Q_{n' \to n}$ ((42), $\bar{n}-(m-r)$ of (43), $m-r$ and (44), $\bar{n}-m$ of (50) and m of (51)), and (14). Further, there are an equation, (22), and seven inequalities ((8), (10), (15), (16), (17), (18) and (37)). Assume that the left-hand side of (18) is bigger than the right-hand side by $\epsilon \geq 0$. That is,

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}) + \epsilon = U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}+1, \bar{G}, \bar{G}). \tag{52}$$

Solving the simultaneous equations system that comprises $5\bar{n}+4$ equations ($5\bar{n}+3$ equations listed above and (52)) and six inequalities ((8), (10), (15), (16), (17) and (37)) yields the following $5\bar{n}+4$ values: $\bar{G}_n$ ($n = 0, ..., \bar{n}$), P_n ($n = 0, ..., \bar{n}$), $2\bar{n}-m$ of $Q_{n' \to n}$ ($\bar{n}$ of $Q_{n \to n+l}$ and $\bar{n}-m$ of $Q_{n \to n+l+1}$), $\bar{n}+m$ of $Q_{n' \to 0 \to n}$ (r of $Q_{0 \to 0 \to n}$, $\bar{n}$ of $Q_{n' \to 0 \to n}$ and

$m-r$ of $Q_{n'\to 0\to n+1}$), w_H, and P_n for $n = \bar{n}+1, ..., N$. By changing ϵ in (52), we obtain the continuous sets of $5\bar{n}+4$ values, though the range will be small.

However, $\bar{n}$ m, r, sets Θ_l ($l = 0, ..., m-1$) and sets Ω_l ($l = 1, ..., m$) are given in the simultaneous equations system. The solution does not necessarily satisfy $Q_{n'\to 0\to n} \in [0, 1]$ and $Q_{n\to l} \in [0, 1]$. Hence, find the combination of $\bar{n}$ m, r, sets Θ_l ($l = 0, ..., m-1$) and sets Ω_l ($l = 0, ..., m$ or $l = 1, ..., m$), under which the solutions of the simultaneous equations satisfy inequalities $Q_{n'\to 0\to n} \in [0, 1]$ and $Q_{n\to l} \in [0, 1]$ as well as (8), (10), (15), (16), (17) and (37).

2.5 Solution Procedure when $\bar{G} = 0$

Given $\bar{n}$, m, r, and sets Θ_l ($l = 0, ..., m-1$), I construct the simultaneous equations system, comprising $4\bar{n}+4$ equations and five inequalities in the case where $\bar{G} = 0$.

That is, $\bar{n}-1$ of equal utility conditions for an individual to place the child in any of given regions ($m-1$ of (9) and $\bar{n}-m$ of (13)), $\bar{n}$ of equal utility conditions for an unskilled worker to reside in either region n or $n+1$ ($\bar{n}-m$ of (11), $m-1$ of (12) and (11) with $n' = 0$, $n = 0$, and l of $1 \in \Theta_l$), $\bar{n}+1$ of conditions of balanced population inflow and outflow (m of (45) and (46), $\bar{n}-2m$ of (48), m of (49), and (21)), $\bar{n}+1$ equations about Q_n and $Q_{n'\to n}$ ((47), $\bar{n}-m$ of (50) and m of (51)), and (14). Further, there are two equations, ((22) and (19)), and five inequalities ((8), (10), (15), (16) and (37)).

Solving the simultaneous equations system that comprises $4\bar{n}+4$ equations and five inequalities yields the following $4\bar{n}+4$ values: $\bar{G}_n$ ($n = 0, ..., \bar{n}$), P_n ($n = 0, ..., \bar{n}$), $2\bar{n}-m$ of $Q_{n'\to n}$ ($\bar{n}$ of $Q_{n\to n+l}$ and $\bar{n}-m$ of $Q_{n\to n+l+1}$), m of Q_n , w_H, and P_n for $n = \bar{n}+1, ..., N$.

However, $\bar{n}$ m, r, and sets Θ_l ($l = 0, ..., m-1$) are given in the simultaneous equations system. The solution does not necessarily satisfy $Q_n \in [0, 1]$ and $Q_{n\to l} \in [0, 1]$. Hence, find the combination of $\bar{n}$ m, r, and sets Θ_l ($l = 0, ..., m-1$) under which the solutions of the simultaneous equations satisfy inequalities $Q_n \in [0, 1]$ and $Q_{n\to l} \in [0, 1]$ as well as (8), (10), (15), (16) and (37).

Appendix B: Proof of Proposition

1 Proof of Proposition: An Unskilled Worker's Behavior

Prior to the proof of Proposition, I show some Lemmas, and prove them with Lemmas 1, 2 and 3 in the maintext.

In order to consider an unskilled worker's behavior when the parent lives in $n' \in \{0, ..., \bar{n}\}$, I compare the following two:

- The optimal behavior and (the minimum level of) the maximum expected lifetime utility when where to live is limited in $\{0, ..., \bar{n}\}$.
- The optimal behavior and the maximum expected lifetime utility when where to live is limited in $\{\bar{n}+1, ..., N\}$.

The following two Lemmas consider the cases.

Lemma 4
Suppose that the parent lives in region $n' \in \{0, ..., \bar{n}\}$ and an unskilled individual can freely choose whether or not she gives a higher education for her child but her choice of where to live is limited in $n \in \{0, ..., \bar{n}\}$.

(i) Suppose that G^o_{t+1} and n' satisfy $0 < \bar{G} \leq (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$. Then (27) necessarily holds, and an unskilled individual prefers $edu = e$ to $edu = 0$, lives in a region in $\{n', ..., \bar{n}\}$, and plans $G^{O,L}_{t+2} \geq 0$ and $G^{O,H}_{t+2} = 0$. Her maximum expected lifetime utility is equal to or at least $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

(ii) Suppose that $0 < (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} < \bar{G}$. In this case (27) does not necessarily holds. If it holds, an unskilled individual prefers $edu = e$ to $edu = 0$, lives in region $\bar{n}$ and plans $G^{O,L}_{t+2} = 0$ and $G^{O,H}_{t+2} = 0$, and her maximum expected lifetime utility is equal to $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. Otherwise, she does not give a higher education for her child ($edu = 0$), lives in region $\bar{n}$ and plans $G^{O,L}_{t+2} = 0$, and her maximum expected lifetime utility is equal to $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

(iii) Suppose that $G^o_{t+1} = 0$ and $\bar{G} = 0$. Then (27) necessarily holds, and an unskilled individual prefers $edu = e$ to $edu = 0$, lives in region $\bar{n}$, and plans $G^{O,L}_{t+2} = 0$ and $G^{O,H}_{t+2} = 0$. Her maximum expected lifetime utility is equal to $U(e, \bar{n}, 0, 0; L, n', 0, 0)$.

Lemma 5
Suppose that the parent lives in region $n' \in \{0, ..., \bar{n}\}$ and an unskilled individual can freely choose whether or not she gives a higher education for her child but her choice of where to live is limited in $n \in \{\bar{n}+1, ..., N\}$.

(i) When $\bar{G} > 0$ or $G^o_{t+1} > 0$, an unskilled individual's maximum expected lifetime utility is $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ or $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

(ii) When $\bar{G} = G^o_{t+1} = 0$, an unskilled individual's maximum expected lifetime utility is $U(e, \bar{n}+1, 0, 0; L, n', 0, 0)$ or $U(0, n, 0, 0; L, n', 0, 0)$ where $n \in \{\bar{n}+1, ..., N\}$.

1.1 Proof of Lemma 1

1.1.1 Lemma 1(i)

An individual's conditional expected lifetime utility $\tilde{U}(S_{t+1}, edu_{t+1}, n_{t+1}, G^{O,H}_{t+2} + \bar{G}, G^{O,L}_{t+2} + \bar{G}; x_{t+1}, n_t, G^o_{t+1} + \bar{G}, \bar{G})$ depends on the utility when she is young, that when she is old and her child becomes a skilled worker, and that when she is old and her child becomes an unskilled worker. Among them, the former two components are equal given the plan of savings S_{t+1}. Therefore, (23) holds if and only if the utility when she plans to provide $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l}/\phi^l$ in the case where her child becomes an unkilled worker is equal to that when she plans to provide $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l+1}/\phi^{l+1}$, given that she lives in n and plans savings in the amount of $S^n_{n+l,n+l+1}$. That is:

$$\begin{aligned}
&u\left((1+r)S^n_{n+l,n+l+1} - \left(\frac{\bar{G}_{n+l}}{\phi^l} - \bar{G}\right)\right) \\
&+ q_{n+l\to n+l+l'} \times v\left(\frac{\bar{G}_{n+l}}{\phi^l} + \phi^l G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'}}{\phi^{l'}} \;\middle|\; L, n, \frac{\bar{G}_{n+l}}{\phi^l}, \bar{G}\right)\right) \\
&+ (1 - q_{n+l\to n+l+l'}) \times v\left(\frac{\bar{G}_{n+l}}{\phi^l} + \phi^l G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'+1}}{\phi^{l'+1}} \;\middle|\; L, n, \frac{\bar{G}_{n+l}}{\phi^l}, \bar{G}\right)\right) \\
=&u\left((1+r)S^n_{n+l,n+l+1} - \left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} - \bar{G}\right)\right) \qquad (53) \\
&+ q_{n+l+1\to n+l+1+l''} \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^{l+1} G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''}}{\phi^{l''}} \;\middle|\; L, n, \frac{\bar{G}_{n+l+1}}{\phi^{l+1}}, \bar{G}\right)\right) \\
&+ (1 - q_{n+l+1\to n+l+1+l''}) \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^{l+1} G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''+1}}{\phi^{l''+1}} \;\middle|\; L, n, \frac{\bar{G}_{n+l+1}}{\phi^{l+1}}, \bar{G}\right)\right),
\end{aligned}$$

where $n + l \in \Theta_{l'}$, $n + l + 1 \in \Theta_{l''}$ and $l' \geq l''$. Similarly, (24) holds if the utility when she plans to provide $G^o_{t+2} + \bar{G} = \bar{G}_{n+l}/\phi^{l-1}$ in the case where her child becomes an unkilled worker is lower than that when she plans to provide $G^o_{t+2} + \bar{G} = \bar{G}_{n+l+1}/\phi^l$, given that she lives in $n+1$ and plans savings in the amount of $S^n_{n+l,n+l+1}$. That is:

$$\begin{aligned}
& u\left((1+r)S^n_{n+l,n+l+1} - \left(\frac{\bar{G}_{n+l}}{\phi^{l-1}} - \bar{G}\right)\right) \\
& + q_{n+l\to n+l+l'} \times v\left(\frac{\bar{G}_{n+l}}{\phi^{l-1}} + \phi^{l-1} G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'}}{\phi^{l'}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l}}{\phi^{l-1}}, \bar{G}\right)\right) \\
& + (1-q_{n+l\to n+l+l'}) \times v\left(\frac{\bar{G}_{n+l}}{\phi^{l-1}} + \phi^{l-1} G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'+1}}{\phi^{l'+1}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l}}{\phi^{l-1}}, \bar{G}\right)\right) \\
< & u\left((1+r)S^n_{n+l,n+l+1} - \left(\frac{\bar{G}_{n+l+1}}{\phi^{l}} - \bar{G}\right)\right) \\
& + q_{n+l+1\to n+l+1+l''} \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l}} + \phi^{l} G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''}}{\phi^{l''}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l+1}}{\phi^{l}}, \bar{G}\right)\right) \\
& + (1-q_{n+l+1\to n+l+1+l''}) \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l}} + \phi^{l} G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''+1}}{\phi^{l''+1}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l+1}}{\phi^{l}}, \bar{G}\right)\right).
\end{aligned} \tag{54}$$

Therefore, in order to prove Lemma 1(i), I show that if (53) holds, then (54) holds.

For the sake of notational simplicity, I let G^S_{n+l}, G^B_{n+l}, G^S_{n+l+1}, and G^B_{n+l+1} represent:

$$\begin{aligned}
G^S_{n+l} &= G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'}}{\phi^{l'}} \;\middle|\; L, n, \frac{\bar{G}_{n+l}}{\phi^{l}}, \bar{G}\right) = G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'}}{\phi^{l'}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l}}{\phi^{l-1}}, \bar{G}\right), \\
G^B_{n+l} &= G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'+1}}{\phi^{l'+1}} \;\middle|\; L, n, \frac{\bar{G}_{n+l}}{\phi^{l}}, \bar{G}\right) = G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+l'+1}}{\phi^{l'+1}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l}}{\phi^{l-1}}, \bar{G}\right), \\
G^S_{n+l+1} &= G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''}}{\phi^{l''}} \;\middle|\; L, n, \frac{\bar{G}_{n+l+1}}{\phi^{l+1}}, \bar{G}\right) = G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''}}{\phi^{l''}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l+1}}{\phi^{l}}, \bar{G}\right), \\
G^B_{n+l+1} &= G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''+1}}{\phi^{l''+1}} \;\middle|\; L, n, \frac{\bar{G}_{n+l+1}}{\phi^{l+1}}, \bar{G}\right) = G_Y\left(\bar{G}, \frac{\bar{G}_{n+l+1+l''+1}}{\phi^{l''+1}} \;\middle|\; L, n+1, \frac{\bar{G}_{n+l+1}}{\phi^{l}}, \bar{G}\right),
\end{aligned}$$

and let q_{n+l} and q_{n+l+1} denote $q_{n+l\to n+l+l'}$ and $q_{n+l+1\to n+l+1+l''}$, respectively. I introduce $\bar{q}_{n+l}$ and $\bar{q}_{n+l+1}$ that satisfy the followings:

$$\begin{aligned}
& q_{n+l} \times v\left(\frac{\bar{G}_{n+l}}{\phi^{l}} + \phi^{l} G^S_{n+l}\right) + (1-q_{n+l}) \times v\left(\frac{\bar{G}_{n+l}}{\phi^{l}} + \phi^{l} G^B_{n+l}\right) \\
= & v\left(\frac{\bar{G}_{n+l}}{\phi^{l}} + \phi^{l}\left(\bar{q}_{n+l} \times G^S_{n+l} + (1-\bar{q}_{n+l}) \times G^B_{n+l}\right)\right),
\end{aligned} \tag{55}$$

and

$$\begin{aligned}
& q_{n+l+1} \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^{l+1} G^S_{n+l+1}\right) + (1-q_{n+l+1}) \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^{l+1} G^B_{n+l+1}\right) \\
= & v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^{l+1}\left(\bar{q}_{n+l+1} \times G^S_{n+l+1} + (1-\bar{q}_{n+l+1}) \times G^B_{n+l+1}\right)\right).
\end{aligned} \tag{56}$$

These are the sum of the second and third terms in the left-hand side, and that in the right-hand side in (53), respectively. Similarly, I introduce $\hat{q}_{n+l}$ and $\hat{q}_{n+l+1}$ that satisfy the followings:

$$\begin{aligned} &q_{n+l} \times v\left(\frac{\bar{G}_{n+l}}{\phi^{l-1}} + \phi^{l-1}G^S_{n+l}\right) + (1-q_{n+l}) \times v\left(\frac{\bar{G}_{n+l}}{\phi^{l-1}} + \phi^{l-1}G^B_{n+l}\right) \\ =&v\left(\frac{\bar{G}_{n+l}}{\phi^{l-1}} + \phi^{l-1}\left(\hat{q}_{n+l} \times G^S_{n+l} + (1-\hat{q}_{n+l}) \times G^B_{n+l}\right)\right), \end{aligned} \tag{57}$$

and

$$\begin{aligned} &q_{n+l+1} \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l}} + \phi^{l}G^S_{n+l+1}\right) + (1-q_{n+l+1}) \times v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l}} + \phi^{l}G^B_{n+l+1}\right) \\ =&v\left(\frac{\bar{G}_{n+l+1}}{\phi^{l}} + \phi^{l}\left(\hat{q}_{n+l+1} \times G^S_{n+l+1} + (1-\hat{q}_{n+l+1}) \times G^B_{n+l+1}\right)\right). \end{aligned} \tag{58}$$

These are the sum of the second and third terms in the left-hand side, and that in the right-hand side in (54), respectively. Due to the convexity of function v, $\bar{q}_{n+l} \geq q_{n+l}$, $\bar{q}_{n+l+1} \geq q_{n+l+1}$, $\hat{q}_{n+l} \geq q_{n+l}$ and $\hat{q}_{n+l+1} \geq q_{n+l+1}$.

In Figure 8, points A and B are coordinates $\big((1+r)S^n_{n+l,n+l+1} - \big(\frac{\bar{G}_{n+l}}{\phi^l} - \bar{G}\big), \frac{\bar{G}_{n+l}}{\phi^l} + \phi^l\big(\bar{q}_{n+l}G^S_{n+l} + (1-\bar{q}_{n+l})G^B_{n+l}\big)\big)$ and $\big((1+r)S^n_{n+l,n+l+1} - \big(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} - \bar{G}\big), \frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^{l+1}\big(\bar{q}_{n+l+1}G^S_{n+l+1} + (1-\bar{q}_{n+l+1})G^B_{n+l+1}\big)\big)$, respectively. From (53), (55) and (56), the points A and B are on the same indifference curve between the consumption of private goods and the total provision of family public goods. Points A' and B' are the coordinates $\big((1+r)S^n_{n+l,n+l+1} - \big(\frac{\bar{G}_{n+l}}{\phi^{l-1}} - \bar{G}\big), \frac{\bar{G}_{n+l}}{\phi^{l-1}} + \phi^{l-1}\big(\hat{q}_{n+l}G^S_{n+l} + (1-\hat{q}_{n+l})G^B_{n+l}\big)\big)$ and $\big((1+r)S^n_{n+l,n+l+1} - \big(\frac{\bar{G}_{n+l+1}}{\phi^{l}} - \bar{G}\big), \frac{\bar{G}_{n+l+1}}{\phi^{l}} + \phi^{l}\big(\hat{q}_{n+l+1}G^S_{n+l+1} + (1-\hat{q}_{n+l+1})G^B_{n+l+1}\big)\big)$, respectively.

I show that the point B' is on the indifference curve that corresponds higher utility than that on which the point A' is. The slope of the line segment AB is:

$$\frac{\frac{\bar{G}_{n+l}}{\phi^l} - \frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^l(\bar{q}_{n+l}G^S_{n+l} + (1-\bar{q}_{n+l})G^B_{n+1}) - \phi^{l+1}(\bar{q}_{n+l+1}G^S_{n+l+1} + (1-\bar{q}_{n+l+1})G^B_{n+l+1})}{\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} - \frac{\bar{G}_{n+l}}{\phi^l}}, \tag{59}$$

and that of the line segment $A'B'$ is:

$$\frac{\frac{\bar{G}_{n+l}}{\phi^{l-1}} - \frac{\bar{G}_{n+l+1}}{\phi^{l}} + \phi^{l-1}(\hat{q}_{n+l}G^S_{n+l} + (1-\hat{q}_{n+l})G^B_{n+1}) - \phi^{l}(\hat{q}_{n+l+1}G^S_{n+l+1} + (1-\hat{q}_{n+l+1})G^B_{n+l+1})}{\frac{\bar{G}_{n+l+1}}{\phi^{l}} - \frac{\bar{G}_{n+l}}{\phi^{l-1}}}. \tag{60}$$

The line segments AB and $A'B'$ have the following three characteristics. First, both of them have a negative slope. In order to increase the total family public goods provision, an individual must contribute more by decreasing the consumption of private goods. Second, the line segment $A'B'$ locates upper than AB. In cases of B and B', an individual places her child in region $n+l+1$ when her child becomes an unskilled worker. In contrast, in cases A and A', she places her child in region $n+l$ by providing larger amount of family public goods. Also, in cases of A' and B', an individual lives in region $n+1$, whereas in cases of A and B she lives in region n. Therefore, in cases of A' and B', her child lives closer to the individual when her child becomes an unskilled worker, and, thus, the individual can enjoy higher value of family public goods provided by her child than in cases of A and B, respectively. Third, the slope of the line segment AB is steeper than $A'B'$. That is, (59) is smaller than (60) (the absolute value of (59) is larger than that of (60)). When an individual increases her provision of family public goods, the value of total family public goods provision also increases but less drastically. That is because her child will decrease her contribution, though her child lives closer to the region where an individual lives. However, when her child initially lives farther away from the region where an individual lives (the case of the line segment AB where an individual lives in region n), the distance mitigates this negative impact. The impact of an increase in the provision of family public goods by an individual on her child is smaller. In addition, the impact of a decrease in the provision of family public goods by her child on that individual is also smaller. Therefore, the total provision of family public goods increase more drastically when she increases her provision by decreasing her consumption of private goods in the case where she lives in n. From (59) and (60), the third characteristic holds if and only if:

$$\begin{aligned}0 &> ((\bar{q}_{n+l}G_{n+l,S} + (1-\bar{q}_{n+l})G_{n+1,B}) - \phi(\bar{q}_{n+l+1}G_{n+l+1,S} + (1-\bar{q}_{n+l+1})G_{n+l+1,B}))\\ &> \phi^{-2}((\hat{q}_{n+l}G_{n+l,S} + (1-\hat{q}_{n+l})G_{n+1,B}) - \phi(\hat{q}_{n+l+1}G_{n+l+1,S} + (1-\hat{q}_{n+l+1})G_{n+l+1,B})). \end{aligned} \tag{61}$$

Unless the gap between $\bar{q}_{n+l}$ and $\hat{q}_{n+l}$ and that between $\bar{q}_{n+l+1}$ and $\hat{q}_{n+l+1}$ are extremely large, it holds.

