To Stay or to Migrate? When Becker Meets Harris-Todaro

Pei-Ju Liao[a], Ping Wang[b], Yin-Chi Wang[c], Chong Kee Yip[d]

[a]National Taiwan University, Taipei, Taiwan
[b]Department of Economics, Washington University in St. Louis, St. Louis, USA & NBER, USA
[c]National Taipei University, New Taipei City, Taiwan
[d]Chinese University of Hong Kong, Hong Kong SAR

This draft: July 18, 2022

Abstract: Allowing migration activity as an integral part of demographic transition and economic development, we construct a dynamic competitive migration equilibrium framework with rural agents heterogeneous in skills and fertility preferences to establish a locational quantity-quality trade-off of children and explore its macroeconomic consequences. We show and characterize a mixed migration equilibrium where high-skilled rural agents with low fertility preferences always migrate to cities, low-skilled agents with high fertility preferences always stay, and only an endogenously determined fraction of high-skilled agents with high fertility preferences or low-skilled agents with low fertility preferences ultimately moves. By calibrating our model to fit the data from China, whose migration and population control policies offer a rich array of issues for quantitative investigation, we find strong interactions – the locational quantity-quality trade-off – between fertility and migration decisions, as well as between these decision-makings and changes in migration and population control policies. We conclude that overlooking the locational quantity-quality trade-off may lead to nonnegligible biases in assessing the implications and effectiveness of government policies.

JEL Classification: E24, O15, R23.

Keywords: Rural-urban migration, fertility, urbanization, development.

Acknowledgments: This is a substantially revised paper from the AEA-ASSA conference paper in 2017 (different title) and the NBER working paper #27767 in 2020. Throughout the long research process, we are grateful for comments/suggestions from Douglas Almond, Carlos Garriga, Chang-Tai Hsieh, Wen-Tai Hsu, Rachel Ngai, Ray Riezman, Michael Song, as well as participants at the AEA-ASSA Annual Meetings, the Asian Meetings of the Econometric Society, the Society for Advanced Economic Theory Meeting, the Society for Economic Dynamics Meeting, the Regional Science Association International Meeting, and the Taiwan Economic Research Conference and various seminars. Travel support from Academia Sinica, the Research Grant Council of Hong Kong, the Center for Dynamic Economics of Washington University, and the Ministry of Science and Technology of Taiwan are gratefully acknowledged. Lei Ye, Hui-wen Chang and Chian-Yu Lin provided excellent research assistance. Needless to say, the usual disclaimer applies.

* E-mail addresses: pjliao@ntu.edu.tw (P.-J. Liao), pingwang@wustl.edu (P. Wang), yinwang@mail.ntpu.edu.tw (Y.-C. Wang, Corresponding Author), chongkeeyip@cuhk.edu.hk (C. K. Yip).

1 Introduction

Fifty years ago, Todaro (1969) and Harris and Todaro (1970) provided a solid foundation for studying the process of rural-urban migration commonly observed in developing countries. With the more recent contribution by Lucas (2004), vast interest has surfaced with respect to examining urbanization and structural transformation within a dynamic general equilibrium framework. While this renewed literature has generated useful insight toward understanding not only urban labor and housing issues, but also various macroeconomic consequences of rural-urban migration, what role demographic transition plays in the dynamic urbanization process remains largely unexplored.

Generally, raising children in the countryside is relatively cheap compared to doing so in urban areas, but cities are usually considered as a better location for children when accounting for education and career opportunities. The chief purpose of this paper is to establish a dynamic internal migration model in which such a *locational quantity-quality trade-off* of children may arise as an equilibrium outcome. In other words, we inquire, when Becker meets Harris-Todaro, whether the interplay of work-based rural-urban migration and fertility decisions may influence the macroeconomy of developing countries in the presence of large migration barriers and active population controls. Specifically, our paper is devoted to addressing the following questions.

Question 1: What are the main channels through which work-based rural-urban migration and differential fertility decisions interplay in the process of economic development to yield a locational quantity-quality trade-off of children?

Question 2: What are the underlying structural transformation and migration- and population-related institutional factors driving such household decisions?

Question 3: What are the macroeconomic consequences of the fertility-migration interactions and the resulting locational quantity-quality trade-off, particularly with regards to per capita output, urban-rural income gap, labor composition, and skill and sectoral wage premia?