Consider the indifference curve on which point A' and point $\left(k\left((1+r)S^n_{n+l,n+l+1} - \left(\frac{\bar{G}_{n+l}}{\phi^l} - \bar{G}\right)\right), k\left(\frac{\bar{G}_{n+l}}{\phi^l} + \phi^l\left(\bar{q}_{n+l}G^S_{n+l} + (1-\bar{q}_{n+l})G^B_{n+l}\right)\right)\right)$ are. Let A'' denote the latter point. In A'', k is larger than unity. Because the utility is the homothetic function of the consumption of private goods and the provision of total family public goods, O, A, and A'' are on the same line. Similarly, let B'' denote the point $\left(k\left((1+r)S^n_{n+l,n+l+1} - \left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} - \bar{G}\right)\right), k\left(\frac{\bar{G}_{n+l+1}}{\phi^{l+1}} + \phi^{l+1}\left(\bar{q}_{n+l+1}G^S_{n+l+1} + (1-\bar{q}_{n+l+1})G^B_{n+l+1}\right)\right)\right)$. Points O, B, and B'' are on the same line, and line AB and line $A''B''$ are parallel. As the provision of family public goods by her child, G^y_{t+2}, is much larger than that by an individual, G^o_{t+2}, the line segments AA' and BB' are extremely steep. Thus, point A' locates north-west of point A'' on the same indifference curve, and point B' locates north-west of point B''. As we saw, line $A'B'$ is less steeper than line AB and $A''B''$. Therefore, point B' locates upper than the line segment $A''B''$. Thus, the utility on point B' is higher than the utility that corresponds the indifference curve on which points A', A'' and B'' are. Therefore, if (53) holds, (54) holds.

1.1.2 Lemma 1(ii)-(iv)

Consider the properties of conditional maximum expected lifetime utility where edu_{t+1}, n_{t+1}, $G^{O,L}_{t+1}$, $G^{O,H}_{t+1}$ are fixed, $U(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^{o}_{t+1}+\bar{G}^P, \bar{G}^P)$. Let $\bar{G}^P$ denote the government's provision of family public goods through the parent when an individual is young, $(1/(1+r))T$ denote the tax on an individual, and $\bar{G}$ denote the government provision of family public goods when she becomes old. In the model, $\bar{G}^P = T = \bar{G}$. A change in $\bar{G}$ has impacts on the expected lifetime utility through various channels. However, to focus on a specific channel, consider the case where one or two of them changes, with the rest being equal.

First, consider the maximization problem when young, given the savings S_{t+1}. That is,

$$\begin{aligned} & U^Y(S_{t+1}, edu_{t+1}, n_{t+1}, T; x_{t+1}, n_t, G^o_{t+1} + \bar{G}^P, \bar{G}^P) = \\ \max_{C^y_{t+1}, G^y_{t+1}} \quad & u(C^y_{t+1}) + v(G^y_{t+1} + \delta(G^o_{t+1} + \bar{G}^P)) \\ \text{s.t.} \quad & C^y_{t+1} + G^y_{t+1} = w(x_{t+1}, n_{t+1}) - edu_{t+1} - (1/(1+r))T - S_{t+1}. \end{aligned}$$

The Lagrangian of the maximization problem when young given the plan of savings S_{t+1} is,

$$\begin{aligned} L = & u(C^y_{t+1}) + v(G^y_{t+1} + \delta(G^o_{t+1} + \bar{G}^P)) \\ & + \lambda \left[w(x_{t+1}, n_{t+1}) - edu_{t+1} - (1/(1+r))T - S_{t+1} - C^y_{t+1} - G^y_{t+1} \right], \end{aligned}$$

where λ is Lagrange multiplier. The first order conditions of the optimiazation are:

$$u'(C^y_{t+1}) = \lambda, \tag{62}$$

$$v'(G^y_{t+1} + \delta(G^o_{t+1} + \bar{G}^P)) = \lambda, \tag{63}$$

and

$$w(x_{t+1}, n_{t+1}) - edu_{t+1} - (1/(1+r))T - S_{t+1} - C^y_{t+1} - G^y_{t+1} = 0. \tag{64}$$

The total derivative of $U^Y(S_{t+1}, edu_{t+1}, n_{t+1}, T; x_{t+1}, n_t, G^O_{t+1} + \bar{G}^P, \bar{G}^P)$ is:

$$\begin{aligned} & dU^Y(S_{t+1}, edu_{t+1}, n_{t+1}, T; x_{t+1}, n_t, G^O_{t+1} + \bar{G}^P, \bar{G}^P) \\ =& u'(C^y_{t+1})dC^y_{t+1} + v'(G^y_{t+1} + \delta(G^o_{t+1} + \bar{G}^P))(dG^y_{t+1} + \delta(dG^o_{t+1} + d\bar{G}^P)). \end{aligned}$$

From (62), (63), and (64), it follows that:

$$\begin{aligned} & dU^Y(S_{t+1}, edu_{t+1}, n_{t+1}, T; x_{t+1}, n_t, G^o_{t+1} + \bar{G}^P, \bar{G}^P) \\ =& \lambda \left[dw(x_{t+1}, n_{t+1}) - (1/(1+r))dT - dS_{t+1} + \delta(dG^o_{t+1} + d\bar{G}^P) \right]. \end{aligned} \tag{65}$$

Second, consider the optimal choice of S_{t+1} that maximizes the expected lifetime utility $U^Y(S_{t+1}; edu_{t+1}, n_{t+1}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P) + (\frac{1}{1+\rho})[(\mu(edu)u((1+r)S_{t+1} - G^{O,H}_{t+2}) + (1-\mu(edu))u((1+r)S_{t+1} - G^{O,L}_{t+2})]$, where $G^{O,H}_{t+2}$ and $G^{O,L}_{t+2}$ are fixed and thus the utility from family public goods when old is fixed. An individual chooses optimal S_{t+1} that satisfies the following:

$$\frac{\partial U_Y(S_{t+1}, edu_{t+1}, n_{t+1}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial S_{t+1}} + \left(\frac{1+r}{1+\rho}\right)\left[\mu(edu)\frac{\partial u((1+r)S_{t+1} - G^{O,H}_{t+2})}{\partial(1+r)S_{t+1}} + (1-\mu(edu))\frac{\partial u((1+r)S_{t+1} - G^{O,L}_{t+2})}{\partial(1+r)S_{t+1}}\right] = 0.$$

From (65), this is rewritten as:

$$-\lambda + \left(\frac{1+r}{1+\rho}\right)\left[\mu(edu)\frac{\partial u((1+r)S_{t+1} - G^{O,H}_{t+2})}{\partial(1+r)S_{t+1}} + (1-\mu(edu))\frac{\partial u((1+r)S_{t+1} - G^{O,L}_{t+2})}{\partial(1+r)S_{t+1}}\right] = 0. \tag{66}$$

The total derivative of $U(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)$ is:

$$\begin{aligned} & dU(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P) \\ =& \left[\frac{\partial U_Y}{\partial S_{t+1}} + \left(\frac{1+r}{1+\rho}\right)\left[\mu(edu)\frac{\partial u((1+r)S_{t+1} - G^{O,H}_{t+2})}{\partial(1+r)S_{t+1}} + (1-\mu(edu))\frac{\partial u(S_{t+1} - G^{O,L}_{t+2})}{\partial(1+r)S_{t+1}}\right]\right] dS_{t+1} \\ &+ \frac{\partial U_Y}{\partial w_{t+1}} dw(n_{t+1}, x_{t+1}) + \frac{\partial U_Y}{\partial T} dT + \frac{\partial U_Y}{\partial G^o_{t+1}} dG^o_{t+1} + \frac{\partial U_Y}{\partial \bar{G}^P} d\bar{G}^P. \end{aligned}$$

From (65) and (66), this is rewritten as:

$$\begin{aligned} & dU(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P) \\ =& \lambda\left[dw(x_{t+1}, n_{t+1}) - (1/(1+r))dT + \delta(dG^o_{t+1} + d\bar{G}^P)\right]. \end{aligned}$$

Therefore,

$$\frac{\partial U(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial w_{t+1}} = \lambda, \tag{67}$$

$$\frac{\partial U(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial G^o_{t+1}} = \delta\lambda. \tag{68}$$

Similarly, when the government's provision of family public goods through her parent $\bar{G}^P$ increases by dG and taxation on an inividual $(1/(1+r))T$ increases by $(1/1+r)dG$ but the government's provision of family public goods when she is old $\bar{G}$ does not change, then

$$\frac{\partial U(edu_{t+1}, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial G} = \left[\delta - \left(\frac{1}{1+r}\right)\right]\lambda. \tag{69}$$

From (67),

$$\frac{\partial^2 U}{\partial G^{O,L}_{t+2}\partial w_{t+1}} = \frac{\partial \lambda}{\partial G^{O,L}_{t+2}} = \frac{\partial \lambda}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial G^{O,L}_{t+2}}. \tag{70}$$

From (68),

$$\frac{\partial^2 U}{\partial G^{O,L}_{t+2}\partial G^{o}_{t+1}} = \delta\frac{\partial \lambda}{\partial G^{O,L}_{t+2}} = \delta\frac{\partial \lambda}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial G^{O,L}_{t+2}}, \tag{71}$$

and

$$\begin{aligned}
&\frac{\partial U(e, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial G^o_{t+1}}\\
-&\frac{\partial U(0, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial G^o_{t+1}}\\
=&\delta\left[\lambda(e+S(e)) - \lambda(0+S(0))\right].
\end{aligned} \tag{72}$$

In this equation λ is the solution of (62), (63) and (64), the conditions of maximization of U^Y. Thus, λ is the function of $w(x_{t+1}, n_{t+1}) - edu_{t+1} - (1/(1+r))T - S_{t+1} + \delta(G^o_{t+1} + \bar{G}^P)$. Therefore, if edu_{t+1} and S_{t+1} change keeping $edu_{t+1} + S_{t+1}$ constant, λ does not change. Thus, I describe λ as a function of $edu_{t+1} + S_{t+1}$. Also, S_{t+1} is the solution of (66). Thus, S_{t+1} is the function of $G^{O,H}_{t+2}$, $G^{O,L}_{t+2}$, and edu. Here, to focus on the impact of edu_{t+1} on S_{t+1}, I describe S_{t+1} as a function of edu_{t+1}, $S(edu_{t+1})$. And from (69),

$$\begin{aligned}
&\frac{\partial U(e, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial G}\\
-&\frac{\partial U(0, n_{t+1}, G^{O,H}_{t+2}+\bar{G}, G^{O,L}_{t+2}+\bar{G}, T; x_{t+1}, n_t, G^o_{t+1}+\bar{G}^P, \bar{G}^P)}{\partial G}\\
=&\left[\delta - \left(\frac{1}{1+r}\right)\right]\left[\lambda(e+S(e)) - \lambda(0+S(0))\right].
\end{aligned} \tag{73}$$

Consider $\partial S_{t+1}/\partial G^{O,L}_{t+2}$, $\partial\lambda/\partial S_{t+1}$ and $\lambda(e+S(e)) - \lambda(0+S(0))$ to investigate the signs of (71), (72), (70) and (73).

From (66), when $G^{O,L}_{t+2}$ differs, the optimal S_{t+1} changes satisfying the following:

$$\begin{aligned}
&\left[-\frac{\partial\lambda}{\partial S_{t+1}} + \left(\frac{(1+r)^2}{1+\rho}\right)\left[\mu(edu)\frac{\partial^2 u((1+r)S_{t+1} - G^{O,H}_{t+2})}{\partial((1+r)S_{t+1})^2} + (1-\mu(edu))\frac{\partial^2 u((1+r)S_{t+1} - G^{O,L}_{t+2})}{\partial((1+r)S_{t+1})^2}\right]\right]dS_{t+1}\\
&-\left(\frac{1+r}{1+\rho}\right)(1-\mu(edu))\frac{\partial^2 u((1+r)S_{t+1} - G^{O,L}_{t+2})}{\partial((1+r)S_{t+1})^2}dG^{O,L}_{t+2} = 0.
\end{aligned}$$

Rearranging it yields:

$$\frac{dS_{t+1}}{dG^{O,L}_{t+2}} = \frac{\left(\frac{1+r}{1+\rho}\right)(1-\mu(edu))\frac{\partial^2 u((1+r)S_{t+1} - G^{O,L}_{t+2})}{\partial((1+r)S_{t+1})^2}}{\Delta}, \tag{74}$$

where

$$\Delta = \left[-\frac{\partial \lambda}{\partial S_{t+1}} + \left(\frac{(1+r)^2}{1+\rho} \right) \left[\mu(edu) \frac{\partial^2 u((1+r)S_{t+1} - G^{O,H}_{t+2})}{\partial((1+r)S_{t+1})^2} + (1-\mu(edu)) \frac{\partial^2 u((1+r)S_{t+1} - G^{O,L}_{t+2})}{\partial((1+r)S_{t+1})^2} \right] \right].$$

From (62), (63) and (64),

$$\begin{pmatrix} u'' & 0 & -1 \\ 0 & v'' & -1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} dC^y_{t+1} \\ dG^y_{t+1} \\ d\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ -v''\delta(dG^o_{t+1} + d\bar{G}^P) \\ dw - dedu_{t+1} - (1/(1+r))dT - dS_{t+1} \end{pmatrix}. \tag{75}$$

Therefore,

$$\frac{\partial \lambda}{\partial S_{t+1}} = \frac{-u''v''}{u'' + v''} > 0. \tag{76}$$

From (76) and $u'' < 0$, the denominator Δ and numerator of (74) are negative. Thus, (74) is positive. From this fact and (76), (70) and (71) are positive. Thus, Lemma 1(ii) follows. From (66), when $G^{O,H}_{t+2} = G^{O,L}_{t+2} = G^o_{t+2}$, between the case where $edu = e$ and the case where $edu = 0$, the optimal S_{t+1} and λ differ as follows:

$$-\lambda(e + S(e)) + \left(\frac{1+r}{1+\rho} \right) \left[\frac{\partial u((1+r)S(e) - G^o_{t+2})}{\partial(1+r)S_{t+1}} \right] = 0, \tag{77}$$

and

$$-\lambda(0 + S(0)) + \left(\frac{1+r}{1+\rho} \right) \left[\frac{\partial u((1+r)S(0) - G^o_{t+2})}{\partial(1+r)S_{t+1}} \right] = 0. \tag{78}$$

From (75),

$$\frac{\partial \lambda}{\partial(edu_{t+1} + S_{t+1})} = \frac{-u''v''}{u'' + v''} > 0. \tag{79}$$

Therefore, $\lambda(0 + S(0)) < \lambda(e + S(0))$. From this fact and (78),

$$-\lambda(e + S(0)) + \left(\frac{1+r}{1+\rho} \right) \left[\frac{\partial u((1+r)S(0) - G^o_{t+2})}{\partial(1+r)S_{t+1}} \right] < 0. \tag{80}$$

For S that satisfies $S = S(0) - e$, it holds that

$$-\lambda(e + S) + \left(\frac{1+r}{1+\rho} \right) \left[\frac{\partial u((1+r)S - G^o_{t+2})}{\partial(1+r)S_{t+1}} \right] > 0. \tag{81}$$

That is because the first term is the same as that in (78). The second term is larger than that in (78) as $S < S(0)$. From (77), (80) and (81), $S(0) - e < S(e) < S(0)$, and, thus, $0 + S(0) < e + S(e)$. From this fact and (76),

$$\lambda(0 + S(0)) < \lambda(e + S(e)).$$

Therefore, (72) is positive and Lemma 1(iii) follows. If an individual lives in a region closer to the hometown, δ is large and thus (73) is positive and Lemma 1(iv) follows.

1.1.3 Lemma 1(v)

If

$$U(e,n,\bar{G},\bar{G}_{n+s}/\phi^s;L,n,G^o_{t+1}+\bar{G},\bar{G}) = U(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G}), \quad (82)$$

where $s=0,...,l,...,\bar{n}-n-1$, then there exist $S^n_{n+l,n+l+1}$ that satisfies:

$$\begin{aligned}&\tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})\\ =&\tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,G^o_{t+1}+\bar{G},\bar{G}),\end{aligned} \quad (83)$$

and

$$S(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G}) \le S^n_{n+s,n+s+1} \le S(e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,G^o_{t+1}+\bar{G},\bar{G}), \quad (84)$$

as in Figure 9. From Lemma 1(i) it follows that:

$$\begin{aligned}&\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^o_{t+1}+\bar{G},\bar{G})\\ >&\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G}),\end{aligned} \quad (85)$$

as in Figure 9.

As $\bar{G}_{n+s+1}/\phi^s < \bar{G}_{n+s+1}/\phi^{s+1}$, $u((1+r)S_{t+1}-\bar{G}_{n+s+1}/\phi^s) > u((1+r)S_{t+1}-\bar{G}_{n+s+1}/\phi^{s+1})$ and, thus, $u'((1+r)S_{t+1}-\bar{G}_{n+s+1}/\phi^s) < u'((1+r)S_{t+1}-\bar{G}_{n+s+1}/\phi^{s+1})$. Therefore, $\partial U^{O,L}(S_{t+1},n+1,\bar{G}_{n+s+1}/\phi^s,\bar{G})/\partial S_{t+1} < \partial U^{O,L}(S_{t+1},n,\bar{G}_{n+s+1}/\phi^{s+1},\bar{G})/\partial S_{t+1}$. The derivatives of $U^{O,H}(S_{t+1},n,\bar{G},\bar{G})$ and $U^{O,H}(S_{t+1},n+1,\bar{G},\bar{G})$ with respect to S_{t+1} are the same. The derivatives of $U^Y(S_{t+1},e,n,L,n,G^o_{t+1}+\bar{G},\bar{G})$ and $U^Y(S_{t+1},e,n+1,L,n+1,G^o_{t+1}+\bar{G},\bar{G})$ with respect to S_{t+1} are also the same. Therefore,

$$\frac{\partial\tilde{U}(S_{t+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^s;L,n+1,G^o_{t+1}+\bar{G},\bar{G})}{\partial S_{t+1}} < \frac{\partial\tilde{U}(S_{t+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})}{\partial S_{t+1}}. \quad (86)$$

From (86), the savings $S(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^s;L,n+1,G^o_{t+1}+\bar{G},\bar{G})$ that maximizes $\tilde{U}(S_{t+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^s;L,n+1,G^o_{t+1}+\bar{G},\bar{G})$ is smaller than the savings $S(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})$ that maximizes $\tilde{U}(S_{t+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})$. As in Figure 9, from this fact and (84),

$$\begin{aligned}S(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^s;L,n+1,G^o_{t+1}+\bar{G},\bar{G}) &< S(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})\\ &\le S^n_{n+s,n+s+1}.\end{aligned} \quad (87)$$

For $S_{t+1} \in [S(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G}), S^n_{n+s,n+s+1}]$, both sides of (86) are negative (savings are excessive), and thus the absolute value of the left-hand side is larger than that in the right-hand side. Thus, when S_{t+1} decreases from $S^n_{n+s,n+s+1}$ to $S(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})$, $\tilde{U}(S_{t+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^s;L,n+1,G^o_{t+1}+$

$\bar{G}, \bar{G})$ increases more dramatically than $\tilde{U}(S_{t+1}, e, n, \bar{G}, \bar{G}_{n+s+1}/\phi^{s+1}; L, n, G^o_{t+1} + \bar{G}, \bar{G})$ does. Therefore,

$$\begin{aligned}
&U(e, n, \bar{G}, \bar{G}_{n+s+1}/\phi^{s+1}; L, n, G^o_{t+1} + \bar{G}, \bar{G}) \\
&\qquad - \tilde{U}(S^n_{n+s,n+s+1}, e, n, \bar{G}, \bar{G}_{n+s+1}/\phi^{s+1}; L, n, G^o_{t+1} + \bar{G}, \bar{G}) \\
<&\tilde{U}(S_{t+1}, e, n+1, \bar{G}, \bar{G}_{n+s+1}/\phi^{s}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}) \\
&\qquad - \tilde{U}(S^n_{n+s,n+s+1}, e, n+1, \bar{G}, \bar{G}_{n+s+1}/\phi^{s}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}) \\
<&U(e, n+1, \bar{G}, \bar{G}_{n+s+1}/\phi^{s}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}) \\
&\qquad - \tilde{U}(S^n_{n+s,n+s+1}, e, n+1, \bar{G}, \bar{G}_{n+s+1}/\phi^{s}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}),
\end{aligned} \tag{88}$$

where $S_{t+1} = S(e, n, \bar{G}, \bar{G}_{n+s+1}/\phi^{s+1}; L, n, G^o_{t+1} + \bar{G}, \bar{G})$.

By the similar discussions,

$$\frac{\partial \tilde{U}(S_{t+1}, e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G})}{\partial S_{t+1}} < \frac{\partial \tilde{U}(S_{t+1}, e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G})}{\partial S_{t+1}}, \tag{89}$$

and

$$S(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}) < S(e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G}). \tag{90}$$

From (84), $S(e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G})$ is necessarily larger than $S^n_{n+s,n+s+1}$, but $S(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G})$ is not necessarily so.

Case 1 $S^n_{n+s,n+s+1} \leq S(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G})$ (The case as in Figure 9.)

From (89), when S_{t+1} increases from $S^n_{n+s,n+s+1}$ to $S(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G})$, $\tilde{U}(S_{t+1}, e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G})$ increases more dramatically than $\tilde{U}(S_{t+1}, e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G})$ does. Therefore,

$$\begin{aligned}
&U(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}) \\
&\qquad - \tilde{U}(S^n_{n+s,n+s+1}, e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}) \\
<&\tilde{U}(S_{t+1}, e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G}) \\
&\qquad - \tilde{U}(S^n_{n+s,n+s+1}, e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G}) \\
<&U(e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G}) \\
&\qquad - \tilde{U}(S^n_{n+s,n+s+1}, e, n, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G}),
\end{aligned} \tag{91}$$

where $S_{t+1} = S(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s-1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G})$.

From (88) and (91), it yields that:

$$\begin{aligned}
&\big[\ U(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})-U(e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,G^o_{t+1}+\bar{G},\bar{G})\ \big]\\
&-\big[\ \tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})\\
&\qquad\qquad -\tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,G^o_{t+1}+\bar{G},\bar{G})\ \big]\\
&<\big[\ U(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^o_{t+1}+\bar{G},\bar{G})\\
&\qquad\qquad -U(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})\ \big]\\
&-\big[\ \tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^o_{t+1}+\bar{G},\bar{G})\\
&\qquad\qquad -\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})\ \big].
\end{aligned} \tag{92}$$

From (82) and (83), the left-hand side of (92) is zero (zero minus zero). From this fact and (85), the first term in the right-hand side of (92) must be positive:

$$U(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})<U(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^o_{t+1}+\bar{G},\bar{G}). \tag{93}$$

Case 2 $S^n_{n+s,n+s+1}>S(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})$
From (87), both of $S(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})$ and $S(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s}; L,n+1,G^o_{t+1}+\bar{G},\bar{G})$ are smaller than $S^n_{n+l,n+l+1}$. When S_{t+1} is smaller than $S^n_{n+l,n+l+1}$, $\tilde{U}(S_{t+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^o_{t+1}+\bar{G},\bar{G})$ is necessarily larger than $\tilde{U}(S_{t+1},e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})$. Therefore,

$$\begin{aligned}
&U(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})\\
<&\tilde{U}(S_{t+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^o_{t+1}+\bar{G},\bar{G})\\
<&U(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,G^o_{t+1}+\bar{G},\bar{G}),
\end{aligned} \tag{94}$$

where $S_{t+1}=S(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,G^o_{t+1}+\bar{G},\bar{G})$.

1.2 Proof of Lemma 2

1.2.1 Lemma 2(i)

The parent of an individual is supposed to live in $n'\in\{0,...,\bar{n}\}$. Fix a region in $\{n',...,\bar{n}\}$ and let n^* denote it. (If $n'=0$, then fix a region in $\{1,...,\bar{n}\}$ and let n^* denote it. Thus, $n^*\geq 1$.) Let l^* denote l that satisfies $n^*\in\Theta_l$.

First, consider the utility when an individual in region n^* places her child in a region in $\{n^*+l^*,...,\bar{n}\}$ in the case where her child becomes an unskilled worker. From (13), for $k=0,...,\bar{n}-(n^*+l^*)-1$ there exists n in $\{n^*,...,\bar{n}\}$ that satisfies:

$$\begin{aligned}
&U(e,n,\bar{G},\bar{G}_{n^*+l^*+k}/\phi^{n^*+l^*+k-n};L,n',\bar{G}_n/\phi^{n-n'},\bar{G})\\
=&U(e,n,\bar{G},\bar{G}_{n^*+l^*+1+k}/\phi^{n^*+l^*+1+k-n};L,n',\bar{G}_n/\phi^{n-n'},\bar{G}).
\end{aligned} \tag{95}$$

The larger k is, the larger n that satisfies (95) is. Since $\bar{G}_{n^*+l^*+1+k}/\phi^{n^*+l^*+1+k-n} < \bar{G}_{n^*+l^*+k}/\phi^{n^*+l^*+k-n}$ and $\bar{G}_n/\phi^{n-n'} \leq \bar{G}_{n^*}/\phi^{n^*-n'}$ when $n \geq n^*$, applying Lemma 1(ii) to (95) yields:

$$\begin{aligned}&U(e,n,\bar{G},\bar{G}_{n^*+l^*+k}/\phi^{n^*+l^*+k-n};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\ \geq &U(e,n,\bar{G},\bar{G}_{n^*+l^*+1+k}/\phi^{n^*+l^*+1+k-n};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G}).\end{aligned}$$

Since n^* is closer to n' than n is ($n \geq n^* \geq n'$), in n^* the value of family public goods her parent provides in n' is larger than that in n. Applying Lemma 1(ii) and (v) to it yields:

$$\begin{aligned}&U(e,n^*,\bar{G},\bar{G}_{n^*+l^*}/\phi^{l^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\ \geq &U(e,n^*,\bar{G},\bar{G}_{n^*+l^*+1}/\phi^{l^*+1};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\ > &U(e,n^*,\bar{G},\bar{G}_{n^*+l^*+2}/\phi^{l^*+2};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\ > &...\\ > &U(e,n^*,\bar{G},\bar{G}_{\bar{n}-1}/\phi^{\bar{n}-1-n^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\ > &U(e,n^*,\bar{G},\bar{G}_{\bar{n}}/\phi^{\bar{n}-n^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G}).\end{aligned} \tag{96}$$

If n^* is not the maximum of Θ_{l^*}, (95) holds with $n = n^*$ and $k = 0$. Therefore, the first one in (96) holds with equality. In contrast, if n^* is the maximum of Θ_{l^*}, (95) holds with $k = 0$ and $n > n^*$. Thus the first one in (96) holds with strict inequality.

Next, we consider the utility when an individual in region n^* places her child in a region in $\{n^*, ..., n^* + l^*\}$ in the case where her child becomes an unskilled worker. From (13), when $n^* \in \{m, ..., \bar{n} - 1\}$, for $h = 0, ..., l^* - 1$ there exists n in $\{1, ..., n^* - 1\}$ that satisfies:

$$\begin{aligned}&U(e,n,\bar{G},\bar{G}_{n^*+l^*-h}/\phi^{n^*+l^*-h-n};L,n',\bar{G}_n/\phi^{n-n'},\bar{G})\\ = &U(e,n,\bar{G},\bar{G}_{n^*+l^*-1-h}/\phi^{n^*+l^*-1-h-n};L,n',\bar{G}_n/\phi^{n-n'},\bar{G}).\end{aligned} \tag{97}$$

From (13), when $n^* \in \{2, ..., m - 1\}$, for $h = 0, ..., n^* + l^* - m - 1$ there exists n in $\{1, ..., n^* - 1\}$ that satisfies (97). Since $\bar{G}_{n^*+l^*-h}/\phi^{n^*+l^*-h} < \bar{G}_{n^*+l^*-1-h}/\phi^{n^*+l^*-1-h}$ and $\bar{G}_n/\phi^{n-n'} > \bar{G}_{n^*}/\phi^{n^*-n'}$ when $n < n^*$, applying Lemma 1(ii) to (97) yields:

$$\begin{aligned}&U(e,n,\bar{G},\bar{G}_{n^*+l^*-h}/\phi^{n^*+l^*-h-n};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\ > &U(e,n,\bar{G},\bar{G}_{n^*+l^*-1-h}/\phi^{n^*+l^*-1-h-n};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G}),\end{aligned} \tag{98}$$

for $h = 0, ..., l^* - 1$ when $n^* \in \{m, ..., \bar{n} - 1\}$, and for $h = 0, ..., n^* + l^* - m - 1$ when $n^* \in \{2, ..., m - 1\}$. The smaller h is, the larger n that satisfies (97) is. Since $0 \leq n < n^*$, applying Lemma 1(ii) and (v) to (98) yields:

$$\begin{aligned}
&U(e,n^*,\bar{G},\bar{G}_{n^*+l^*}/\phi^{l^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\
>&U(e,n^*,\bar{G},\bar{G}_{n^*+l^*-1}/\phi^{l^*+1};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\
>&...\\
>&U(e,n^*,\bar{G},\bar{G}_{m+1}/\phi^{m+1-n^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\
>&U(e,n^*,\bar{G},\bar{G}_{m}/\phi^{m-n^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G}).
\end{aligned} \tag{99}$$

From (9), when $n^* \in \{1,...,m-1\}$ for $h = n^*+l^*-m,...,l^*-1$ there exists n'' in $\{0,...,\bar{n}\}$ that satisfies $n'' \in \Omega_{n^*+l^*-h} \cap \Omega_{n^*+l^*+1-h}$:

$$\begin{aligned}
&U(e,0,\bar{G},\bar{G}_{n^*+l^*-h}/\phi^{n^*+l^*-h};H,n'',\bar{G},\bar{G})\\
=&U(e,0,\bar{G},\bar{G}_{n^*+l^*-1-h}/\phi^{n^*+l^*-1-h};H,n'',\bar{G},\bar{G}).
\end{aligned} \tag{100}$$

Since $\bar{G}_{n^*+l^*-h}/\phi^{n^*+l^*-h} < \bar{G}_{n^*+l^*-1-h}/\phi^{n^*+l^*-1-h}$, from Lemma 1(v), when a skilled worker moves to n^* but her substantial income does not change, her utility when she places her child in region n^*+l^*-h by providing $\bar{G}_{n^*+l^*-h}/\phi^{l^*-h}$ is higher than that when she places her child in region n^*+l^*-1-h by providing $\bar{G}_{n^*+l^*-1-h}/\phi^{l^*-1-h}$. In reality, her wage becomes w_L. From (37), $w_L + \bar{G}_{n^*} < w_H$. Thus, applying Lemma 1(ii) yields:

$$\begin{aligned}
&U(e,n^*,\bar{G},\bar{G}_{m}/\phi^{m-n^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\
>&U(e,n^*,\bar{G},\bar{G}_{m-1}/\phi^{m-1-n^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\
>&...\\
>&U(e,n^*,\bar{G},\bar{G}_{n^*+1}/\phi;L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G})\\
>&U(e,n^*,\bar{G},\bar{G}_{n^*};L,n',\bar{G}_{n^*}/\phi^{n^*-n'},\bar{G}).
\end{aligned} \tag{101}$$

(96), (99) and (101) hold for all $n^* \in \{1,...,\bar{n}-1\}$ that is in Θ_l where $l = 1,...,m$ (all Θ_l except Θ_0). Therefore, for all n in Θ_l where $l = 1,...,m$, $U(e,n,\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',\bar{G}_n/\phi^{n-n'},\bar{G})$ is maximized when $G^{O,L}_{t+2}+\bar{G} = \bar{G}_{n+l}/\phi^l$ and $G^{O,L}_{t+2}+\bar{G} = \bar{G}_{n+l+1}/\phi^{l+1}$ in the case where n is not the maximum in Θ_l, and it is maximized only when $G^{O,L}_{t+2}+\bar{G} = \bar{G}_{n+l}/\phi^l$ in the case where n is the maximum in Θ_l. And (96) holds for all n^* in Θ_0 except $\bar{n}$, and thus for all n in Θ_0 except $\bar{n}$, $U(e,n,\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',\bar{G}_n/\phi^{n-n'},\bar{G})$ is maximized when $G^{O,L}_{t+2}+\bar{G} = \bar{G}_n$ and $G^{O,L}_{t+2}+\bar{G} = \bar{G}_{n+1}/\phi$. Therefore, Lemma 2(i)a follows.

Consider $G^o_{t+1}+\bar{G}$ under which there exist $\bar{G}_{n+s+1}/\phi^{n+s+1}$ and $\bar{G}_{n+s}/\phi^{n+s}$ that satisfy:

$$\begin{aligned}
&U(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{n+s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})\\
=&U(e,n,\bar{G},\bar{G}_{n+s}/\phi^{n+s};L,n,G^o_{t+1}+\bar{G},\bar{G}).
\end{aligned} \tag{102}$$

If $G^o_{t+1} + \bar{G} = \bar{G}_n/\phi^{n-n'}$ where n is in Θ_l but not the maximum in it, then $s = l$. From Lemma 1(ii), the larger $G^{O,L}_{t+2}+\bar{G}$ is, the more dramatically $U(e, n, \bar{G}, G^{O,L}_{t+2}+\bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ increases when G^o_{t+1} increases. Therefore, when $G^o_{t+1}+\bar{G}$ becomes a bit larger than that in (102), $U(e, n, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n, G^o_{t+1} + \bar{G}, \bar{G})$ is the largest when $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+s}/\phi^{n+s}$. In this sense, $G^o_{t+1} + \bar{G}$ in (102) is the critical amount above which an individual in region n switches her choice from placing her child in region $n + s + 1$ to placing her child in region $n + s$.

From Lemma 1(v), if (102) holds, then

$$\begin{aligned} &U(e, n+1, \bar{G}, \bar{G}_{n+s+1}/\phi^{s+1}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}) \\ >&U(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n+1, G^o_{t+1} + \bar{G}, \bar{G}). \end{aligned} \tag{103}$$

This means that when an individual moves farther away from the city and her hometown by one region, she prefers $G^{O,L}_{t+2}+\bar{G} = \bar{G}_{n+s+1}/\phi^{s+1}$ to $\bar{G}_{n+s}/\phi^s$ even when the parent also moves from n to $n + 1$ and thus the value of family public goods provided by her parent evaluated in the region where she lives is unchanged. In reality, however, her parent does not move and thus the value of the family public goods falls. Therefore, from Lemma 1(ii):

$$\begin{aligned} &U(e, n+1, \bar{G}, \bar{G}_{n+s+1}/\phi^{s+1}; L, n, G^o_{t+1} + \bar{G}, \bar{G}) \\ >&U(e, n+1, \bar{G}, \bar{G}_{n+s}/\phi^{s}; L, n, G^o_{t+1} + \bar{G}, \bar{G}). \end{aligned} \tag{104}$$

From (102) and (104) and Lemma 1(ii), the critical value of $G^O_{t+1}+\bar{G}$ at which an individual in region $n + 1$ switches her choice of where she will place her child is larger than that at which an individual in region n that is closer to the city will do so. Thus, Lemma 2(i)b follows.

1.2.2 Lemma 2(ii)

Case 1 $n \in \{n', ..., N\}$ and $n' \in \{0, ..., N\}$ (when $n' = 0$, $n \in \{1, ..., N\}$)
Let $\bar{U}(e, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ denote the maximum lifetime utility conditional on that an individual lives in region n and gives a higher education for her child, and her child remains in the region where she lives, regardless of whether or not she becomes a skilled worker and even if she chooses $G^{O,H}_{t+2} = 0$ and $G^{O,L}_{t+2} = 0$. Also, let $\bar{U}(0, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ denote the maximum lifetime utility conditional on that an individual lives in region n and does not give a higher education for her child, and her child remains in the region where she lives even if the individual does not provide family public goods when old ($G^{O,L}_{t+2} = 0$) as defined in section 4.3.3. Obviously, as long as $e > 0$, it holds that:

$$\bar{U}(e, n; L, n', G^o_{t+1} + \bar{G}, \bar{G}) < \bar{U}(0, n; L, n', G^o_{t+1} + \bar{G}, \bar{G}), \tag{105}$$

where $n \in \{n', ..., \bar{n}\}$ and $n' \in \{0, ..., \bar{n}\}$ (when $n' = 0$, $n \in \{1, ..., \bar{n}\}$).

Consider an individual's expected lifetime utility placing her child in a region other than the city even when she becomes a skilled worker. It is lower, or at most equal to, $\bar{U}(e, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. The reason is as follows. When her child becomes an unskilled worker and she is willing to attract her child in region $n+l$, she needs to provide $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l}/\phi^l$. Therefore, if she could attract her child in the region where she lives with $G^{O,L}_{t+2} = 0$, her utility would be higher. Moreover, when her child becomes a skilled worker but she is willing to attract her child in region $n+l$, she needs to provide $G^{O,H}_{t+2} + \bar{G}$ that is larger than, or at least equal to $\bar{G}_{n+l}/\phi^l$. Her child will be eager to live in the city, and thus the parent's provision of family public goods is large enough to overcome the incentive. Therefore, if she could attract her child in the region where she lives with $G^{O,H}_{t+2} = 0$, her utility would be higher.

Therefore, if it holds that:

$$\bar{U}(e, n; L, n', G^o_{t+1} + \bar{G}, \bar{G}) \leq U(e, n, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}), \tag{106}$$

then for such G^o_{t+1} and n, an individual chooses $G^{O,H}_{t+2} = 0$. Below, I show that (106) holds for $n \in \{n', ..., \bar{n}\}$ where $n' \in \{0, ..., \bar{n}\}$ and any G^o_{t+1} (when $n' = 0$, $n \in \{1, ..., \bar{n}\}$). However, for $n \in \{\bar{n}, ..., N\}$, (106) may not hold when G^o_{t+1} is extremely large.

First, consider the case where an individual lives in $n \in \{n', ..., \bar{n}\}$ where $n' \in \{0, ..., \bar{n}\}$ (when $n' = 0$, $n \in \{1, ..., \bar{n}\}$). When an individual and the parent live in the same region $\bar{n}$ and the parent does not privately provide family public goods, (15) holds. Next, suppose that an individual lives in region $\bar{n}$ but the parent lives in region n' that is closer to the city and privately provides family public goods. If G^o_{t+1} and n' satisfy $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} = \bar{G}$, for an individual in region $\bar{n}$ the family public goods provided by her parent in region n' is equal to $\bar{G}$ in the value in region $\bar{n}$. From (15), for such G^o_{t+1} and n', (27) holds. When $n = \bar{n}$, her child will live in that region in the case where her child becomes an unskilled worker even when $G^{O,L}_{t+2} = 0$, and, thus,

$$\bar{U}(0, \bar{n}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) = U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \tag{107}$$

From (27), (105), and (107), when G^o_{t+1} and n' satisfy $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} = \bar{G}$, it holds that:

$$\bar{U}(e, \bar{n}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) < U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \tag{108}$$

When $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ is bigger than $\bar{G}$, both sides of (108) grow equally. Thus, (108) holds. Also, when $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ is smaller than $\bar{G}$, both sides of (108) decrease equally, and, thus, (108) holds for $n' \in \{0, ..., \bar{n}\}$ and any G^o_{t+1}.

Suppose that an individual moves closer to her hometown, n'. This entails the following three effects. First, the value of family public goods provided by her parent rises. This increases both sides of (108) equally. Second, population density rises. This equally reduces both sides of (108). Third, the value of family public goods provided by the child in the city when the child becomes a skilled worker rises. It increases only the right-hand side. Therefore, for n in $\{n', ..., \bar{n}\}$ where $n' \in \{0, ..., \bar{n}\}$ and any G^o_{t+1}, (106) holds. An individual's maximum expected lifetime utility conditional on that $G^{O,H}_{t+2} = 0$ is higher than, or at least equal to, the right-hand side of (106), the maximum expected lifetime utility where the condition $G^{L,O}_{t+2} = 0$ is also imposed. And her maximum expected lifetime utility when she attracts her child in a region other than the city by providing positive $G^{O,H}_{t+2}$ and $G^{O,L}_{t+2}$ is lower than, or at most equal to, the left-hand side of (106), the utility when her child remains in the region where she lives even if $G^{O,H}_{t+2} = 0$ and $G^{O,L}_{t+2} = 0$. Therefore, when she lives in $n \in \{n', ..., \bar{n}\}$ where $n' \in \{0, ..., \bar{n}\}$, an individual will not choose positive $G^{O,H}_{t+2}$, whatever the parent's provision of family public goods G^o_{t+1} is.

Next, consider the case where an individual lives in $n \in \{\bar{n}+1, ..., N\}$. When $n \in \{\bar{n}+1, ..., N\}$, it holds that:

$$\bar{U}(0, \bar{n}+1; L, n', G^o_{t+1} + \bar{G}, \bar{G}) = U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \quad (109)$$

From (16), (105) and (109), which is higher, $\bar{U}(e, \bar{n}+1; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ or $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$, is ambiguous. However, the farther the region where she lives is from the city and the hometown (the larger n is), the lower $U(e, n, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ is. Though $\bar{U}(e, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is also lower, but less dramatically so. Therefore, when n is very large, (106) may not hold, and, thus, an individual is willing to provide large $G^{O,H}_{t+2}$ to attract her child in a region closer to n.

Note that this is the case where edu and n are fixed as $edu = e$ and a region in $n \in \{\bar{n}+1, ..., N\}$, respectively. In reality, however, an individual can freely choose edu and $G^{O,H}_{t+2}$. As we will see in Lemma 5, a combination of $(edu, G^{O,H}_{t+2})$ where $edu = e$ and $G^{O,H}_{t+2} > 0$ will not be chosen.

Case 2 $n = n' = 0$
When an unskilled worker lives in the city, gives a higher education for her child and plans $C^{O,H}_{t+2} = 0$, her child will also live in the city in the case where her child becomes a skilled worker. The larger $C^{O,H}_{t+2}$ is, the more her child is eager to live closer to the unskilled worker in the city, yet the lower the unskilled worker's utility is. Therefore, the unskilled worker will choose $C^{O,H}_{t+2} = 0$.

1.3 Proof of Lemma 3

I prove Lemma 3 by the following three steps. First, I show that the closer to the city the region where an individual lives is (the smaller n in $\{n', ..., \bar{n}\}$ is), the more dramatically the maximum expected lifetime utility conditional on that she lives in that region, $U(e, n, \bar{G}, \bar{G}_{n+s}/\phi^s; L, n', G^o_{t+1} + \bar{G}, \bar{G})$, in which she optimally chooses $n+s$, the region where she places her child in the case that her child becomes an unskilled worker, rises, as $G^o_{t+1} + \bar{G}$ increases. Therefore, the slope of the graph of $U(e, n, \bar{G}, \bar{G}_{n+s}/\phi^s; L, n', G^o_{t+1} + \bar{G}, \bar{G}) - U(e, n+1, \bar{G}, \bar{G}_{n+1+s'}/\phi^{1+s'}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ in which $n+s$ and $n+1+s'$ are optimally chosen has a positive slope, as in Figure 5.

Second, I show that an individual's maximum expected lifetime utility conditional on that she lives in region n and $edu = e$ is equal to the one conditional on that she lives in region $n+1$ and $edu = e$ only at $G^o_{t+1} + \bar{G} = \bar{G}_n/\phi^{n-n'}$ for $n = n', ..., \bar{n} - 1$ as shown in the first step. The former increases more dramatically than the latter as $G^o_{t+1} + \bar{G}$ increases. Therefore, if $G^o_{t+1} + \bar{G}$ is larger (smaller) than $\bar{G}_n/\phi^{n-n'}$, an individual's maximum expected lifetime utility when she lives in region n is higher (lower) than that when she lives in region $n+1$. Similarly, an individual's maximum expected lifetime utility conditional on that she lives in region $n-1$ and $edu = e$ is equal to the one conditional on that she lives in region n and $edu = e$ only at $G^o_{t+1} + \bar{G} = \bar{G}_{n-1}/\phi^{n-n'-1}$ for $n = n'+1, ..., \bar{n}$. If $G^o_{t+1} + \bar{G}$ is larger (smaller) than $\bar{G}_{n-1}/\phi^{n-n'-1}$, an individual's maximum expected lifetime utility when she lives in region $n-1$ is higher (lower) than that when she lives in region n. Since $\phi\bar{G}_{n-1} > \bar{G}_n$ $(n = 1, ..., \bar{n})$, then we see that there exists the range of $G^o_{t+1} + \bar{G}$ that satisfies $\bar{G}_n/\phi^{n-n'} \leq G^o_{t+1} + \bar{G} < \bar{G}_{n-1}/\phi^{n-n'-1}$ with $n = n'+1, ..., \bar{n}$. When $G^o_{t+1} + \bar{G}$ is in this range, an individual's maximum expected lifetime utility when she lives in region n is higher than that when she lives in region $n+1$ and that when she lives in region $n-1$. Iterating this discussion for all $n = n'+1, ..., \bar{n}$, we see that when $G^o_{t+1} + \bar{G}$ is in $\bar{G}_n/\phi^{n-n'} \leq G^o_{t+1} + \bar{G} < \bar{G}_{n-1}/\phi^{n-n'-1}$ with $n = n'+1, ..., \bar{n}$, the maximum expected lifetime utility conditional on that she lives in region n is the highest for the ones conditional on that she lives in any other region in $\{n'+1, ..., \bar{n}\}$.