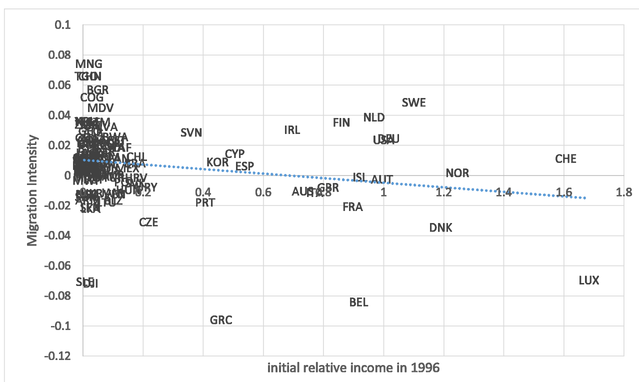
To further motivate our study, it is informative to investigate across countries how migration and fertility are related to their development stages. We choose a sample period from 1996 to 2015 to accommodate data availability, especially for underdeveloped countries. Both the development stage and the fertility measure are measured using common standards with development being measured by initial relative income (to the United States) in 1996 and the fertility measure being assessed as the total fertility rate (TFR). Both measurements are based on the World Development Indicators (WDI). To measure migration activity, however, the process is not straightforward. First, the typically used census-based IPUMS data do not separate various types of migration (i.e., rural-rural, rural-urban, urban-rural, urban-urban). To suit for our study, we thus create a measure of rural to urban migration intensity of country i in year t (MI_{it}, see Appendix IA for details):

$$MI_{it} = \frac{\text{urban-rural employment ratio of } i \text{ in } t+1 \; - \; \text{urban-rural employment ratio of } i \text{ in } t}{1 + \text{urban-rural employment ratio of } i \text{ in } t+1}$$

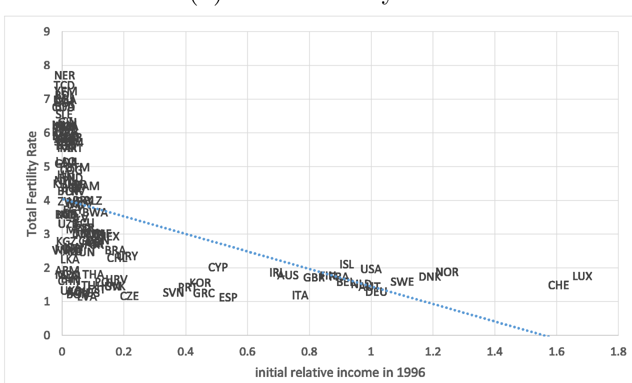
where urban and rural employment data are taken from the Global Jobs Indicators Database (JOIN) of the World Bank. We then average the time-series data of each country to obtain cross-country measures of migration intensity. This process enables us to have a large sample with 33 developed countries and 83 developing countries (classified by the World Bank in our initial year of 1996). Figures 1(a) and 1(b) provide cross-country scatter plots of migration intensity and TFR, respectively, against each country's initial relative income.[1] Our results indicate that the less developed a country, the higher the migration intensity and the TFR. The correlation coefficient of MI and initial relative income is -0.208 (significant at the 5% level), whereas the correlation coefficient of TFR and initial relative income is -0.501 (significant at the 1% level). We further examine whether the mean MIs or mean TFRs between the group of 33 developed and 83 developing countries differ. Because neither MI nor TFR is normally distributed, we perform the Mann-Whitney test and find the former significant at the 5% level and the latter at the 1% level. We thus conclude a casual relationship: Migration and fertility decisions are both related to the stage of development, with less advanced featuring higher fertility and more intensive internal migration to urban areas. Put differently, migration activity is most plausibly an integral part of demographic transitions and economic development, thereby motivating the present paper.

Figure 1: Initial Relative Income, Migration Intensity, and Total Fertility Rate

(a) Migration Intensity



(b) Total Fertility Rate



[1] In Appendix IA, we provide scatter plots using IPUMS-based merged data from Bernard, Bell, and Cooper (2018) – a measure of total rather than rural-urban MI – and find similar patterns.