Third, I show that if $\bar{G}_{n'} \leq G^o_{t+1} + \bar{G}$, the maximum expected lifetime utility conditional on that she lives in region n' is higher than the one conditional on that she lives in any region in $\{1, ..., n'-1\}$. If so, to live in n' is optimal since it is higher than the maximum expected lifetime utility conditional on that she lives in any region in $\{n'+1, ..., \bar{n}\}$ from the discussion above. Then we see that Lemma 3(i) holds.

First, I show that the smaller n is, the more dramatically $U(e, n, \bar{G}, \bar{G}_{n+s}/\phi^s; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ in which $n+s$ is optimally chosen increases when $G^o_{t+1} + \bar{G}$ increases. From Lemma 2(i)b, if an individual lives in region n and her optimal choice of $G^{O,L}_{t+2} + \bar{G}$ is $\bar{G}_{n+s}/\phi^s$, then her optimal choice of $G^{O,L}_{t+2} + \bar{G}$ is at least $\bar{G}_{n+s}/\phi^{s+1}$ when she lives in region $n-1$. It is G_{n+s-1}/ϕ^s or larger than that. From Lemma 1(ii), $U(e, n-1, \bar{G}, \bar{G}_{n+s}/\phi^{s+1}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ and $U(e, n-1, \bar{G}, G^{O,L}_{t+2} + \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ where $G^{O,L}_{t+2} + \bar{G}$ is equal to or larger than G_{n+s-1}/ϕ^s increase more dramatically than $U(e, n, \bar{G}, \bar{G}_{n+s}/\phi^s; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ does when $G^o_{t+1} + \bar{G}$ increases, since $\bar{G}_{n+s}/\phi^s < \bar{G}_{n+s}/\phi^{s+1} < \bar{G}_{n+s-1}/\phi^s$. In addition, in region $n-1$ the value of family public goods provided by the parent in n' increases more drastically than that in region n does when the parent increases $G^o_{t+1} + \bar{G}$.

Second, I show that an individual's maximum expected lifetime utility conditional on that she lives in region n and $edu = e$ is equal to the one conditional on that she lives in region $n+1$ at $G^o_{t+1} + \bar{G} = \bar{G}_n/\phi^{n-n'}$ for $n \in \{n', ..., \bar{n}-1\}$.

If $n \in \Theta_l$ but not the maximum in Θ_l, from Lemma 2(i)a, the maximum expected lifetime utility conditional on that she lives in region n and $edu = e$ but freely chooses G^o_{t+2} is achieved by choosing $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l}/\phi^l$ or $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l+1}/\phi^{l+1}$ when her parent in region n' provides family public goods in the amount $\bar{G}_n/\phi^{n-n'}$ and thus equal to $U(e, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ and $U(e, n, \bar{G}, \bar{G}_{n+l+1}/\phi^{l+1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$. This is equal to (13) and the left-hand side of (11). On the other hand, the maximum expected lifetime utility conditional on that she lives in region $n+1$ and $edu = e$ but freely chooses G^o_{t+2} is equal to $U(e, n+1, \bar{G}, \bar{G}_{n+l+1}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ (place her child in region $n+l+1$ that is l regions farther away from region $n+1$) when her parent in region n' provides family public goods in the amount $\bar{G}_n/\phi^{n-n'}$. This is equal to the right-hand side of (11). The reason is as follows. Since n is not the largest in Θ_l, $n+1$ is also in Θ_l. Therefore, from Lemma 2(i)b, if her parent provides family public goods in the amount $G^o_{t+1} + \bar{G}$ in the range between $\bar{G}_{n+1}/\phi^{n+1-n'}$ (that is strictly smaller than $\bar{G}_n/\phi^{n-n'}$) and the value that is strictly larger than $\bar{G}_n/\phi^{n-n'}$, the maximum expected lifetime utility conditional on that she lives in region $n+1$ and $edu = e$ but freely chooses G^o_{t+2} is achieved by choosing $G^{O,L}_{t+2} + \bar{G} = \bar{G}_{n+l+1}/\phi^l$ and equal to $U(e, n+1, \bar{G}, \bar{G}_{n+l+1}/\phi^l; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. From these facts and (11), the maximum expected lifetime utility conditional on that $edu = e$ and she lives in region n is equal to the one conditional on that $edu = e$ and she lives in region $n+1$ when her parent chooses $\bar{G}_n/\phi^{n-n'}$.

If n is the maximum in Θ_l, from Lemma 2(i)a, the maximum expected lifetime utility conditional on that she lives in region n and $edu = e$ but freely chooses G^o_{t+2} is equal to $U(e, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ when her parent in region n' provides family public goods in the amount $\bar{G}_n/\phi^{n-n'}$. This is equal to the left-hand side of (12). On the other hand, the maximum expected lifetime utility conditional on that she lives in region $n+1$ and $edu = e$ but freely chooses G^o_{t+2} is equal to $U(e, n+1, \bar{G}, \bar{G}_{n+l}/\phi^{l-1}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G})$ when her parent in region n' provides family public goods in the amount $\bar{G}_n/\phi^{n-n'}$. This is equal to the right-hand side of (12). The reason is as follows. Since n is the largest in Θ_l, $n+1$ is in Θ_{l-1}. Therefore, from Lemma 2(i)b, if her parent provides family public goods $G^o_{t+1} + \bar{G}$ in the range between $\bar{G}_{n+1}/\phi^{n+1-n'}$ (that is strictly smaller than $\bar{G}_n/\phi^{n-n'}$) and the value that is strictly larger than $\bar{G}_n/\phi^{n-n'}$, the maximum expected lifetime utility conditional on that she lives in region $n+1$ and $edu = e$ but freely chooses G^o_{t+2} is equal to $U(e, n+1, \bar{G}, \bar{G}_{n+l}/\phi^{l-1}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. From these facts and (12), the maximum expected lifetime utility conditional on that $edu = e$ and she lives in region n is equal to the one conditional on that $edu = e$ and she lives in region $n+1$ when her parent chooses $\bar{G}_n/\phi^{n-n'}$.

Third, I show that an individual's maximum expected lifetime utility conditional on that $edu = e$ and she lives in region n' when her parent provides family public goods that is larger than $\bar{G}_{n'}$ is higher than the one conditional on that she lives in a region that is closer to the city, $n \in \{1, ..., n'-1\}$.

When she lives closer to the city, an individual's expected utility rises since the value of family public goods provided by her child in the city when her child becomes a skilled worker rises. On the other hand, her disutility rises as the population is larger. The rise in the disutility overwhelms that in the expected utility.

If her parent lives in a region closer to the city than the individual initially lives, she can enjoy higher value of family public goods provided by her parent by moving closer to the city. Especially, when her parent provides family public goods in the amount of $\bar{G}_n/\phi^{n-n'}$ and she moves from region $n+1$ to region n $(n+1 > n')$, the rise of the value just compensates the gap between a rise in the expected utility and that in the disutility in equilibrium as discussed above.

In contrast, if her parent lives in a region that is the same or farther away from the city than the region that the individual initially lives, the value of family public goods provided by her parent falls when she moves closer to the city. Therefore, her utility necessarily falls. And when her parent's provision increases, the fall in the value by moving closer to the city becomes more drastic. Therefore, an individual will not live closer to the city than the parent does.

1.4 Proof of Lemma 4

When an individual and the parent live in the same region $\bar{n}$ and the parent does not privately provide family public goods, (15) holds and an individual prefers $edu = e$. Next, suppose that an individual lives in region $\bar{n}$ but the parent lives in region n' that is closer to the city and privately provides family public goods. If G^o_{t+1} and n' satisfy $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} = \bar{G}$, for an individual in region $\bar{n}$ the family public goods provided by her parent in region n' is equal to $\bar{G}$ in the value in region $\bar{n}$. From (15), for such G^o_{t+1} and n', (27) holds. When $n = \bar{n}$, her child will live in that region in the case where she becomes an unskilled worker even when $G^{L,O}_{t+2} = 0$, and, thus, (107) holds. Thus, when $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} = \bar{G}$ and $n = \bar{n}$, from (27) and (107), it follows that:

$$\bar{U}(0, \bar{n}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) \leq U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \tag{110}$$

where $\bar{U}(0, n; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is the maximum lifetime utility conditional on that an individual lives in region n and does not give a higher education for her child, and her child remains in the region where she lives even if the individual does not provide family public goods when old $(G^{O,L}_{t+2} = 0)$.

When $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ is bigger than $\bar{G}$, both sides of (110) grow. However, from Lemma 1(iii), the right-hand side grows more drastically. Thus (110) holds when $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} \geq \bar{G}$.

Suppose that an individual moves closer to the region where her parent lives. This entails the following three effects. First, the value of family public goods provided by her parent rises. This increases right-hand side more dramatically. That is because the rise in the value of family public goods provided by the parent increases the substantial income and it in turn reduces the substantial costs of education as Lemma 1(iii) indicates. Second, population density rises. This reduces both sides of (110) equally. Third, the value of family public goods provided by the child in the city when the child becomes a skilled worker rises. It increases only the right-hand side. Therefore, for n in $\{n', ..., \bar{n}\}$ it follows that:

$$\bar{U}(0, n; L, n', G^o_{t+1} + \bar{G}, \bar{G}) \leq U(e, n, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \tag{111}$$

An individual's maximum expected lifetime utility conditional on that $edu = e$ is higher than, or at least equal to, the right-hand side of (111), the maximum expected lifetime utility where the conditions $G^{O,L}_{t+2} = 0$ and $G^{O,H}_{t+2} = 0$ are imposed. Moreover, her maximum expected lifetime utility conditional on that $edu = 0$ is lower than, or at most equal to, the left-hand side of (111), the utility when her child remains in the region where she lives even if $G^{O,L}_{t+2} = 0$. Therefore, when she lives in $n \in \{n', ..., \bar{n}\}$ where $n' \in \{0, ..., \bar{n}\}$, an individual prefers $edu = e$ to $edu = 0$. Also, where she lives and where she places her child by providing $G^{O,L}_{t+2}$ and $G^{O,H}_{t+2}$ depend on the size of $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ as Lemmas 2 and 3 indicate. Therefore, her maximum expected lifetime utility when an individual freely chooses $edu = \{0, e\}$ and G^o_{t+2} but n is limited in $\{0, ..., \bar{n}\}$ is equal to or at least $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ when $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} \geq \bar{G}$. Thus Lemma 4(i) holds.

When $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ is smaller than $\bar{G}$, (27) does not necessarily hold even if (15) holds, since the right-hand side of (27) decreases more dramatically than the left-hand side as G^o_{t+1} becomes smaller than the one in $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} = \bar{G}$.

First, consider the case where (27) holds. When (27) holds, as we saw above, (111) holds. Then an individual prefers $edu = e$. She lives in region $\bar{n}$ as Lemma 3 indicates, and chooses $G^{O,L}_{t+2} = 0$ and $G^{O,H}_{t+2} = 0$ as Lemma 2 indicates, and, thus, her maximum expected lifetime utility is equal to $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

Next, we consider the case where (27) does not hold. That is,

$$U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) > U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \tag{112}$$

The maximum utility conditional on that $edu = e$ and n is limited in $\{0, ..., \bar{n}\}$ is achieved when $n = \bar{n}$ and equal to the right-hand side. From (112), $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is larger than that. It is also larger than $U(0, n, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ with $n = 0, ..., \bar{n} - 1$. Therefore, an individual will not give a higher education for her child ($edu = 0$), and lives in $\bar{n}$. The maximum expected lifetime utility is $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. Therefore, Lemma 4(ii) holds.

When $G^o_{t+1} = 0$ and $\bar{G} = 0$, (15) and (110) are the same. As an individual moves closer to the city, the gap between the right-hand side and the left-hand side in (111) expands. Thus, an individual prefers $edu = e$ to $edu = 0$. Also, when her choice of residential region is limited in $\{0, ..., \bar{n}\}$, from Lemmas 2 and 3 she chooses $n = \bar{n}$, $G^{O,L}_{t+2} = 0$ and $G^{O,H}_{t+2} = 0$. Therefore, her maximimum expected lifetime utility when an individual freely chooses $edu = \{0, e\}$ and G^o_{t+2} but n is limited in $\{0, ..., \bar{n}\}$ is equal to $U(e, \bar{n}, 0, 0; L, n', 0, 0)$. Therefore, Lemma 4(iii) holds.

1.5 Proof of Lemma 5

First, consider an individual's choices of where to live and $G^{O,L}_{t+2}$ when her choice of education is given as $edu = 0$. Second, consider an individual's choice of where to live and $G^{O,L}_{t+2}$ when her choices of education and $G^{O,H}_{t+2}$ are given as $edu = e$ and $G^{O,H}_{t+2} = 0$, respectively. Finally, consider whether an individual can enhance her utility by choosing positive $G^{O,H}_{t+2}$.

First, consider the choices of n and $G^{O,L}_{t+2}$ conditional on that $edu = 0$. Wherever in regions $n \in \{\bar{n} + 1, ..., N\}$ she lives, an individual can make her child stay in the region where she lives without providing family public goods privately in case where her child becomes an unskilled worker and thus she is willing to do so ($G^{O,L}_{t+2} = 0$). From (21), population is the same in regions $n \in \{\bar{n} + 1, ..., N\}$. The difference in the utility in the regions stems from the difference in the value of family public goods provided by her parent $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$. The values are the highest when she lives in $\bar{n} + 1$, that is the closest to the region where her parent lives. Thus, the maximum expected lifetime utility in this case is $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$. However, if $\bar{G} = G^o_{t+1} = 0$, her expected lifetime utility is the same wherever in $n \in \{\bar{n} + 1, ..., N\}$ she lives.

Second, consider the choices of n and $G^{O,L}_{t+2}$ conditional on that $edu = e$ and $G^{O,H}_{t+2} = 0$. From the discussions that is the same as above, $G^{O,L}_{t+2} = 0$. As $edu = e$ and $G^{O,H}_{t+2} = 0$, her child will move to the city when she becomes a skilled worker. The difference in the utility in the regions stems from the differences in the value of family public goods provided by her child in the city when her child becomes a skilled worker as well as that provided by her parent $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$. The values are the highest when she lives in $\bar{n} + 1$, that is the closest to the region where her parent lives and the city. Thus, the maximum expected lifetime utility in this case is $U(e, \bar{n} + 1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

Third, consider whether an individual can enhance her utility by choosing positive $G^{O,H}_{t+2}$. From Lemma 2(ii), when she lives in a region in $n \in \{\bar{n} + 1, ..., N\}$, she may be able to attract her child in a closer region by choosing a large amount of $G^{O,H}_{t+2}$ even when her child becomes a skilled worker, and may be willing to do so. However, the maximum expected lifetime utility in this case is at most equal to $\bar{U}(e, \bar{n} + 1; L, n', G^o_{t+1} + \bar{G}, \bar{G})$, the utility level when she can attract her child in region n without providing G^o_{t+2}, though in reality it is impossible. Also, from (105) and (107), $\bar{U}(e, \bar{n}+1; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ is strictly lower than $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$, the utility when she does not give a higher education to her child. Therefore, a combination of $edu = e$ and $G^{O,H}_{t+2} > 0$ will not take place.

In summary, the maximum expected lifetime utility is $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ or $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ when $\bar{G} > 0$ or $G^o_{t+1} > 0$. And it is $U(e, \bar{n}+1, 0, 0; L, n', 0, 0)$ or $U(0, n, 0, 0; L, n', 0, 0)$ where $n \in \{\bar{n}+1, ..., N\}$ when $\bar{G} = 0$ and $G^o_{t+1} = 0$.

1.6 Proof of Proposition, Case L(i): $n' \in \{0, ..., \bar{n}\}$ and $0 < \bar{G} \leq (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ or $\bar{G} = 0$ and $G^o_{t+1} > 0$

Suppose that the parent lives in region $n' \in \{0, ..., \bar{n}\}$ and provides family public goods in the amount $G^o_{t+1} + \bar{G}$ that satisfies $0 < \bar{G} \leq (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$. When her choice of where to live is limited in $n \in \{0, ..., \bar{n}\}$, an individual prefers $edu = e$ to $edu = 0$ and her maximum expected lifetime utility is equal to or at least $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ from Lemma 4(i). When $n \in \{\bar{n}+1, ...\}$, her her maximum expected lifetime utility is $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ or $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ from Lemma 5(i). Therefore, compare the sizes of these three, $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$, $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ and $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$.

First, consider which is higher, $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ or $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$. If (16) holds, it necessarily follows that:

$$U(e, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}) < U(0, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}). \tag{113}$$

(16) is the case where the individual and the parent live in the same region, $\bar{n}+1$. In contrast, (113) is the case where the parent lives in region $\bar{n}$. Therefore, the family public goods provided by her parent through the government in the value in region $\bar{n}+1$ is smaller in (113) than in (16), and thus both sides in (113) are lower than those in (16). However, from Lemma 1(iii), the gap in the left-hand sides is larger than that in the right-hand sides, and thus (113) holds. From (17) and (113), it holds that:

$$U(e, \bar{n}+1, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}) < U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}). \tag{114}$$

If $0 < \bar{G} \leq (G^o_{t+1}+\bar{G})\phi^{(\bar{n}-n')}$, $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ and $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ are higher than the left-hand side and right-hand side of (114), respectively. However, from Lemma 1(iii), the difference between $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', \bar{G}, \bar{G})$ and $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ is larger than that between $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', \bar{G}, \bar{G})$ and $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$, and thus it holds that:

$$U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G}) < U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G}). \tag{115}$$

Next, consider which is higher, $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ or $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$. When $0 < \bar{G} \leq (G^o_{t+1}+\bar{G})\phi^{(\bar{n}-n')}$, from (17) and Lemma 1(iii), (28) holds ($U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is equal to or higher than $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$).

In summary, $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is the highest, and thus the optimal edu is $edu = e$. The optimal n and $G^{O,L}_{t+2}$ depend on $G^o_{t+1} + \bar{G}$ as Lemmas 2 and 3 indicate. Thus, the behaviors (3) described in Case L(i) where $0 < \bar{G} \leq (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ are optimal.

If (17) holds with bind, (28) holds with bind when $(G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} = \bar{G}$. Therefore, both of $(e, n) = (e, \bar{n})$ and $(e, n) = (0, \bar{n}+1)$ are optimal choices. Otherwise, the optimal choice of edu is $edu = e$. Thus, the behaviors (3) described in Case L(i) where $0 < \bar{G} = (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')}$ and (17) holds with bind are optimal.

1.7 Proof of Proposition, Case L(i): $n' \in \{0, ..., \bar{n}\}$ and $0 < (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} < \bar{G}$

Suppose that the parent lives in region $n' \in \{0, ..., \bar{n}\}$ and provides family public goods in the amount $G^o_{t+1} + \bar{G}$ that is small enough to satisfy $0 < (G^o_{t+1} + \bar{G})\phi^{(\bar{n}-n')} < \bar{G}$. In this case, (27) and (28) do not necessarily hold. When her choice of where to live is limited in $n \in \{0, ..., \bar{n}\}$, an individual prefers $edu = e$ to $edu = 0$ and her maximum expected lifetime utility is equal to $U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ in the case that (27) holds. If (27) does not hold, she prefers $edu = 0$ and her maximum expected lifetime utility is equal to $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ from Lemma 4(ii). When $n \in \{\bar{n}+1, ...\}$, her maximum expected lifetime utility is $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ or $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ from Lemma 5(i).

First, consider the case where (27) does not hold ((112) holds). In this case, compare the sizes of $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$, $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ and $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. From the fact that (27) does not hold and Lemma 1(iii), it holds that:

$$U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) \leq U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \quad (116)$$

Thus, I compare $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ and $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$.

If (28) holds, from (28) and (112), $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is the higher than $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. Her optimal choice is $edu = 0$ and $n = \bar{n}$, as in (29) where $\psi(L, n', G^o_{t+1} + \bar{G}, \bar{G}) = \bar{n}$.

In contrast, if (28) does not hold, her optimal choice is $edu = 0$ and $n = \bar{n}$ or $\bar{n}+1$. Since neither (27) nor (28) holds, both of $U(0, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ and $U(0, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ are higher than $U(e, \bar{n}+1, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. However, whih is larger is ambiguous. Therefore, depending on whether or not (30) holds, she decides where to live, as in (29).

Next, consider the case where (27) holds but (28) does not hold. When (27) holds, an individual's expected lifetime utility when n is limited in $\{0,...,\bar{n}\}$ is maximized when an individual chooses $edu = e$ and $n = \bar{n}$, to be $U(e,\bar{n},\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$ as Lemma 4(ii) indicates. Then compare the sizes of $U(e,\bar{n},\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$, $U(e,\bar{n}+1,\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$ and $U(0,\bar{n}+1,\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$. If (16) holds, then (116) holds. That is because, the left-hand side and the right-hand side of (116) are lower than those in (16), as $(G^o_{t+1}+\bar{G})\phi^{(\bar{n}-n')} \leq \bar{G}$. From Lemma 1(iii), the gap in the left-hand sides is larger. And as (28) does not hold, $U(0,\bar{n}+1,\bar{G},\bar{G};L,n',G^o_{t+1},\bar{G})$ is higher than $U(e,\bar{n},\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$. Therefore, the optimal choice is $e = 0$ and $n = \bar{n}+1$, as in (29) where $\psi(L,n',G^o_{t+1}+\bar{G},\bar{G}) = \bar{n}+1$.

Finally, from Lemma 2, she will place her child in region $\bar{n}$ without providing family public goods privately ($G^{O,L}_{t+2} = 0$). Thus, the behaviors described in Case L(i) where $0 \leq (G^o_{t+1}+\bar{G})\phi^{(\bar{n}-n')} < \bar{G}$ are optimal.

1.8 Proof of Proposition, Case L(ii): $n' \in \{\bar{n}+1,...\}$ and $\bar{G}$ is positive

First, I show that when her parent lives in $n' \in \{\bar{n}+1,...\}$, an individual's optimal choice conditional on that she does not give a higher education for her child is such that she remains in region n' where she was born, and plans not to provide family public goods when old, $G^{O,L}_{t+2} = 0$. The maximum expected lifetime utility conditional on that $edu = 0$ is $U(0,n',\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$. After that I discuss the conditions under which the expected lifetime utility conditional on that she gives a higher education exceeds $U(0,n',\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$.

Given that $edu = 0$, she will choose $n = n'$ and $G^{O,L}_{t+2} = 0$. The reason is as follows. When edu is fixed as $edu = 0$, her child will necessarily become an unskilled worker and thus will not live in the city. Thus, she is not willing to live in the city or regions close to the city, $n \in \{0,...,\bar{n}\}$, that has larger population than region $n \in \{\bar{n}+1,...,N\}$ has. When the parent lives in n' and the individual lives in n, the value of family public goods provided by her parent is $(G^o_{t+1}+\bar{G})\phi^{(n-n')}$. The value is the highest when she lives in $n = n'$. When she lives in a region in $n \in \{\bar{n}+1,...\}$, an individual can make her child stay in the region where she lives without providing family public goods privately in case where her child becomes an unskilled worker and thus she is willing to do so ($G^{O,L}_{t+2} = 0$).

Next, consider the conditions under which she will give a higher education for her child. I consider the existence of the critical value of $G^o_{t+1}+\bar{G}$ above which the individual swiches edu from 0 to e. I will show that if it exists, it is larger than $\bar{G}$. I also show that the larger n' is, the larger it is and the less plausible it exists.

From (E2) and (F2), when $G^o_{t+1} = 0$, $U(0,n',\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$ is higher than $U(e,n,\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$ where $n = \bar{n},\bar{n}+1,...,n',...,N$. When G^o_{t+1} grows, $U(e,n,\bar{G},\bar{G};L,n',G^o_{t+1}+\bar{G},\bar{G})$ with n that is n' or close to it increases more dramatically than $U(0, n', \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ does. Let $G^{**}_{n'}$ denote $G^o_{t+1} + \bar{G}$ that satisfies:

$$U(e, n', \bar{G}, \bar{G}; L, n', G^{**}_{n'}, \bar{G}) = U(0, n', \bar{G}, \bar{G}; L, n', G^{**}_{n'}, \bar{G}), \tag{117}$$

and let $\tilde{n}$ denote the region farthest from the city in regions where (117) holds (the largest n' with which $G^{**}_{n'}$ exists). Obviously, $\bar{n} \leq \tilde{n}$. From Lemma 1(iii), $G^{**}_{n'}$ must be larger than, or at least equal to $\bar{G}$. Moreover, the largher n' is, the larger $G^{**}_{n'}$ that makes (117) hold. That is because, the larger n' is, the lower $U(e, n', \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is, whereas $U(0, n', \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ is constant. Therefore, with larger $G^o_{t+1} + \bar{G}$, these two are balanced.

Let $G^*_{n'}$ denote $G^o_{t+1}+\bar{G}$ that satisfies (31). When $n' \leq \tilde{n}$, by definition, $U(e, n', \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ is equal to $U(0, n', \bar{G}, \bar{G}; L, n', G^o_{t+1}+\bar{G}, \bar{G})$ with $G^o_{t+1}+\bar{G} = G^{**}_{n'}$. However, $U(e, n, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ with $n \in \{\bar{n}, ..., n' - 1\}$ may be larger than that. If so, with $G^*_{n'}$ that is smaller than $G^{**}_{n'}$, (31) holds. In contrast, when $n' \geq \tilde{n} + 1$, $G^{**}_{n'}$ does not exist. However, for $n \in \{\bar{n}, ..., \tilde{n} - 1\}$ and very large $G^*_{n'}$, $U(e, n, \bar{G}, \bar{G}; L, n', G^*_{n'}, \bar{G})$ may be equal to $U(0, n', \bar{G}, \bar{G}; L, n', G^*_{n'}, \bar{G})$. [26]

Below, I show that even if $G^*_{n'}$ exists, $G^*_{n'+1}$ does not necessarily exist. I also show that if $G^*_{n'+1}$ exists, $G^*_{n'+1} > G^*_{n'}$.

First, consider the case where $n' + 1 \leq \tilde{n}$. From (31) it yields that:

$$\begin{aligned} & U(0, n', \bar{G}, \bar{G}; L, n' + 1, G^*_{n'}/\phi, \bar{G}) \\ = MAX[\ & U(e, n', \bar{G}, \bar{G}; L, n' + 1, G^*_{n'}/\phi, \bar{G}), \\ & ... \\ & U(e, \bar{n}, \bar{G}, \bar{G}; L, n' + 1, G^*_{n'}/\phi, \bar{G})\ ] \end{aligned} \tag{118}$$

However, $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^*_{n'}/\phi, \bar{G}) > U(0, n', \bar{G}, \bar{G}; L, n'+1, G^*_{n'}/\phi, \bar{G})$. Therefore, when the parent lives in region $n' + 1$ and chooses $G^o_{t+1} = G^*_{n'}/\phi$, an individual will live in region $n' + 1$. Depending on which is larger, $U(e, n' + 1, \bar{G}, \bar{G}; L, n' + 1, G^*_{n'}/\phi, \bar{G})$ or $U(0, n' + 1, \bar{G}, \bar{G}; L, n' + 1, G^*_{n'}/\phi, \bar{G})$, an individual chooses $edu = e$ or not.

If $n' + 1 \leq \tilde{n}$ and if $U(e, n' + 1, \bar{G}, \bar{G}; L, n' + 1, G^*_{n'}/\phi, \bar{G}) > U(0, n' + 1, \bar{G}, \bar{G}; L, n' + 1, G^*_{n'}/\phi, \bar{G})$, from Lemma 1(iii), $G^o_{t+1}+\bar{G}$ that equalizes $U(e, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+$

[26] As shown in the proof of Lemma 3, when she initially lives in $n \leq \bar{n}$ and moves to $n-1$ that is closer to the city, the value of family public goods provided by her child in the city when her child becomes a skilled worker rises, but it is overwhelmed by a rise in disutility caused by larger population. In addition, when her parent lives farther away from the city than the individual, the value of family public goods provided by her parent falls. Thus an individual strictly prefers to live in n to $n-1$. However, given that she chooses $edu = e$ and she moves from $n \geq \bar{n} + 1$ to $n - 1$, her utility necessarily rises without taking into account the change in the value of family public goods by her parent. That is because regions $\bar{n}, \bar{n} + 1, ..., N$ have the same population and thus there is no rise in disutility stemming from a difference in population. Even when we take into account a change in the value of family public goods by her parent, living farther away from the hometown region n' and closer to the city may improve her expected lifetime utility.

$\bar{G}, \bar{G})$ and $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ is strictly smaller than $G^*_{n'}/\phi$. By definition, such $G^o_{t+1}+\bar{G}$ is $G^{**}_{n'+1}$. And none of $U(e, n, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}, \bar{G})$ where $n \in \{\bar{n}, ..., n'\}$ exceeds $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}, \bar{G})$. Hence, $G^*_{n'+1} = G^{**}_{n'+1} < G^*_{n'}/\phi$. As we saw, $G^{**}_{n'+1} > G^{**}_{n'}$. Also, the value of $G^*_{n'}$ is smaller than, or at most equal to $G^{**}_{n'}$ ($G^*_{n'} \leq G^{**}_{n'}$). Thus, $G^*_{n'} < G^*_{n'+1}$.

In contrast, if $n'+1 \leq \tilde{n}$ but if $U(e, n'+1, \bar{G}, \bar{G}; L, n'+1, G^*_{n'}/\phi, \bar{G}) \leq U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^*_{n'}/\phi, \bar{G})$, the critical $G^*_{n'+1}$ above which an individual will choose $edu = e$, is equal to, or larger than $G^*_{n'}/\phi$. If $G^o_{t+1}+\bar{G}$ increases from $G^*_{n'}/\phi$ to $G^{**}_{n'+1}$, an individual will choose $edu = e$ and $n = n'+1$. However, before $G^o_{t+1}+\bar{G}$ reaches $G^{**}_{n'+1}$, one of $U(e, n, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ where $n \in \{\bar{n}, ..., n'\}$ may catch up with $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$. In either case, $G^*_{n'+1} > G^*_{n'}/\phi > G^*_{n'}$.

Second, consider the case where $n'+1 > \tilde{n}$. In this case, $G^{**}_{n'+1}$ does not exist. In this case as well, $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^*_{n'}/\phi, \bar{G}) > U(0, n', \bar{G}, \bar{G}; L, n'+1, G^*_{n'}/\phi, \bar{G})$. Therefore, when the parent lives in region $n'+1$ and chooses $G^o_{t+1}+\bar{G} = G^*_{n'}/\phi$, an individual will live in region $n'+1$. Moreover, an individual will not give a higher education for her child ($edu = 0$), since $U(e, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ is necessarily lower than $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$. From Lemma 1(iii), when her parent increases G^o_{t+1}, $U(e, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ increases more dramatically than $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ does. When $G^o_{t+1}+\bar{G}$ is increasing from $G^*_{n'}/\phi$, $U(e, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ does not catch up with $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$, yet one of $U(e, n, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ where $n \in \{\bar{n}, ..., \tilde{n}\}$ may catch up with it. By definition, such a large $G^o_{t+1}+\bar{G}$ is $G^*_{n'+1}$. However, the longer the distance between the region where her parent resides and that where an individual resides, the less plausible it is that $U(e, n, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ increase more dramatically than $U(0, n'+1, \bar{G}, \bar{G}; L, n'+1, G^o_{t+1}+\bar{G}, \bar{G})$ does. Therefore, $G^*_{n'+1}$ does not necessarily exist, especially when $n'+1$ is very large. Even if it exists, $G^*_{n'+1}$ is strictly larger than $G^*_{n'}/\phi$.

Therefore, if her parent lives in $n' \in \{\bar{n}+1, ...\}$ and provides family public goods $G^o_{t+1}+\bar{G}$ that is so large as to satisfy $G^*_{n'} \leq G^o_{t+1}+\bar{G}$, then an individual will give a higher education for her child, and live in region n that gives the maximum of the right-hand side of (31), as (32) in Case L(ii) indicates. In contrast, if $G^o_{t+1}+\bar{G}$ is smaller, an individual's choice is (4) in Case L(ii).

1.9 Proof of Proposition, Case L(iii): $G^o_{t+1} = 0$ and $\bar{G} = 0$

When the parent lives in region $n' \in \{0, ..., \bar{n}\}$ and does not provide family public goods, $G^o_{t+1} = \bar{G} = 0$, an individual's maximum expected lifetime utility conditional on that $n \in \{0, ..., \bar{n}\}$ is equal to $U(e, \bar{n}, 0, 0, 0; L, n', 0, 0)$ from Lemma 4(iii), the one conditional on that $n \in \{\bar{n}+1, ...\}$ is $U(e, \bar{n}+1, 0, 0, 0; L, n', 0, 0)$ or $U(0, n, 0, 0, 0; L, n', 0, 0)$ for all n in $\{\bar{n}+1, ...\}$ from Lemma 5(ii). Therefore, I compare the sizes of these three, $U(e, \bar{n}, 0, 0, 0; L, n', 0, 0)$, $U(e, \bar{n}+1, 0, 0, 0; L, n', 0, 0)$ and $U(0, n, 0, 0, 0; L, n', 0, 0)$ where

$n \in \{\bar{n}+1, ...\}$.

From (16), $U(e, \bar{n}+1, 0, 0, 0; L, n', 0, 0)$ is lower than, or at most equal to $U(0, n, 0, 0, 0; L, n', 0, 0)$ with $n \in \{\bar{n}+1, ...\}$. From (19), $U(e, \bar{n}, 0, 0, 0; L, n', 0, 0)$ is equal to $U(0, n, 0, 0, 0; L, n', 0, 0)$ with $n \in \{\bar{n}+1, ...\}$. Therefore, both of $(e, n) = (e, \bar{n})$ and $(e, n) = (0, n)$ where $n \in \{\bar{n}+1, ...\}$ are optimal choices.

2 Proof of Proposition: A Skilled Worker's Behavior

Prior to the proof of Proposition, I prove the following Lemma.

Lemma 6

Suppose the parent of a skilled individual lives in region n' that is closer to the city ($n' \in \{0, ..., \bar{n}\}$). Suppose, for a skilled individual, education and where to live are fixed, in that she gives a higher education for her child and lives in the city.

(i) Plan of $G^{O,L}_{t+2}$, the provision of family public goods in the case where the child becomes an unskilled worker. The higher the value of family public goods provided by the parent in region n' but evaluated in the city, $(G^o_{t+1}+\bar{G})\phi^{n'}$, is, the closer the region in which a skilled individual will place her child by providing larger $G^{O,L}_{t+2}+\bar{G}$ in the case where her child becomes an unskilled worker. More specifically,

- *if $0 < \bar{G} < (G^o_{t+1}+\bar{G})\phi^{n'}$, a skilled individual will place her child in region 1 by providing $\frac{\bar{G}_1}{\phi}$.*
- *If $0 < \bar{G} = (G^o_{t+1}+\bar{G})\phi^{n'}$, a skilled individual will place her child in a region in $l = 1, ..., r-1$ or r by providing $\frac{\bar{G}_l}{\phi^l}$.*
- *If $0 < \bar{G}\phi^n \leq (G^o_{t+1}+\bar{G})\phi^{n'} < \bar{G}\phi^{n-1}$ where $n \in \Omega_l$ ($n = 1, ..., \bar{n}$ and $l = r, ..., m$), then a skilled individual will place her child in region l by providing $\bar{G}_l/\phi^l$. If $\bar{G}\phi^n = (G^o_{t+1}+\bar{G})\phi^{n'}$ and if n is included in Ω_{l+1} as well, placing her child in $l+1$ by providing $\bar{G}_{l+1}/\phi^{l+1}$ is also optimal.*
- *If $\bar{G} = 0$ and $G^o_{t+1} = 0$, then a skilled individual will place her child in a region in $l = 1, ..., m-1$ or m by providing $\frac{\bar{G}_l}{\phi^l}$.*
- *If $\bar{G} = 0$ and $G^o_{t+1} > 0$, then a skilled individual will place her child in region 1 by providing $\frac{\bar{G}_1}{\phi}$.*

(ii) Plan of $G^{O,H}_{t+2}$, the provision of family public goods in the case where the child becomes a skilled worker.

Whatever the parent's provision of family public goods including the one through the government, $G^o_{t+1}+\bar{G}$, is, a skilled individual will not privately provide family public goods in the case where her child becomes a skilled worker ($G^{O,H}_{t+2} = 0$).

2.1 Proof of Lemma 6: $\bar{G} > 0$

2.1.1 Lemma 6(i), $\bar{G} > 0$

Fix a region in $\{1, ..., m-1\}$ and let l^* denot it. And let n^* denote n that is in multiple Ω_l (Ω_{l^*} and Ω_l with $l > l^*$). Specifically, the city ($n = 0$) is included in all of Ω_l where $l \in \{1, ..., r\}$ (when $l^* = 1$, $n^* = 0$). Also, several regions in $\{1, ..., \bar{n}\}$ are included in $\Omega_{l^*} \cap \Omega_{l^*+1}$ with $l^* \in \{r, ..., m-1\}$.

First, I show that if a skilled worker who lives in the city with her parent ($n^* = 0$) and her parent does not privately provide family public goods ($G^o_{t+1} = 0$), placing her child in regions $1, ..., r$ in the case where her child becomes an unskilled worker is optimal. I also show that if a skilled worker lives in the city but the parent lives in region n^* other than the city ($n^* > 0$) and does not privately provide family public goods ($G^o_{t+1} = 0$), placing her child in region l^* and placing her child $l^* + 1$ ($l^* \in \{r, ..., m-1\}$) in the case where her child becomes an unskilled worker are optimal.

I start with the case of a skilled worker who lives in the city and has the parent in region n^* other than the city ($n^* > 0$).

From (9),

$$\begin{aligned}
&U(e, 0, \bar{G}, \bar{G}_r/\phi^r; H, 0, \bar{G}, \bar{G}) \\
=&U(e, 0, \bar{G}, \bar{G}_{r-1}/\phi^{r-1}; H, 0, \bar{G}, \bar{G}) \\
=&... \\
=&U(e, 0, \bar{G}, \bar{G}_1/\phi; H, 0, \bar{G}, \bar{G}).
\end{aligned} \tag{119}$$

For a skilled worker who lives in the city and has the parent in region n^* other than the city ($n^* > 0$), the value of family public goods provided by the parent in n^* but evaluated in the city, $\bar{G}\phi^{n^*}$, is lower than that of a skilled worker who has the parent in the city, $\bar{G}$. Applying Lemma 1(ii) to (119) yields:

$$\begin{aligned}
&U(e, 0, \bar{G}, \bar{G}_r/\phi^r; H, n^*, \bar{G}, \bar{G}) \\
>&U(e, 0, \bar{G}, \bar{G}_{r-1}/\phi^{r-1}; H, n^*, \bar{G}, \bar{G}) \\
>&... \\
>&U(e, 0, \bar{G}, \bar{G}_1/\phi; H, n^*, \bar{G}, \bar{G}).
\end{aligned} \tag{120}$$

Above is the case where a skilled worker places her child in a region in $\{1, ..., r\}$ in the case where her child becomes an unskilled worker. Turn to the case where she places her child in a region in $\{r, ..., l^*, l^* + 1\}$. From (9), for l^* in $\{r, ..., m-1\}$ and $h = 0, ..., l^* - r$, there exists n' that satisfies:

$$\begin{aligned}
&U(e, 0, \bar{G}, \bar{G}_{l^*-h}/\phi^{l^*-h}; H, n', \bar{G}, \bar{G}) \\
=&U(e, 0, \bar{G}, \bar{G}_{l^*+1-h}/\phi^{l^*+1-h}; H, n', \bar{G}, \bar{G}).
\end{aligned} \tag{121}$$

When $h = 0$, $n' = n^*$. The larger h is, the smaller n' is. Therefore, $1 \leq n' \leq n^*$. For a skilled worker in the city, the value of family public goods provided by the parent and evaluated in the city is higher when the parent lives in region n' than that when the parent lives in region n^* ($\bar{G}\phi^{n'} \geq \bar{G}\phi^{n^*}$). Applying Lemma 1(ii) to (121) yields:

$$\begin{aligned}
&U(e, 0, \bar{G}, \bar{G}_{l^*+1}/\phi^{l^*+1}; H, n^*, \bar{G}, \bar{G}) \\
=&U(e, 0, \bar{G}, \bar{G}_{l^*}/\phi^{l^*}; H, n^*, \bar{G}, \bar{G}) \\
>&U(e, 0, \bar{G}, \bar{G}_{l^*-1}/\phi^{l^*-1}; H, n^*, \bar{G}, \bar{G}) \\
>&... \\
>&U(e, 0, \bar{G}, \bar{G}_{r+1}/\phi; H, n^*, \bar{G}, \bar{G}) \\
>&U(e, 0, \bar{G}, \bar{G}_r; H, n^*, \bar{G}, \bar{G}).
\end{aligned} \tag{122}$$

Moreover, consider the utility when a skilled worker in the city places her child in a region in $\{l^* + 1, ..., m\}$ in the case where her child becomes an unskilled worker. From (9), for l^* in $\{r, ..., m - 1\}$ and $k = 1, ..., m - l^* - 1$ there exists n' in $\{n^* + 1, ..., \bar{n}\}$ that satisfies:

$$\begin{aligned}
&U(e, 0, \bar{G}, \bar{G}_{l^*+k}/\phi^{l^*+k}; H, n', \bar{G}, \bar{G}) \\
=&U(e, 0, \bar{G}, \bar{G}_{l^*+1+k}/\phi^{l^*+1+k}; H, n', \bar{G}, \bar{G}).
\end{aligned} \tag{123}$$

The larger k is, the large n' is. Since $n' > n^*$, $\bar{G}\phi^{n'} < \bar{G}\phi^{n^*}$. Applying Lemma 1(ii) to (123) yields:

$$\begin{aligned}
&U(e, 0, \bar{G}, \bar{G}_{l^*+1}/\phi^{l^*+1}; H, n^*, \bar{G}, \bar{G}) \\
>&U(e, 0, \bar{G}, \bar{G}_{l^*+2}/\phi^{l^*+2}; H, n^*, \bar{G}, \bar{G}) \\
>&... \\
>&U(e, 0, \bar{G}, \bar{G}_{m-1}/\phi^{m-1}; H, n^*, \bar{G}, \bar{G}) \\
>&U(e, 0, \bar{G}, \bar{G}_m/\phi^m; H, n^*, \bar{G}, \bar{G}).
\end{aligned} \tag{124}$$

Finally, consider the utility when a skilled worker in the city places her child in a region in $\{m, ..., \bar{n}\}$ in the case where her child becomes an unskilled worker. From (13), for l in $\{r, ..., m-1\}$ and $k = m-l, ..., \bar{n}-l-1$ there exist n in $\{1, ..., \bar{n}\}$ and n' in $\{1, ..., n\}$ that satisfy $n \in \Theta_{l+k-n}$ and thus:

$$\begin{aligned}
&U(e, n, \bar{G}, \bar{G}_{l+k}/\phi^{l+k-n}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}) \\
=&U(e, n, \bar{G}, \bar{G}_{l+1+k}/\phi^{l+1+k-n}; L, n', \bar{G}_n/\phi^{n-n'}, \bar{G}).
\end{aligned} \tag{125}$$

From (37), $w_L + \bar{G}_n < w_H$. Thus, applying Lemma 1(ii) and (v) to (125) yields:

$$\begin{aligned} &U(e,0,\bar{G},\bar{G}_m/\phi^m;H,n^*,\bar{G},\bar{G}) \\ >&U(e,0,\bar{G},\bar{G}_{m+1}/\phi^{m+1};H,n^*,\bar{G},\bar{G}) \\ >&... \\ >&U(e,0,\bar{G},\bar{G}_{\bar{n}-1}/\phi^{\bar{n}-1};H,n^*,\bar{G},\bar{G}) \\ >&U(e,0,\bar{G},\bar{G}_{\bar{n}}/\phi^{\bar{n}};H,n^*,\bar{G},\bar{G}). \end{aligned} \tag{126}$$

From (120), (122), (124) and (126), for a skilled worker whose parent lives in region n^* other than the city and does not privately provide family public goods ($G^o_{t+1} = 0$), placing her child in region l^* by providing $G^{O,L}_{t+2} + \bar{G} = \bar{G}^{l^*}/\phi^{l^*}$ and placing her child in region l^*+1 by providing $G^{O,L}_{t+2} + \bar{G} = \bar{G}^{l^*+1}/\phi^{l^*+1}$, where $n^* \in \Omega_{l^*} \cap \Omega_{l^*+1}$, are optimal. This holds for all $n^* \in \Omega_{l^*} \cap \Omega_{l^*+1}$ where $l^* \in \{r, ..., m-1\}$.