What, then, are the costs and benefits associated with rural-urban migration? The major costs crucial for the interplay of migration and fertility are pecuniary and time costs associated with migration, the opportunity cost associated with forgone rural earnings and farm land-use rights, and the higher childrearing cost in urban areas. Such migration is beneficial, however, due to better pay and job perspectives, a better education and future for children, and better urban amenities net of urban congestions/pollution and the rising cost of living. Should the marginal benefit exceed the marginal cost, rural workers are expected to continue migrating to cities. Since both the accrued cost and the opportunity cost of childrearing are high in urban areas, a consequence of such migration decisions is a reduction in childbearing in conjunction with an improved quality of living locationally.

The basic story mentioned above is simple and intuitive, but formalizing it in a dynamic model with fertility and locational choices by heterogenous agents is by no means straightforward. To enable the analysis in a tractable and quantifiable manner, we construct an infinite-horizon dynasty model à la Aghion and Bolton (1997) with agents of each generation taking various actions at different points during agents' one-period lifetime and with inter-generational linkages through warm-glow bequests. To be consistent with the reality that rural workers migrate to cities at young ages, in their one-period lifetime, rural agents first make migration decisions, then work at their chosen locations, obtain urban residency if lucky, consume, give birth to children, and leave bequests to children prior to exiting the market. This framework is natural for studying the nexus between parents and children with parents making fertility, bequest, and migration decisions and children inheriting parental givings, skills, and residency. Within this framework, we can characterize the dynamic competitive migration equilibrium by allowing agents to have two-dimensional heterogeneities in skills (high or low) and in fertility preferences (high or low) and by allowing a rich array of frictions and public policies.

The consideration of a dynamic setting is valuable for analyzing the nexus between fertility and migration choice, and macroeconomic outcomes. Specifically, we address the first research question by considering two important dynamic channels:

(i) Direct Feedback Channel: Fertility choice would affect migration cost and incentives, thus influencing the migration decision; migration choice affects cost and the opportunity cost of childrearing and, hence, the fertility decision. Notably, such feedback depends crucially on individual perceived value where the bequest and the valuation of migration in the dynasty setting interact to affect each other during the decision-making process.

(ii) Dynamic General Equilibrium Channel: Fertility and migration decisions would affect labor composition, market wage, and output, which in turn influence fertility and migration choices. Specifically, we will model a temporal general equilibrium effect induced by changes in labor composition that tend to generate a downward pressure on the wages of abundant workers; moreover, we allow both fertility and migration decisions to affect population dynamics. Such dynastic and locational dynamic changes subsequently impact the workforce, leading to sectoral shifts not only *between* rural and urban areas but also *within* the urban area between different sectors. Further, as more agents migrate to cities in the process of economic development, the general equilibrium effect on urban wages subsequently reduces

> migration incentives and, as our paper will show, the model economy would eventually approach a *balanced growth path* (BGP) featuring common growth and mixed migration, along which a locational quantity-quality trade-off between payoffs from locational quality and fertility is present. Along the BGP, we will establish that a better urban amenity, a larger childrearing subsidy, a lower urban-rural childrearing cost differential, a lower urban fertility penalty, or a more secured rural land entitlement speeds up the urbanization process.

More specifically, we establish conditions to support a most plausible "mixed migration equilibrium" along the BGP: High-skilled rural residents with low fertility preferences always desire to migrate to cities, low-skilled rural residents with high fertility preferences always stay in rural areas, and only an endogenously determined fraction of those high-skilled residents with high fertility preferences or low-skilled residents with low fertility preferences ultimately moves. In a less developed economy under proper regularity conditions, we show that the fraction of high-skilled movers with high fertility preferences is lower than that of low-skilled agents with low fertility preferences. The moving fraction differential shrinks when a tightened population policy exists that binds agents with high fertility preferences. A key insight stems from allowing for interrelated fertility and migration decisions: the establishment of a locational quantity-quality trade-off of children in the process of economic development in which more rural-urban migration is accompanied by lower fertility but higher overall per capita output.