Next, I consider the case of a skilled worker whose parent also lives in the city and does not privately provide family public goods. For a skilled worker, the value of family public goods provided by the parent is higher if she has the parent in the city than otherwise ($\bar{G} > \bar{G}\phi^{n'}$). Therefore, applying Lemma 1(ii) to (121) yields:

$$\begin{aligned} &U(e,0,\bar{G},\bar{G}_{l^*+1}/\phi^{l^*+1};H,0,\bar{G},\bar{G}) \\ <&U(e,0,\bar{G},\bar{G}_{l^*}/\phi^{l^*};H,0,\bar{G},\bar{G}) \\ <&U(e,0,\bar{G},\bar{G}_{l^*-1}/\phi^{l^*-1};H,0,\bar{G},\bar{G}) \\ <&... \\ <&U(e,0,\bar{G},\bar{G}_{r+1}/\phi;H,0,\bar{G},\bar{G}) \\ <&U(e,0,\bar{G},\bar{G}_r;H,0,\bar{G},\bar{G}). \end{aligned} \tag{127}$$

This is the inverse of (122). In contrast, (124) and (126) hold if n^* in these equations is replaced by 0. (Inequalities in (124) and (126) become more prominent). Therefore, from (119) and (127), and (124) and (126) where n^* is replaced by 0, for a skilled worker whose parent also lives in the city and does not privately provide family public goods ($G^o_{t+1} = 0$), placing her child in regions $\{1, ..., r\}$ in the case where her child becomes an unskilled worker is optimal.

Consider the case where the parent lives in a region other than n^* and/or privately provides a positive amount of family public goods.

When the parent increases the private provision of family public goods (G^o_{t+1} is larger) and/or lives in region n' closer to the city (n' is smaller), the value of family public goods provided by the parent in region n' but evaluated in the city, $(G^o_{t+1}+\bar{G})\phi^{n'}$, increases. From Lemma 1(ii), the larger $G^{O,L}_{t+2}+\bar{G}$, the more dramatically $U(e,0,\bar{G},G^{O,L}_{t+2}+\bar{G};H,n',G^o_{t+1}+\bar{G},\bar{G})$ rises when the value of family public goods provided by the parent and evaluated in the city rises. Therefore, if n', $\bar{G}$ and G^o_{t+1} satisfy $\bar{G}\phi^n \leq (G^o_{t+1}+\bar{G})\phi^{n'} < \bar{G}\phi^{n-1}$ where $n \in \Omega_l$, for a skilled worker placing her child in region l by providing $\bar{G}_l/\phi^l$ in the case where her child becomes an unskilled worker is optimal. Only when $\bar{G}\phi^n = (G^o_{t+1}+\bar{G})\phi^{n'}$ and $n \in \Omega_l \cap \Omega_{l+1}$, placing her child in region $l+1$ is also optimal. If $\bar{G} = (G^o_{t+1}+\bar{G})\phi^{n'}$, placing her child in regions $1, ..., r$ in the case where her child becomes an unskilled worker is optimal. If $\bar{G} < (G^o_{t+1}+\bar{G})\phi^{n'}$, a skilled worker will place her child in region 1.

2.1.2 Lemma 6(i), $\bar{G} = 0$

First, consider the case where $G^o_{t+1} = 0$. From (9),

$$\begin{aligned}&U(e,0,0,\bar{G}_m/\phi^m;H,n',0,0)\\=&U(e,0,0,\bar{G}_{m-1}/\phi^{m-1};H,n',0,0)\\=&...\\=&U(e,0,0,\bar{G}_1/\phi;H,n',0,0).\end{aligned} \tag{128}$$

When $\bar{G} = 0$ and $G^o_{t+1} = 0$, the value of family public goods is zero wherever an individual and her parent live. Therefore, (119) holds even if $n' \neq 0$.

From (13), for $k = 0, ..., \bar{n}-m-1$ there exists n in $\{0, ..., \bar{n}\}$ that satisfies $n \in \Theta_{m+k-n}$ and thus:

$$\begin{aligned}&U(e,n,\bar{G}_{m+k}/\phi^{m+k-n},0;L,n',\bar{G}_n/\phi^{n-n'},0)\\=&U(e,n,\bar{G}_{m+k+1}/\phi^{m+k+1-n},0;L,n',\bar{G}_n/\phi^{n-n'},0).\end{aligned} \tag{129}$$

From (37), $w_L + \bar{G}_n < w_H$. Thus, applying Lemma 1(ii) and (v) to (129) yields:

$$\begin{aligned}&U(e,0,0,\bar{G}_m/\phi^m;H,n',0,0)\\>&U(e,0,0,\bar{G}_{m+1}/\phi^{m+1};H,n',0,0)\\>&...\\>&U(e,0,0,\bar{G}_{\bar{n}-1}/\phi^{\bar{n}-1};H,n',0,0)\\>&U(e,0,0,\bar{G}_{\bar{n}}/\phi^{\bar{n}};H,n',0,0).\end{aligned} \tag{130}$$

From (128) and (130), for a skilled worker whose parent does not privately provide family public goods ($G^o_{t+1} = 0$), placing her child in regions $\{1, ..., m\}$ in the case where her child becomes an unskilled worker is optimal.

Consider the case where the parent privately provides a positive amount of family public goods. From Lemma 1(ii), the larger $G^{O,L}_{t+2}$, the more dramatically $U(e, 0, 0, G^{O,L}_{t+2}; H, n',$

$G^o_{t+1}, 0)$ rises when G^o_{t+1} increases. Therefore, from (128), for $G^o_{t+1} > 0$ it holds that:

$$\begin{aligned} &U(e, 0, 0, \bar{G}_m/\phi^m; H, n', G^o_{t+1}, 0) \\ <&U(e, 0, 0, \bar{G}_{m-1}/\phi^{m-1}; H, n', G^o_{t+1}, 0) \\ <&... \\ <&U(e, 0, 0, \bar{G}_1/\phi; H, n', G^o_{t+1}, 0), \end{aligned} \tag{131}$$

and inequalities in (130) becomes more prominent. Therefore, a skilled worker will place her child in region 1.

2.1.3 Lemma 6(ii)

When a skilled worker lives in the city, gives a higher education for her child and plans $C^{O,H}_{t+2} = 0$, her child will also live in the city in the case where her child also becomes a skilled worker. The larger $C^{O,H}_{t+2}$ is, the more her child is eager to live in the city, yet the lower the skilled worker's utility is. Therefore, the skilled worker will choose $C^{O,H}_{t+2} = 0$.

2.2 Proof of Proposition, Case H(i): $n' \in \{0, ..., \bar{n}\}$

Compare a skilled worker's maximum expected lifetime utility conditional on that where she lives is limited in regions other than the city, $n \in \{1, ..., N\}$ and the one conditional on that she lives in the city, $n = 0$. Suppose that n' and $G^o_{t+1} + \bar{G}$ satisfy

$$U(e, 0, \bar{G}, \bar{G}^{l'}/\phi^{l'}; H, n', G^o_{t+1} + \bar{G}, \bar{G}) \geq U(edu, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', G^o_{t+1} + \bar{G}, \bar{G}), \tag{132}$$

where $n' \in \{0, ..., \bar{n}\}$, $n \in \{0, ..., N\}$, $l' \in \{0, ..., m\}$, $l \in \{0, ..., m\}$ and $edu \in \{0, e\}$, and the relationship between $(edu, \bar{G}_{n+l}/\phi^l, n)$ and $(n', G^o_{t+1} + \bar{G})$ is as described in Case L(i), and the relationship between $\bar{G}_{l'}/\phi^{l'}$ and $(n', G^o_{t+1} + \bar{G})$ is as described in Lemma 6. Then a skilled worker will live and work in the city. This is because the right-hand side is the maximum expected lifetime utility conditional on that where a skilled worker lives is limited in regions other than the city. In this case, her skill does not contribute and thus her wage is w_L. Thus, a skilled worker's choice depends on her parent's choice of n' and G^o_{t+1}, exactly in the same manner as an unskilled worker's choice depends on n' and G^o_{t+1}. Thus, $U(edu, n, \bar{G}, \bar{G}_{n+l}/\phi^l; H, n', G^o_{t+1} + \bar{G}, \bar{G}) = U(edu, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. The left-hand side of (132) is a skilled worker's maximum expected lifetime utility when she lives in the city. Therefore, if (132) holds, a skilled worker prefers to live in the city.

Below, I show that if the private provision of family public goods by her parent is zero or small, (132) holds and a skilled worker will live in the city.

Consider a skilled worker's maximum expected lifetime utility conditional on that where she lives is limited in regions other than the city. When G^o_{t+1} and n' satisfy $(G^o_{t+1} +$

$\bar{G})\phi^{\bar{n}-n'} = \bar{G}$, from the discussions of an unskilled worker's optimal behavior in Case L(i), she will live in region $\bar{n}$, give a higher education for her child, and place her child in region $\bar{n}$ when her child becomes an unskilled worker, and the maximum expected lifetime utility is

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}) = U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}). \tag{133}$$

In contrast, when G^o_{t+1} and n' satisfy $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} < \bar{G}$, from the discussions of an unskilled worker's optimal behavior when $0 \leq (G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} < \bar{G}$ in Case L(i), she will live in region $\bar{n}$ or $\bar{n}+1$. Also, she may opt not to give a higher education for her child ($edu = 0$). In any case, the maximum expected lifetime utility becomes lower than $U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G})$:

$$U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G}) \geq U(edu, \bar{n}, \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G}). \tag{134}$$

That is, $U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G})$ is the highest possible value of maximum expected lifetime utility when a skiled worker lives in a region other than the city and when $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} \leq \bar{G}$.

Next, consider a skilled worker's maximum expected lifetime utility conditional on that she lives in the city. The value of the family public goods provided by the parent in region n' and evaluated in the city is the lowest when $n' = \bar{n}$ and $G^o_{t+1} = 0$. Therefore, it holds that:

$$U(e, 0, \bar{G}, \bar{G}^{l'}/\phi^{l'}; H, n', G^o_{t+1} + \bar{G}, \bar{G}) \geq U(e, 0, \bar{G}, \bar{G}^m/\phi^m; H, \bar{n}, \bar{G}, \bar{G}), \tag{135}$$

where $n' \in \{0, ..., \bar{n}\}$ and $l' \in \{0, ..., m\}$. That is, $U(e, 0, \bar{G}, \bar{G}^m/\phi^m; H, \bar{n}, \bar{G}, \bar{G})$ is the lowest possible value of maximum expected lifetime utility when she lives in the city.

(8) states that $U(e, \bar{n}, \bar{G}, \bar{G}; L, \bar{n}, \bar{G}, \bar{G})$, the highest possible value of maximum expected lifetime utility when a skilled worker lives in a region other than the city and when $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} \leq \bar{G}$, is lower than $U(e, 0, \bar{G}, \bar{G}^m/\phi^m; H, \bar{n}, \bar{G}, \bar{G})$, the lowest possible value of maximum expected lifetime utility when she lives in the city. From this fact, (133), (134) and (135), when $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} \leq \bar{G}$, (132) holds with $n' \in \{0, ..., \bar{n}\}$, $n = \bar{n}$, $l' \in \{0, ..., m\}$, $l \in \{0, ..., m\}$, and $edu =\in \{0, e\}$. Thus, a skilled worker will live in the city when $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} \leq \bar{G}$.

If $G^{\#}_{n'}$ and n that satisfy (34) exist, (34) is written as

$$U(e, 0, \bar{G}, \bar{G}^{l'}/\phi^{l'}; H, n', G^{\#}_{n'}, \bar{G}) = U(edu, n, \bar{G}, \bar{G}_{n+l}/\phi^l; L, n', G^{\#}_{n'}, \bar{G}). \tag{136}$$

If (8) holds with equality, (132) with $n' = \bar{n}$, $n = \bar{n}$, $l' = m$, $l = 0$ and $edu = e$ also holds with bind, and, thus, from (136), $G^{\#}_{\bar{n}} = \bar{G}$. In contrast, if (8) holds with inequality, or if $n' \neq \bar{n}$, (132) with $n' \in \{0, ..., \bar{n}\}$, $n = \bar{n}$, $l' \in \{0, ..., m\}$, $l \in \{0, ..., m\}$, and $edu =\in \{0, e\}$, holds with inequality, as long as $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} \leq \bar{G}$, as we saw. However, the larger

G^o_{t+1} is, the more hardly (132) holds, as when $G^o_{t+1} + \bar{G}$ increases, the right-hand side rises more dramatically than the left-hand side. Also, in the right-hand side, n tends to be smaller and l tends to be larger, depending on G^o_{t+1}. That is because region n is closer to region n' where the parent lives and provides it, than the city is. Therefore, if (8) holds with inequality, $G^{\#}_{n'}$ is strictly larger than $G^o_{t+1} + \bar{G}$ that satisfies $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} = \bar{G}$. Thus, $G^{\#}_{\bar{n}} \geq \bar{G}$ and $G^{\#}_{n'} > \bar{G}$ for $n' \in \{0, ..., \bar{n} - 1\}$. Therefore, if the parent's private provision of family public goods, G^o_{t+1}, is zero or very small, a skilled worker will live in the city. Where she plans to place her child in the case where her child becomes an unskilled worker depends on $G^o_{t+1} + \bar{G}$ as Lemma 6 indicates.

When $G^o_{t+1} + \bar{G}$ is larger than $G^{\#}_{n'}$, the choices of n in $\{n', ..., N\}$, edu in $\{0, e\}$, and $G^{O,L}_{t+2}$ depend on n' and $G_{t+1} + \bar{G}$ that is larger than $G^{\#}_{n'}$ in the same manner as an uskilled worker's choice depend on them in Case L(i). Since $G^{\#}_{n'} > \bar{G}$ for $n' \in \{0, ..., \bar{n} - 1\}$, $(G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'}$ is not smaller than $\bar{G}$. Therefore, behaviors in Case L(i) where $0 < (G^o_{t+1} + \bar{G})\phi^{\bar{n}-n'} < \bar{G}$ are not optimal, and thus n is in $\{n', ..., \bar{n}\}$ and $edu = e$ are as in (3).

Thus, the behaviors described in Case H(i) are optimal.

2.3 Proof of Proposition, Case H(ii): $n' \in \{\bar{n} + 1, ...\}$

I compare the following three: the expected lifetime utility conditional on that a skilled worker does not either live in the city or give a higher education for her child ($n \neq 0$ and $edu = 0$), the one conditional on that she does not live in the city but she gives a higher education ($n \neq 0$ and $edu = e$), and the one conditional on that she lives in the city and gives a higher eduation ($n = 0$ and $edu = e$).

From the discussion in the proof of proposition of Case L(ii), given that a skilled worker does not either live in the city or give a higher education for her child ($n \neq 0$ and $edu = 0$), her optimal choice is to stay in the region where her parent lives, $n = n'$, and not to provide family public goods when old, $G^{O,L}_{t+2} = 0$. Thus the maximum expected lifetime utility conditional on that $n \neq 0$ and $edu = 0$ is $U(0, n', \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$. For all n' in $\{\bar{n} + 1, ..., N\}$, this is constant.

First, compare $U(0, n', \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ and the expected lifetime utility conditional on that a skilled worker lives in the city, $U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^o_{t+1} + \bar{G}, \bar{G})$. When her parent in region n' increases $G^o_{t+1} + \bar{G}$, an individual's expected lifetime utility when she lives in the city, $U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^o_{t+1} + \bar{G}, \bar{G})$ rises less drastically than that when she lives in region n' and does not give a higher eduation for her child, $U(0, n', \bar{G}, \bar{G}; L, n', G^o_{t+1} + \bar{G}, \bar{G})$ does, due to the distance from n' to the city. Therefore, if $G^o_{t+1} + \bar{G}$ is larger than $G^{\flat}_{n'}$ that satisfies (36), a skilled worker will not move to the city but stay in region n'. The farther the hometown is from the city (the larger n' is), the smaller the threshold $G^{\flat}_{n'}$ is. That is because the larger n' is, the smaller

$U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ is, whereas $U(0, n', \bar{G}, \bar{G}; L, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ is constant, and thus with smaller $G^{o}_{t+1} + \bar{G}$ these two are equalized. Thus, even when $G^{o}_{t+1} + \bar{G}$ is small, she hesitates to move to the city.

Next, compare $U(0, n', \bar{G}, \bar{G}; L, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ and the expected lifetime utility conditional on that a skilled worker does not lives in the city but she gives a higher eduation for her child, ($n \neq 0$ and $edu = e$). In these cases, a skilled worker does not live in the city. When a skilled worker lives in a region other than the city, her skill does not contribute and thus her wage is w_L. Therefore, her behavior depends on her parent's choice exactly in the same manner as an unskilled worker's choice depends on it as we saw in the discussions in the proof of proposition of Case L(ii). That is, if $G^{*}_{n'}$ exists and if $G^{o}_{t+1} + \bar{G}$ exceeds it, an individual will live in a region between n' and $\bar{n}$, and give a higher education for her child. In contrast, if $G^{o}_{t+1} + \bar{G}$ is smaller than $G^{*}_{n'}$, or if $G^{*}_{n'}$ does not exist, an individual will live in n' and will not give a higher education for her child.

Summarizing the discussions thus far, consider the behavior of a skilled worker in the case where $G^{\flat}_{n'} \leq G^{*}_{n'}$. If $\bar{G} \leq G^{o}_{t+1} + \bar{G} \leq G^{\flat}_{n'}$, then a skilled worker will move to the city. In contrast, $G^{o}_{t+1} + \bar{G}$ becomes so large as to satisfy $G^{\flat}_{n'} \leq G^{o}_{t+1} + \bar{G}$, it attracts a skilled worker in her hometown. If it is smaller than $G^{*}_{n'}$, she will not give a higher education for her child. In contrast, if it exceeds $G^{*}_{n'}$, she will give a higher education for her child and live in a region in $\{\bar{n}, ..., n'\}$.

Next, consider the behavior of a skilled worker when $G^{*}_{n'} < G^{\flat}_{n'}$. When $G^{o}_{t+1} + \bar{G} \leq G^{*}_{n'}$, $G^{o}_{t+1} + \bar{G} < G^{\flat}_{n'}$. Therefore, $U(0, n', \bar{G}, \bar{G}; L, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ is lower than or equal to the right-hand side of (31), yet it is still strictly lower than $U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^{o}_{t+1} + \bar{G}, \bar{G})$. Therefore, the right-hand side of (31) is also strictly lower than $U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^{o}_{t+1} + \bar{G}, \bar{G})$. Thus, an individual will move to the city and give a higher education for her child.

When her parent in region n' increases $G^{o}_{t+1} + \bar{G}$ and it exceeds $G^{*}_{n'}$, an individual continues to give a higher education for her child. Where she lives depends on whether $G^{o}_{t+1} + \bar{G}$ is larger than $G^{\#}_{n'}$ that satisfies (35). When her parent in region n' increases $G^{o}_{t+1} + \bar{G}$, an individual's expected lifetime utility when she lives in the city, $U(e, 0, \bar{G}, G^{O,L}_{t+2} + \bar{G}; H, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ rises less drastically than $U(e, n, \bar{G}, \bar{G}; L, n', G^{o}_{t+1} + \bar{G}, \bar{G})$ with $n \in \{\bar{n} + 1, ...\}$ does, due to the distance from n' to the city. If $G^{o}_{t+1} + \bar{G}$ is larger than $G^{\#}_{n'}$, a skilled worker will be attracted in her hometown or a region close to it, and give a higher education for her child.

Appendix C: Functional and Numerical Examples

Specify utility and related functions as:

$$u(C^y_{t+1}) = \ln(C^y_{t+1}),$$

$$v\left(G^y_{t+1} + \delta\left(n_t - n_{t+1}\right) G^o_{t+1}\right) = \ln\left(G^y_{t+1} + \phi^{|n_t - n_{t+1}|} G^o_{t+1}\right),$$

$$c(P_n) = 2P_n.$$

$$w(Ph) = 16.88Ph^{0.1}.$$

Set parameters as $\alpha = 1$, $\phi = 0.988$, $\mu = 0.5$, $\rho = r = 1$, $e = 0.2$, and $wL = 10$. Total number of regions is $\bar{N} = 65$ and the total population is unity ($P = 1$).

All figures and table uses the example of monocentric equilibrium without social security ($\bar{G} = 0$). In the equilibrium, $\bar{n} = 13$. That is, in total, in 27 regions (city plus $2\bar{n}$) investment in higher education takes place. The population in the city is 0.336.

Figures 6 and 7 show the examples of multicentric equilibrium without social security and that with social security ($\bar{G} = 0.24$). In the equilibrium without social security, $\bar{n} = 9$. Therefore, in total, in 38 regions (two cities and $4\bar{n}$) investment in higher education takes place. The population in each city is 0.192. In the equilibrium with social security, $\bar{n} = 11$. Therefore, in total, in 46 regions (two cities and $4\bar{n}$) investment in higher education takes place. The population in each city is 0.222.

References

[1] Aarland, K., Davis, J.C., Henderson, J. V., Ono, Y. 2007. Spatial Organization of Firms: The Decision to Split Production and Administration. RAND Journal of Economics 38(2): 480-494.

[2] Arrow, K. 1962. The Economic Implications of Learning by Doing. Review of Economic Studies 29(3), 155-173.

[3] Banerjee, A.V., Newman, A. F. 1993. Occupational Choice and the Process of Development. Journal of Political Economy 101(2), 274-298.

[4] Barro, R. J., Sala-i-Martin, X. 1992. Convergence. Journal of Political Economy 100(2), 223-251.

[5] Bernheim, B. D, Shleifer, A., Summers, L. H. 1985. The Strategic Bequest Motive. Journal of Political Economy 93(6), 1045-1076.

[6] Burzynski, M., Docquier. C., Docquier, F. 2020. Geography of Skills and Global Inequality. Journal of Development Economics 142, 1023333.

[7] Camargo, B., Stein, G. 2022. Credit Constraints and Human Capital Policies. Journal of Public Economics 208, 104624.

[8] Card, D. 2009. Immigration and Inequality. American Economic Review 99(2), 1-21.

[9] Caucutt, E. M., Lochner, L. 2020. Early and Late Human Capital Investments, Borrowing Constraints, and the Family. Journal of Political Economy 128(3): 1065-1147.

[10] Davis, J.C., Henderson, J. V. 2008. The agglomeration of headquarters. Regional Science and Urban Economics 38(5): 445-460.

[11] Docquier, F., Rapoport, H. 2012. Globalization, Brain Drain, and Development. Journal of Economic Literature 50(3): 681-730.

[12] Duranton, G., Puga, D. 2005. From sectoral to functional urban specialisation. Journal of Urban Economics 57(2): 343-370.

[13] Ehrlich, I., Lui, F. T. 1991. Intergenerational Trade, Longevity, and Economic Growth. Journal of Political Economy 99(5), 1029-1059.

[14] Ellison, G., Glaeser, E. 1997. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. Journal of Political Economy 105(5), 889-927.

[15] Forslid, R., Ottaviano, G. I. P. 2014. An Analytically Solvable Core-Periphery Model. Journal of Economic Geography 3(3): 229-240.

[16] Fujita, M., Krugman, P. Venables, A. J., 1999. The Spatial Economy: Cities, Regions, and International Trade, MIT Press.

[17] Gabe, T. M., Abel, J. R. 2012. Specialized Knowledge and the Geographic Concentration of Occupation. Journal of Economic Geography 12(2), 435-453.

[18] Galor, O., Moav, O., 2004. From Physical to Human Capital Accumulation: Inequality in the Process of Development. Review of Economic Studies 71(4): 1001–1026.

[19] Galor, O., Zeira, J. 1993. Income Distribution and Macroeconomics. Review of Economic Studies 60(1): 35–52.

[20] Ganong, P., Shoag, D. 2017. Why Has Regional Income Convergence in the U.S. declined? Journal of Urban Economics 102: 76-90.

[21] Glomm, G., Kaganovich, M. 2008. Social Security, Public Education and the Growth-Inequality Relationship, European Economic Review 52(6): 1009-1034.

[22] Glomm, G., Ravikumar, B. 1992. Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality. Journal of Political Economy 100(4): 818-834.

[23] Greenstone, M., Hornbeck. R., Moretti, E. 2010. Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. Journal of Political Economy 118(3): 536-598.

[24] Grossman, G. M., Helpman, E. 1991. Innovation and Growth in the Global Economy. MIT Press.

[25] Gyourko, J., Mayer, C., Sinai, T. 2013. Superstar Cities. American Economic Journal: Economic Policy 5(4): 167-199.

[26] Harris, J. R., Todaro, M. P. 1970. Migration, Unemployment and Development: A Two Sector Analysis. American Economic Review 60(1): 126-142.

[27] Jaffe, A. B. 1986. Technological Opportunity and Spillovers of R& D: Evidence from Firms' Patents, Profits, and Market Value. American Economic Review 76 (5):984–1001.

[28] Jaffe, A. B., Trajtenberg, M., Henderson, R. M. 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citation. Quarterly Journal of Economics 108(3): 577-598.

[29] Jensen, R. 2012. Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India. Quarterly Journal of Economics 127(2): 753-792.

[30] Kaplan, G., Schulhofer-Wohl, S. 2017. Understanding the long-run decline in interstate migration. International Economic Review 58(1): 57-94.

[31] Keller, W. 2002. Geographic Localization of International Technology Diffusion, American Economic Review 92(1): 120-142.

[32] Kerr, S.P., Kerr, W., Özden, C., Parsons, C. 2016. Global Talent Flows. Journal of Economic Perspectives 30(4): 83-106.

[33] Konrad, K. A., Künemund, H., Lommerud, K. E., Robledo, J. R. 2002. Geography and Family. American Economic Review 92(4): 981-998.

[34] Krugman, P. 1991. Increasing Returns and Economic Geography. Journal of Political Economy 99(3): 483-499.

[35] Krugman, P., Venables, A. J. 1995. Globalization and the Inequality of Nations. Quarterly Journal of Economics 110(4): 857–880.

[36] Lewis, W. A. 1954. Economic development with unlimited supplies of labour. The Manchester School 22(2):115-227.

[37] Lucas, R. E. Jr. 1988. On the Mechanics of Economic Development. Journal of Monetary Economics 22(1): 3-42.

[38] Lychagin, S., Pinkse, J., Slade, M. E., Van Reenen, J. 2016. Spillovers in Space: Does Geography Matter? Journal of Industrial Economics 64(2): 295-335.

[39] Mookherjee, d., Ray, D. 2003. Persistent Inequality. Review of Economic Studies 70(2):369-393.

[40] Moretti, E. 2004. Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions. American Economic Review 94 (3): 656–90.

[41] Orlando, M. 2004. Measuring Spillovers from Industrial R&D: On the Importance of Geographic and Technological Proximity, RAND Journal of Economics 35(4): 777-786.

[42] Ottaviano, G. I. P., Peri, G. 2012. Rethinking the Effect of Immigration on Wages. Journal of European Economic Association 10(1), 152-197.

[43] Restuccia, D., Urrutia, C. 2004. Intergenerational Persistence of Earnings: The Role of Early and College Education. American Economic Review 94(5): 1354-1378.

[44] Roland Bénabou, R. 1996. Inequality and Growth, NBER Macroeconomics Annual 11: 11-74.

[45] Romer, P. M. 1986. Increasing Returns and Long-Run Growth. Journal of Political Economy 94(5): 1002-1037.

[46] Uzawa. H. 1965. Optimum Technical Change in An Aggregative Model of Economic Growth. International Economic Review 6(1): 18-31.

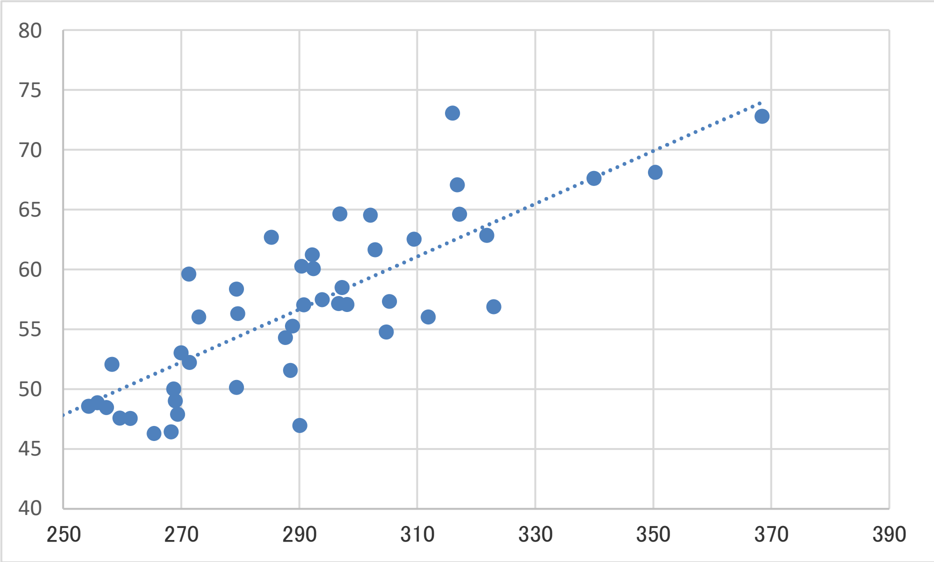

Figure 1 (a)

College enrollment rates (%) and average wages across prefectures

Correlation coef. =0.809

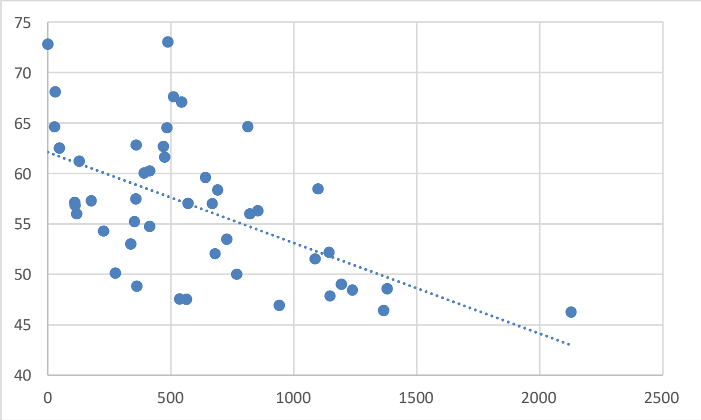

Figure 1 (b)

College enrollment rates (%) and the distances to Tokyo (km) across prefectures

Correlation coef. =-0.563

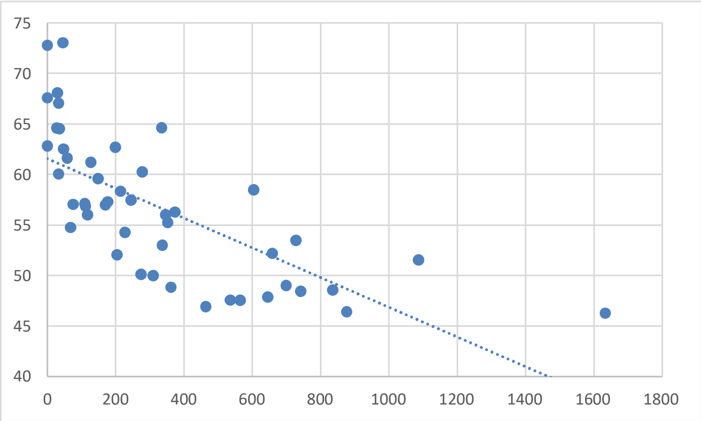

Figure 1 (c)

College enrollment rates (%) and the distances to the nearest three big cities across prefectures

Correlation coef. =-0.702

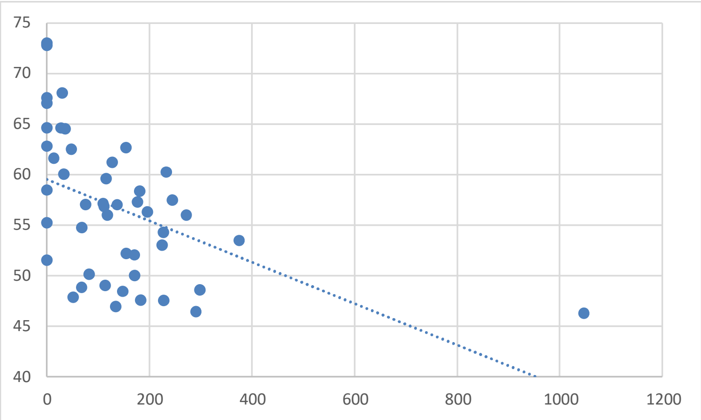

Figure 1 (d)

College enrollment rates (%) and the distances to the nearest nine big cities across prefectures

Correlation coef. =-0.486

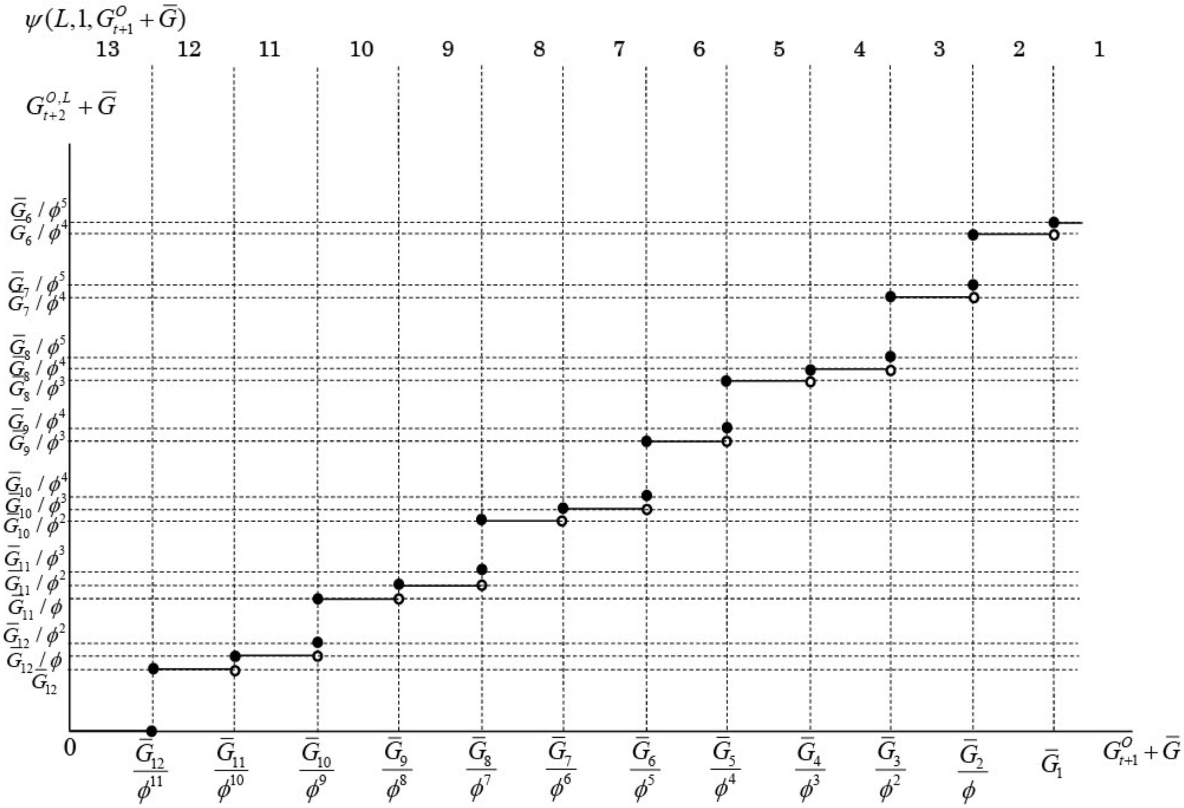

Figure 2: Optimal correspondences

The optimal choices of n and $G^{O,L}_{t+2}+\bar{G}$ by an unskilled worker when the parent lives in region one ($n'=1$) and provides $G^o_{t+1}+\bar{G}$. The value of $G^{O,L}_{t+2}+\bar{G}$ at which $0 < Q(\bar{G}, G^{O,L}_{t+2}+\bar{G}|L,1,G^o_{t+1}+\bar{G}) \leq 1$. The optimal choice of edu is e when $0 \leq G^o_{t+1}+\bar{G}$

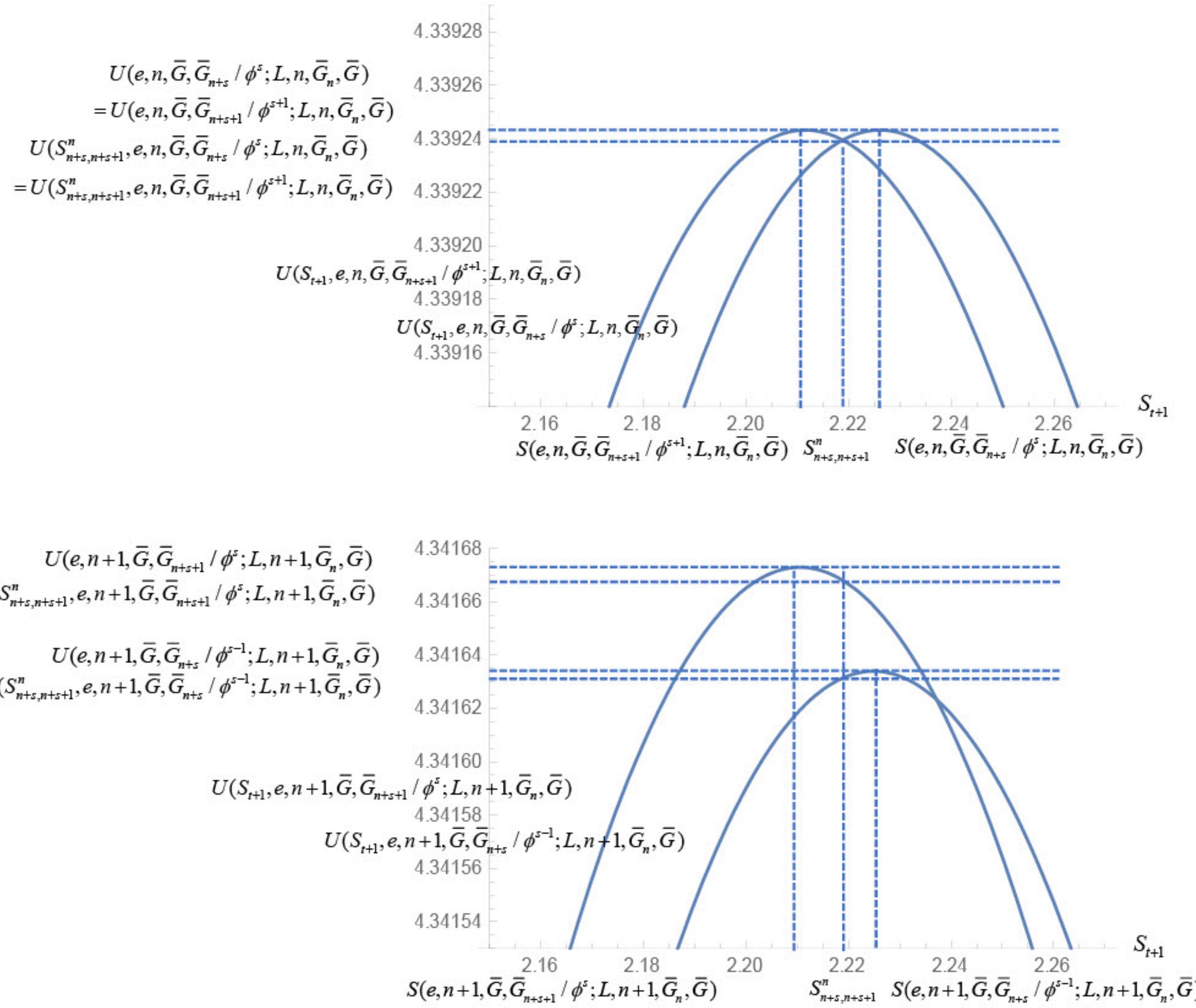

Figure 3: Maximum conditional expected lifetime utility

An individual's expected lifetime utility $\tilde{U}(S_{t+1},e,n,\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',\bar{G}_n/\phi^{n-n'},\bar{G})$, and the maximum conditional expected lifetime utility $U(e,n,\bar{G},G^{O,L}_{t+2}+\bar{G};L,n',\bar{G}_n/\phi^{n-n'},\bar{G})$ where $G^{O,L}_{t+2}+\bar{G}=\bar{G}_{n+l-1}/\phi^{l-1}$, $\bar{G}_{n+l}/\phi^{l}$, $\bar{G}_{n+l+1}/\phi^{l+1}$, and $\bar{G}_{n+l+2}/\phi^{l+2}$. In the example in this Figure, $n=3$ and $l=4$. If she resides in a region that is closer to the city such as $n-1$ and her parent provides larger amount of public goods such as $\bar{G}_{n-1}/\phi^{n-1-n'}$, the maximum expected lifetime utility when she places her child in region $n+l$ with $\bar{G}_{n+l}/\phi^{l+1}$ is equal to the one when she places her child in closer region $n+l-1$ with larger $\bar{G}_{n+l-1}/\phi^{l}$. However, in the case where an individual resides in region n and her parent provides family public goods in a smaller amount of $\bar{G}_n/\phi^{n-n'}$, the former becomes strictly higher than the latter. By moving from region $n-1$ to region n and thus approaching her child, the value of the family public goods provided by her child G^{y}_{t+2} is more dramatically amplified in the case where her child lives in a more distant region $n+l$ than in the case where her child lives in a closer region $n+l-1$, since in the former case the amount of G^{y}_{t+2} her child provides is larger. In addition to this, the value of family public goods provided by her parent, $\bar{G}_n$, is lower than $\bar{G}_{n-1}$. Therefore, when she plans $G^{O,L}_{t+2}+\bar{G}$ that is larger than $\bar{G}_{n+l-1}/\phi^{l-1}$, the total provision of family public goods is excessive. From a similar logic, inversely, when she plans $G^{O,L}_{t+2}+\bar{G}$ to be equal to or smaller than $\bar{G}_{n+l+2}/\phi^{l+2}$, the total provision of family public goods is exceedingly small relative to the consumption of private goods.