Just how important is the locational quantity-quality trade-off of children for economic development, particularly in view of the second and the third research questions? To address this quantitative question, we calibrate our model economy to fit the data from China, which is particularly interesting because of its migration and population policies. These policies include the household registration system, a so-called *hukou system* (HS) that imposes large barriers on rural-urban migration, the *differential one-child policy* (DOCP) with stronger enforcement in cities and in public sectors. These three policies allow us to perform several interesting policy experiments and counterfactual analyses to assess the migration, fertility, and macroeconomic consequences of migration and population control policies. In all cases, we find that the locational quantity-quality trade-off of children plays a crucial role for economic development where more rural-urban migration is accompanied by lower fertility but a higher overall per capita output. In response to changes in population control policies, a negative association exists between the overall per capita output and skill premium; in response to changes in migration-related policies, however, such association turns to positive.

Notably, our findings of completely removing population control align with the pure fertility choice model of Liao (2013): Skill premium increases and fertility rebounds, especially the fertility of those who were seriously affected by population control policies. Our results further indicate that, by allowing rural-urban migration, the increase in fertility is accompanied by lower migration – especially a larger reduction in the high-skilled migration and thereby a lower overall output per capita. Overall, relaxing urban population control to the rural level is better than entirely removing population control as the local control induces more high-skilled migration, leading to higher urban output and overall output per capita. Stricter population control policies in urban areas may not be ideal in lowering fertility rates from a nationwide perspective because such a tightened fertility policy in urban areas deters high-skilled workers with stronger fertility preferences from migrating to cities. This effect is undesirable because had such workers migrated to urban areas and faced a higher childrearing cost, the workers would have naturally adjusted their fertility down. Moreover, we find that an urban sprawl control policy eliminating urban benefits to all rural migrants will raise urban benefits as planned but would severely deter the incentives for rural-urban migration, which is undesirable for advancing the economy. An urban promotion policy such as providing full benefits to all rural migrants would induce more migration overall.

Furthermore, by conducting counterfactual analysis, we verify quantitatively strong interactions between dynamic decision-makings for migration and fertility. We further show that, under a generalized framework allowing for reverse migration, left-behind children, upward skill mobility, and various directed urban benefits, the importance of the locational quantity-quality trade-off of children remains whereas the main policy outcomes continuously hold true. Finally, various robustness checks are performed to ensure the validity of our main findings.

Wrapping up, we now deliver our answers to the two remaining research questions.

Answer to Question 2: In addition to the key driver of structural transformation via urban TFP advancement, we emphasize two mutually connected migration and population institution/policy factors in China:

- **(i)** The **Hukou System** (HS): HS limits rural-urban mobility, affects urban sectoral choice, and influences the extent to which migrants may enjoy urban benefits.
- **(ii)** The **Differential One-Child Policy** (DOCP): DOCP not only restricts fertility choice but also leads to differential childrearing costs, which subsequently affect the incentive to migrate to urban areas that impose tighter population controls and have larger penalties on violation (including monetary cost and demotion or layoff from the state-owned enterprises).

Answer to Question 3: Our main findings linking migration and fertility decision-making to macroeconomic outcomes are threefolds:

- **(i)** **Fertility-Migration Interactions**: Migration decisions not only affect fertility decisions due to tighter population control in urban areas but also impact the skill and sectoral composition of urban workers, thereby influencing urban wages and per capita output. The quantitative and qualitative decisions in fertility affect the direct incentives for rural-urban migration and, through bequest giving and skill and residency inheritance, influence the skill and sectoral compositions of urban workers, thus impacting urban wages and per capita output.
- **(ii)** **Structural Transformation**: The key drivers of structural transformation constitute the urban-rural technology gap, the human-capital advantages of urban workers, and the productivity benefits from high- to low-skilled workers, thus widening urban-rural income gaps.

(iii) Institutions and Policies: HS-induced mobility frictions and migration policies and DOCP are both important for the interplays of migration and fertility decisions and their macroeconomic consequences, especially because the effects on the locational quantity-quality trade-off are interconnected. Thus, the interactions between migration and fertility decisions not only generate a locational quantity-quality trade-off but also reinforce structural transformation, leading to more rapid urbanization and advancing economic development.

The main takeaway is that, when formulating policies for developing countries, one shall not neglect policy impacts via migration and fertility responses. Ignoring the consequential locational quantity-quality trade-off of children may lead to nonnegligible biases in assessing the consequences and effectiveness of government policies.

Literature Review Previous works on reallocating abundant and over-employed labor from the rural agricultural sector to the urban manufacturing sector can be traced back to Lewis (1954). The research focusing on rural-urban migration was pioneered by Todaro (1969) and Harris and Todaro (1970). Since then, economists have attempted to understand the forces driving rural-to-urban migration and how rural-urban migration impacts the development process.

Some related studies have examined the causes and consequences of rural-urban migration using calibrated dynamic models. The pivotal study is by Lucas (2004), who proposes that human capital accumulation in cities induces better earnings and hence migration from rural to urban areas. Rural-urban migration ceases when the values of earnings in the two locations equalize. Laing, Park, and Wang (2005) build a dynamic search model to illustrate how reductions in urban labor market frictions may yield higher wages and induce more rural-urban migration. In a companion study, Liao, Wang, Wang, and Yip (2022) examine the rural-urban migration in China with a focus on education-based migration and establish that the contribution of education-based migration on urban output shares are comparable to that of work-based migration. In two independent related papers, Ngai, Pissarides, and Wang (2019) illustrate how much the household registration system in China has slowed industrialization and urbanization in China, whereas Ma and Tang (2020) connect external frictions in trade to internal migration frictions using data from China where the interplays lead to rich local labor market and welfare outcomes. For a comprehensive literature review of internal migration from a macro perspective, the reader is referred to Lagakos (2020).

Limited papers have been devoted to studying fertility and migration decision-making: Sato and Yamamoto (2005), Sato (2007), and Cheung (2022). Within a static setting, the reduction in child mortality led to urbanization in Sato and Yamamoto (2005), whereas urban agglomeration economies and congestion in Sato (2007) could interact with fertility choices to yield a negative correlation between income and fertility across regions. In an independent work within the overlapping-generations framework, Cheung (2022) highlights rural education reform in the United States in the early 1900s as the key driver of the demographic transition and the sector shift from rural farming to urban manufacturing.

In our paper, we also emphasize the structural transformation via urban TFP advancement that widens the urban-rural productivity gap. If one may assume, as in Lucas (2004), that the possible negative free-rider problem is dominated by the direct positive productivity effect or a positive learning effect, agglomeration economies can be captured by our urban TFP advancement in an observational equivalent sense. Thus, structural transformation-led urban productivity advantages may be viewed as a common driver in Sato (2007), Cheung (2022), and the present paper. Beyond this, none of the other factors play much of a role in the case of China where child mortality was low by 1980, no major rural education reforms were launched throughout the period of our study, and the fertility-urban congestion relationship has not yet been identified empirically. Rather, the *hukou* system that affects migration and labor mobility and the DOCP between rural and urban areas, together with continual TFP advancement, are central in the process of urbanization and structural transformation in China and are the drivers considered in the present paper.

To the best of our knowledge, the novelty of our paper is to examine rural-urban migration with the endogenous fertility choice within a dynamic general equilibrium framework, under which we can infer decision rules of migration for agents with heterogeneous skills and preferences toward children. Also, based on the framework we can quantify the extent to which migration and population policies affect rural-urban migration and economic development.

2 The Model

As discussed in the introduction, migration intensities and TFRs are related to economic development. Our model will be built upon these facts. To allow for policy analysis, our model will incorporate a rich array of migration- and fertility-related policies prevalent in developing countries.

Consider an infinite-horizon dynasty model with agents of each generation taking various actions at different points in time during the agents' one-period lifetime, a tractable dynamic setting developed by Aghion and Bolton (1997). Although tractability follows from the non-overlapping feature of the generations, inter-generational linkages are restored via "warm-glow" bequests from parents to children (Andreoni, 1989). In addition to bequests considered in the model, agents of two adjacent generations in our economy are also connected via fertility decisions and locational choices. There are two geographical locations, urban and rural. In rural areas, production is simply backyard farming, using land as an input. There are two sectors in cities: a private sector (P) and a state-owned enterprises sector (S) that contains government agencies and public enterprises (SOE). While workers in the SOE sector all hold urban residency, the private sector employees may not. Urban workers with urban residency are entitled with urban benefits, whereas rural farmers are entitled with the use rights of rural land. These rights are often observed in developing countries under various usage and zoning restrictions, such as regulations imposed on agricultural land.