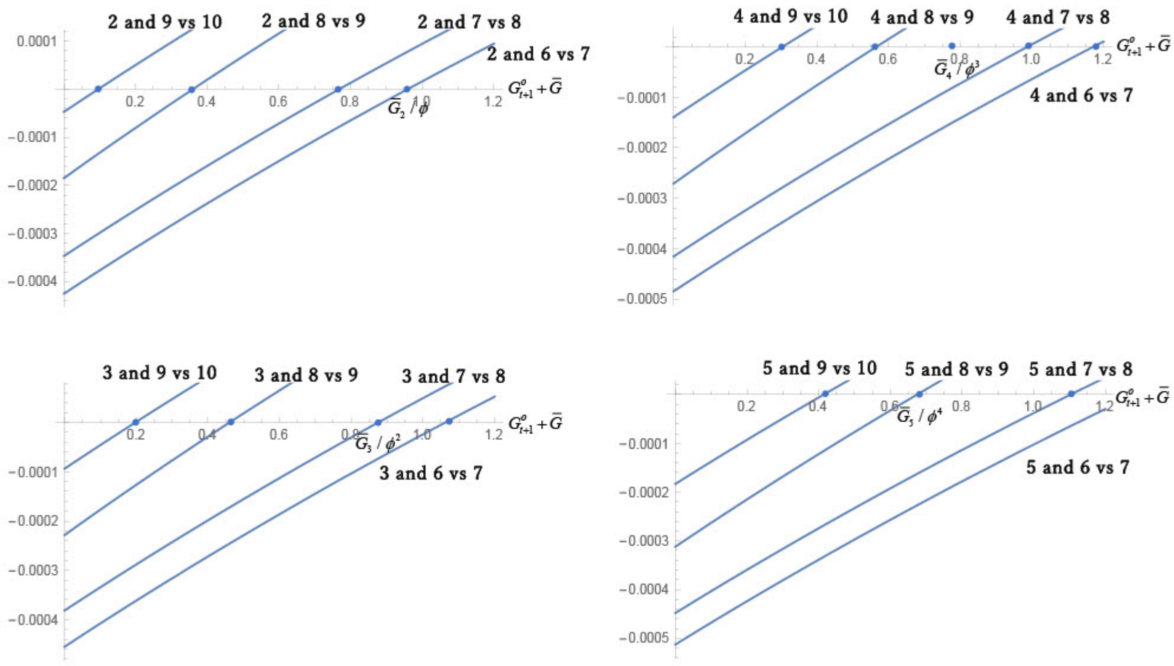

Figure 4: Maximum conditional expected lifetime utility

The graph "n and $n+s$ vs $n+s+1$" means an individual's maximum conditional expected lifetime utility when she lives in region n and chooses $G^{O,L}_{t+2}+\bar{G}=\bar{G}_{n+s}/\phi^s$ minus that when she lives in region n and chooses $G^{O,L}_{t+2}+\bar{G}=\bar{G}_{n+s+1}/\phi^{s+1}$, $U(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,G^o_{t+1}+\bar{G},\bar{G})-U(e,n,\bar{G},\bar{G}_{n+s}/\phi^s;L,n,G^o_{t+1}+\bar{G},\bar{G})$. The parent lives in region 1. When $G^o_{t+1}+\bar{G}$ provided by her parent increases, every $U(e,n,\bar{G},G^{O,L}_{t+2}+\bar{G};L,n,G^o_{t+1}+\bar{G},\bar{G})$ rises, but more importantly, the one with larger $G^{O,L}_{t+2}+\bar{G}$ rises more drastically. However, the size of $G^{O,L}_{t+2}+\bar{G}$ provided by her parent with which for an individual in region n, placing her child in region $n+s$ and placing her child in closer region $n+s-1$ yield indifferent outcomes is bigger than that with which, for an individual in region $n-1$, these two yield indifferent outcomes.

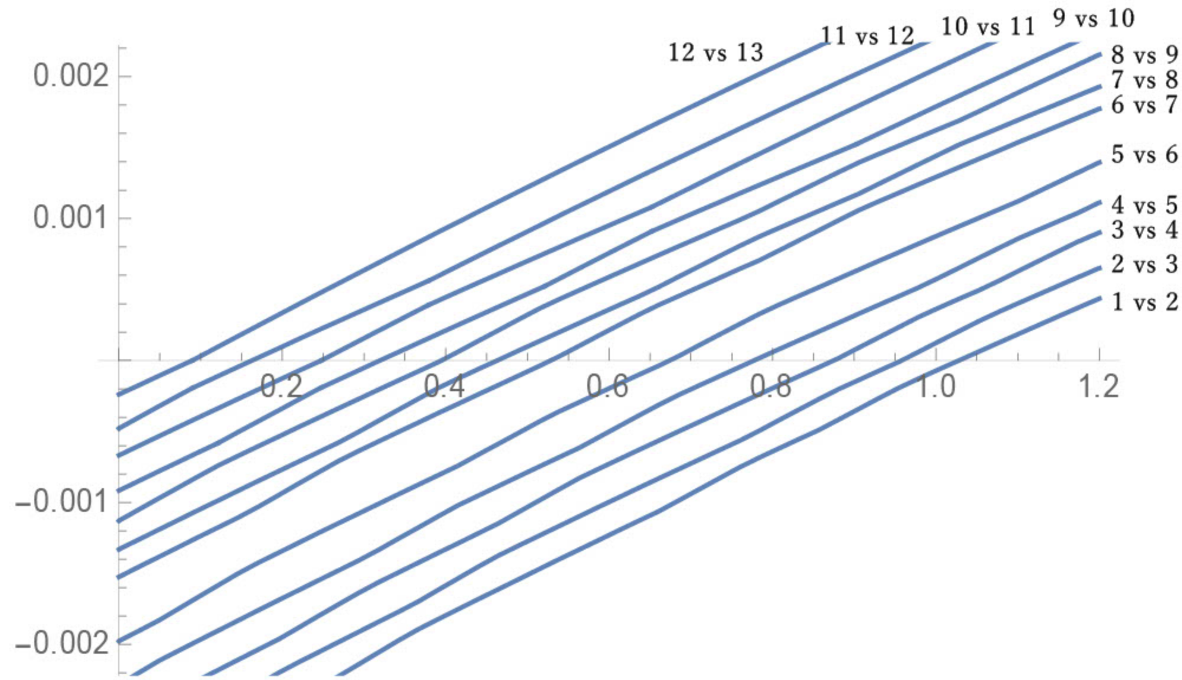

Figure 5: Decision on where to live

The graph "n vs $n+1$" means an individual's expected lifetime utility when she lives in region $n+1$ minus that when she lives in region n for $n \in \{1, ..., \bar{n}-1\}$. In the example in this figure, $\bar{n} = 13$. The parent lives in region 1. If regions have the same population, all the graphs locate above the horizontal line. If a region closer to the city has a larger population to the extent that it just cancels out the higher value of family public goods provided by the child in the city when the child becomes a skilled worker, all the graphs start from the origin. In reality, as the population density in a region closer to the city is higher to the extent that it more than just cancels out the higher value of family public goods provided by the child in the city, the graph of a region that is closer to the city is located more downwardly, and, thus, intersects with the horizontal axis at larger $G^o_{t+1} + \bar{G}$. Unless her parent provides family public goods in the amount that is large enough to overcome the prominently large population, an individual is not willing to live closer to the region one where her parent lives.

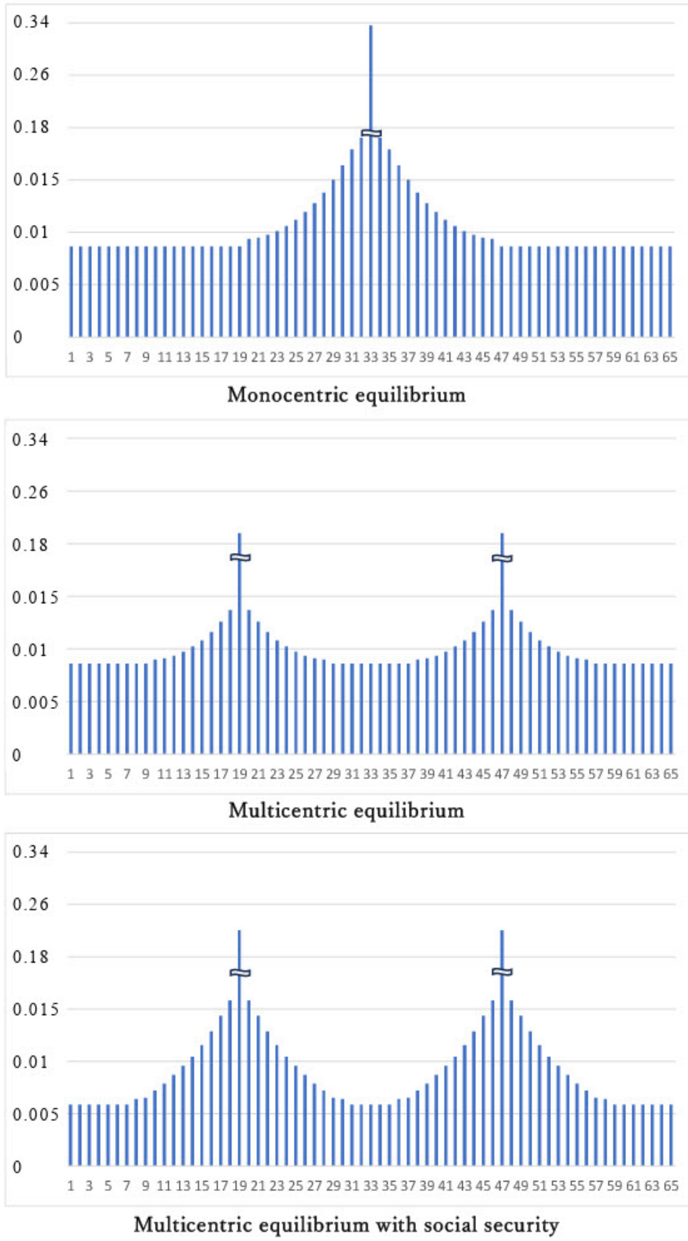

Figure 6: Population distributions in monocentric and multicentric equilibrium

Population in regions in monocentric equilibrium, in multicentric equilibrium, and multicentric equilibrium with social security. The total population in cities, the population with higher education, and the number of regions where higher education is observed in multicentric equilibrium exceed those in monocentric equilibrium. However, population size in each city is smaller. Social security policy increases the net benefits of higher education. The policy makes the child depend more on the individual, especially when the child becomes an unskilled worker and remains in the hometown with the parent. In this case, the child reduces the provision of family public goods more drastically. Therefore, an increase in the public provision of family public goods increases the total provision more dramatically in the case where her child becomes a skilled worker. The population with higher education and that of skilled workers in the city increase, and it raises the wage. It further induces individuals to give a higher education and, thus, the urban size expands in its population and area.

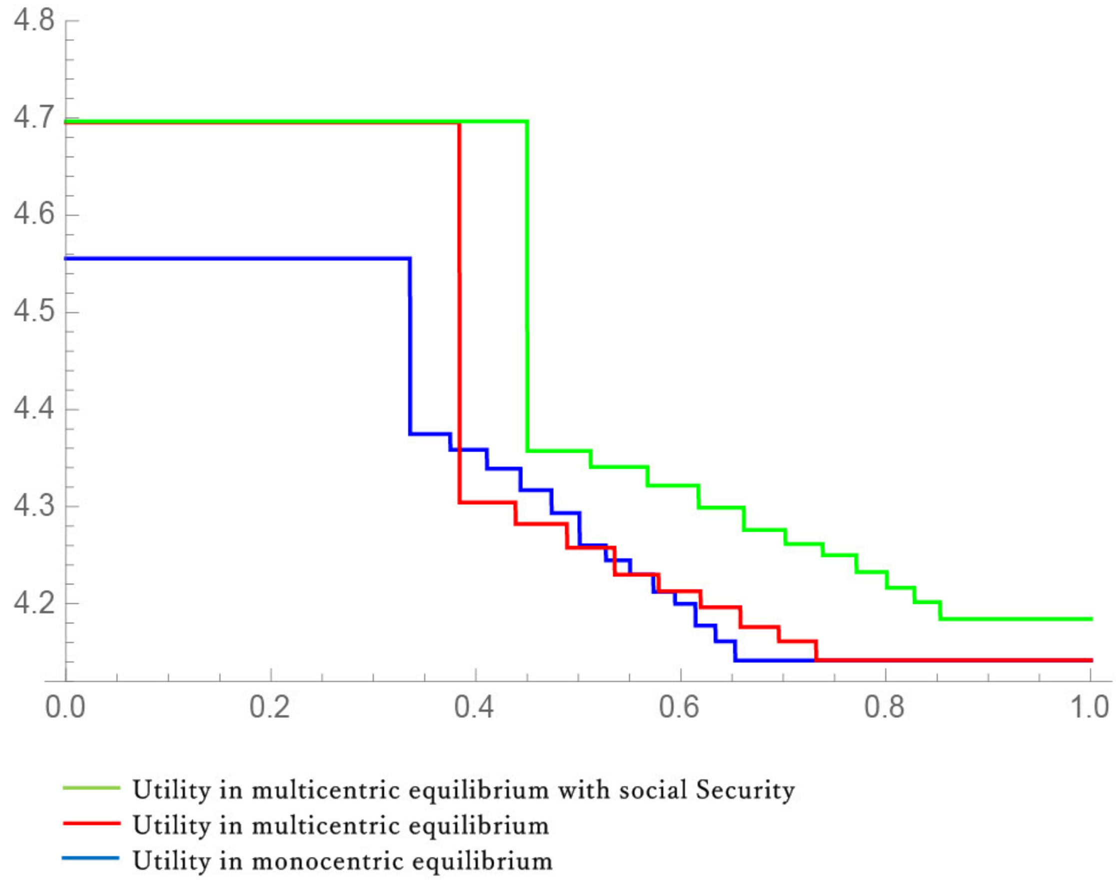

Figure 7: Utilities in monocentric and multicentric equilibrium

Lifetime expected utility in monocentric equilibrium, in multicentric equilibrium, and multicentric equilibrium with social security. The utilities are lined up in descending order. In multicentric equilibrium, many individuals can be better off, and the inequality problem is mitigated. Population size in each city in multicentric equilibrium becomes smaller than that in monocentric equilibrium. The utility loss caused by a fall in wages can be sufficiently compensated by the utility gain caused by less congestion in the city. A larger population will have higher education, and the total population in cities in multicentric equilibrium is larger. Therefore, a larger population is better off. In addition, in multicentric equilibrium, the advantage of unskilled workers in regions near the city enjoying large amounts of family public goods provided by the parents in the city is less prominent, as the wage in the city is lower. Except unskilled workers who luckily have the parent in the city and live in regions near the city, individuals are better off in multicentric equilibrium. The introduction of the social security policy induces higher education, which expands the urban size in its population and area. A larger population will have higher education. It raises the wage in cities, and enhances the utilities of unskilled workers near the cities. Therefore, all the individuals can be better off.

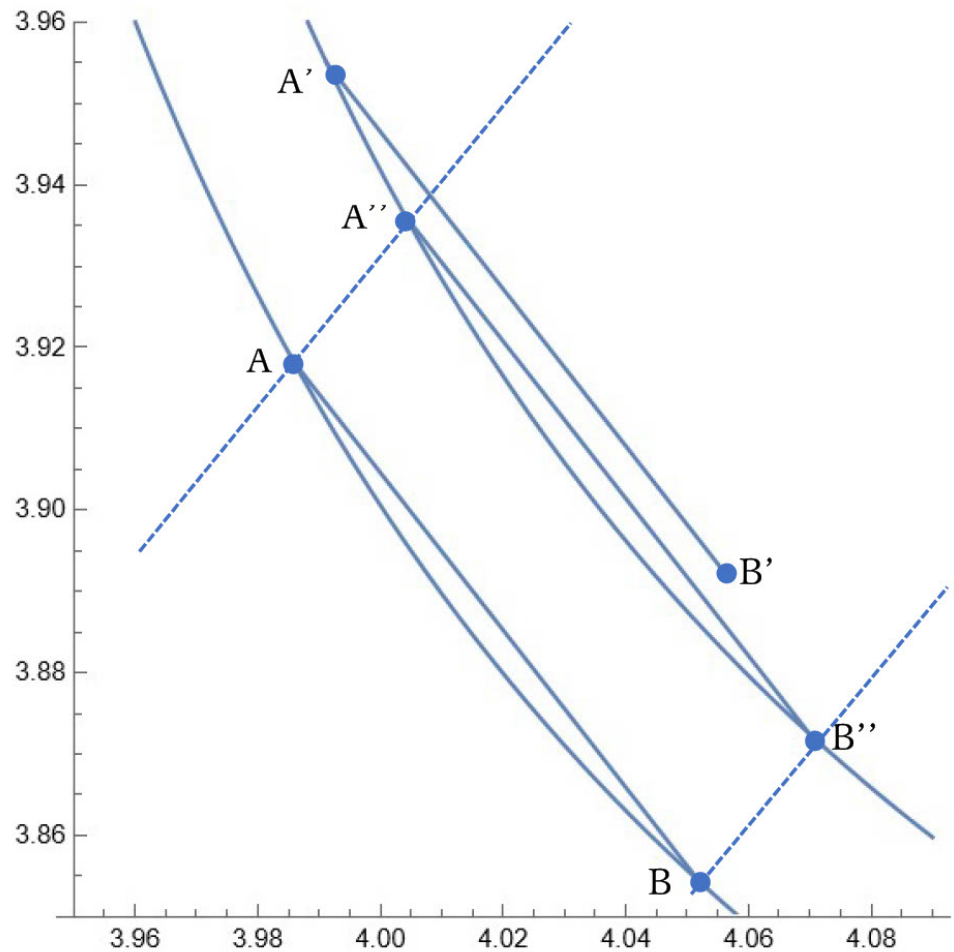

Figure 8: Consumption of private goods and provision of family public goods

Consumption of private goods and the total provision of family public goods when an individual is old and her child becomes an unskilled worker. In cases of B and B', an individual places her child in region $n+l+1$ when her child becomes an unskilled worker. In contrast, in cases A and A', she places her child in region $n+1$ by providing larger amount of family public goods. Also, in cases of A' and B', an individual lives in region $n+1$, whereas in cases of A and B she lives in region n. Therefore, in cases of A' and B', her child lives closer to the individual when her child becomes an unskilled worker, and, thus, the individual can enjoy higher value of family public goods provided by her child than in cases of A and B, respectively. The slope of the line segment AB is steeper than $A'B'$. When an individual increases her provision of family public goods, the value of total family public goods provision also increases but less drastically. That is because her child will decrease her contribution. However, when her child initially lives far away from the region where an individual lives (the case of the line segment AB), the distance mitigates this negative impact. Points A' and B' locate north-west of point A'' and B'', respectively. As line $A'B'$ is less steeper than line AB $A''B''$, point B' locates upper than the line segment $A''B''$. Thus, the utility on point B' is higher than the utility that corresponds the indifference curve on which points A', A'' and B'' are.

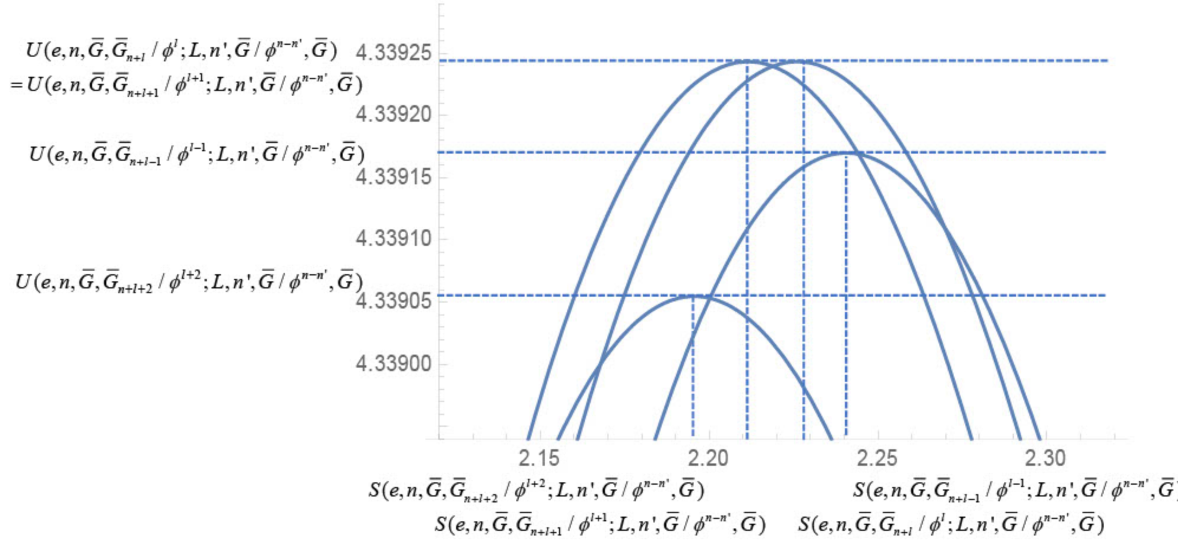

Figure 9: Maximum conditional expected lifetime utility

An individual's expected lifetime utilities $\tilde{U}(S_{t+1},e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,\bar{G}_n,\bar{G})$, $\tilde{U}(S_{t+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,\bar{G}_n,\bar{G})$, $\tilde{U}(S_{t+1},e,n{+}1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n{+}1,\bar{G}_n,\bar{G})$, and $\tilde{U}(S_{t+1},e,n{+}1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,\bar{G}_n,\bar{G})$. In the example in this Figure, $n=3$ and $s=4$ (as $3\in\Theta_4$, $l=s=4$). For $S_{t+1}\in[S(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,\bar{G}_n,\bar{G}),S^n_{n+s,n+s+1}]$, the absolute value of the slope of $\tilde{U}(S_{t+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,\bar{G}_n,\bar{G})$ is larger than that of $\tilde{U}(S_{t+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,\bar{G}_n,\bar{G})$. Therefore, the gap between $U(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,\bar{G}_n,\bar{G})$ and $\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n{+}1,\bar{G}_n,\bar{G})$ is larger than the gap between $U(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,\bar{G}_n,\bar{G})$ and $\tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,\bar{G}_n,\bar{G})$. Similarly, as the slope of $\tilde{U}(S_{t+1},e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,\bar{G}_n,\bar{G})$ is larger than that of $\tilde{U}(S_{t+1},e,n{+}1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n{+}1,\bar{G}_n,\bar{G})$ for $S_{t+1}\in[S^n_{n+s,n+s+1},S(e,n{+}1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n{+}1,\bar{G}_n,\bar{G})]$, the gap between $U(e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,\bar{G}_n,\bar{G})$ and $\tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,\bar{G}_n,\bar{G})$ is larger than the gap between $U(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,\bar{G}_n,\bar{G})$ and $\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,\bar{G}_n,\bar{G})$. Therefore, the gap between $U(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,\bar{G}_n,\bar{G})-U(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,\bar{G}_n,\bar{G})$ and $\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,\bar{G}_n,\bar{G})-\tilde{U}(S^n_{n+s,n+s+1},e,n{+}1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n{+}1,\bar{G}_n,\bar{G})$ is larger than that of $U(e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,\bar{G}_n,\bar{G})-U(e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,\bar{G}_n,\bar{G})$ and $\tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s+1}/\phi^{s+1};L,n,\bar{G}_n,\bar{G})-\tilde{U}(S^n_{n+s,n+s+1},e,n,\bar{G},\bar{G}_{n+s}/\phi^{s};L,n,\bar{G}_n,\bar{G})$. The latter is zero and $\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,\bar{G}_n,\bar{G})-\tilde{U}(S^n_{n+s,n+s+1},e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,\bar{G}_n,\bar{G})$ is positive, thus $U(e,n+1,\bar{G},\bar{G}_{n+s+1}/\phi^{s};L,n+1,\bar{G}_n,\bar{G})-U(e,n+1,\bar{G},\bar{G}_{n+s}/\phi^{s-1};L,n+1,\bar{G}_n,\bar{G})$ is positive.