As agents make migration decisions within the one period of their lifetime, in any location or sector, the beginning-of-the-period population and the actual-working population stocks are different, with the differences resulting from the net migration flows of workers. To avoid confusion, the number of workers actually working in the sector during a period *after migration occurs* is distinguished by a superscript "+". We use the subscript t to index time and dismiss it whenever doing so does not create confusion.

Agents live for one period and are heterogeneous in skill levels and fertility preferences. There are two different skill levels, high (H) and low (L), and two types of quantity-based altruistic factors (β), high ($\beta = \bar{\beta}$) and low ($\beta = \underline{\beta}$), with the high type in favor of having more children. Children inherit skills and residency status from parents, but the quantity-based altruistic factors are assumed to be redrawn for every generation. Agents consume and give birth to n children right before the end of life. Agents are altruistic, deriving utility from both the quantity and quality of children, owning one unit of labor time and supplying labor inelastically throughout life.

In what follows, we will first describe the sectoral production in the economy and then study the household optimization problems in urban and rural locations. Once we define the value functions of the agents, we then delineate how rural agents make decisions about migrating to urban areas. Finally, we discuss the evolution of workers and study the equilibrium of the economy.

2.1 Production

We begin with rural production, followed by the sectoral production activities in the urban location.

2.1.1 Rural production

Rural land within a period is given at an exogenous level Q but can vary across periods. Use rights for rural land are evenly distributed to rural workers at the beginning of each period. Rural production uses land as inputs and requires workers to stay on site to operate the production technology, but does not require skills. Denote R as the beginning-of-the-period farmers in rural areas and R^+ as the mass of farmers residing on the rural land during a period after migration occurs. For migrants moving to cities and leaving their land behind, they retain their entitlement of land for $\delta \in [0, 1]$ fraction of lifetime before losing it. The use rights of the idle land will then be evenly reallocated to farmers staying in rural areas. Total output in the rural location is:

$$X = zQ\left(1 - \tilde{\delta}\right), \tag{1}$$

where $z > 0$ is the farming technology and $\tilde{\delta} \equiv (R - R^+)\,\delta/R$ is the fraction of land left idle from the total land. Define $q \equiv Q/R$ as the beginning-of-the-period land per farmer. Rural workers' per capita income is given by,

$$x \equiv \frac{X}{R^+} = zq\left[1 + \frac{(R - R^+)(1 - \delta)}{R^+}\right], \tag{2}$$

where the first term in the bracket times q is the initially-distributed land and the second term times q is the extra land obtained from land reallocation during the period.

2.1.2 Urban production

There are two types of firms in urban areas: SOE and private firms. The SOE operates a linear technology, using relatively high-skilled workers as inputs to produce output:

$$Y_S = A_S \eta S^+, \tag{3}$$

where A_S is the technology scaling factor of SOE, S^+ is the number of high-skilled workers hired by SOE, and η is the quality index of high-skilled workers that captures the relative productivity of the high- to low-skilled workers.

Unlike SOE, private firms hire both high- and low-skilled workers in production. Denote P^{H+} and P^{L+} as the quantity of high- and low-skilled workers, respectively, hired by private firms. The production of private firms takes the following CES form:

$$Y_P = A_P \left[\alpha \left(\eta P^{H+}\right)^\sigma + (1-\alpha)\left(P^{L+}\right)^\sigma\right]^{\frac{1}{\sigma}}, \quad \eta > 1,\, \alpha \in (0,1)\,, \sigma < 1, \tag{4}$$

where A_P is the technology scaling factor of private firms, α is the share parameter, and $1/(1-\sigma)$ is the elasticity of substitution between the high- and low-skilled workers in production. Under competitive markets, the wage rates in SOE and private firms are:

$$\begin{aligned}
w_S &= A_S \eta, && (5)\\
w_P^H &= A_P \left[\alpha \left(\eta P^{H+}\right)^\sigma + (1-\alpha)\left(P^{L+}\right)^\sigma\right]^{\frac{1}{\sigma}-1} \alpha \eta^\sigma \left(P^{H+}\right)^{\sigma-1}, && (6)\\
w_P^L &= A_P \left[\alpha \left(\eta P^{H+}\right)^\sigma + (1-\alpha)\left(P^{L+}\right)^\sigma\right]^{\frac{1}{\sigma}-1} (1-\alpha) \left(P^{L+}\right)^{\sigma-1}. && (7)
\end{aligned}$$

It is then straightforward to derive the wage ratio:

$$\frac{w_P^H}{w_P^L} = \frac{\alpha}{1-\alpha}\eta^\sigma \left(\frac{P^{H+}}{P^{L+}}\right)^{-(1-\sigma)} \equiv \omega_1\left(\frac{P^{H+}}{P^{L+}}\right), \tag{8}$$

which depends negatively on the high- to low-skilled employment ratio (P^{H+}/P^{L+}) but positively on the quality of high-skilled workers (η). Specifically, we have $\omega_1(0) = \infty$, $\omega_1(\infty) = 0$, and $\omega_1'\left(\frac{P^{H+}}{P^{L+}}\right) < 0$. Likewise, we specify another useful wage ratio:

$$\frac{w_S}{w_P^H} = \frac{A_S}{\alpha A_P}\left[\alpha + \frac{1-\alpha}{\eta^\sigma}\left(\frac{P^{H+}}{P^{L+}}\right)^{-\sigma}\right]^{-\frac{1-\sigma}{\sigma}} \equiv \omega_2\left(\frac{P^{H+}}{P^{L+}}\right), \tag{9}$$

where $\omega_2(0) = 0$ and $\omega_2'\left(\frac{P^{H+}}{P^{L+}}\right) > 0$. This wage ratio depends positively on three factors: the high- to low-skilled employment ratio (P^{H+}/P^{L+}), the quality of high-skilled workers (η), and the relative technological productivity of the SOE to the private sector. In the benchmark case, we assume that $w_S/w_P^H \leq 1$ for two reasons. First, workers in the S sector enjoy a higher level of the urban benefits than those in the P sector so that there is a SOE discount in the relative wage. Second, productivity is commonly higher in the private sector than in the SOE sector.[2]

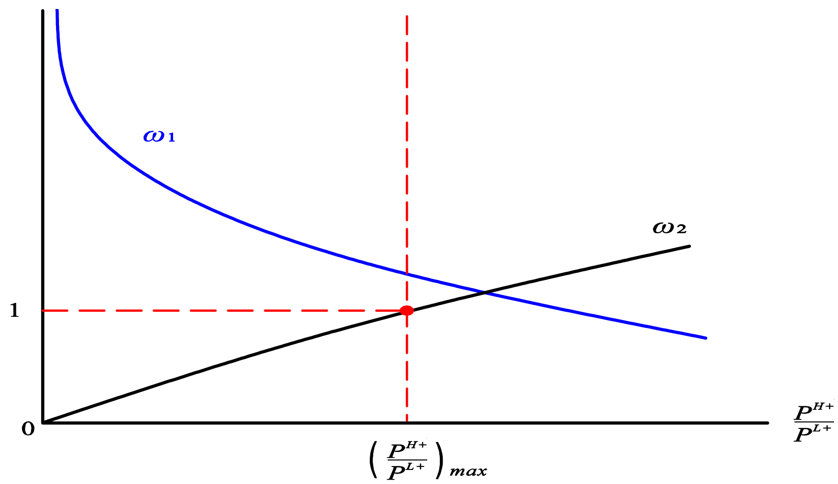
[2] This issue of SOE workers enjoying more urban benefits than private workers is further discussed in Subsection 2.2.2. A counterfactual exercise on the contrary case of relative productivity can be found in Appendix ID.

Define $\omega_2\left(\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}\right)=1$ so that $\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}=\left[\alpha^{-\frac{1}{1-\sigma}}\left(\frac{A_S}{A_P}\right)^{\frac{\sigma}{1-\sigma}}-1\right]^{-\frac{1}{\sigma}}\left(\frac{\alpha}{1-\alpha}\eta^{\sigma}\right)^{-\frac{1}{\sigma}}$. To satisfy $w_S/w_P^H\leq 1$, the equilibrium labor market must be subject to the restriction $\frac{P^{H+}}{P^{L+}}\leq\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}$. To assure that $\frac{w_P^H}{w_P^L}>1\geqslant\frac{w_S}{w_P^H}$ for all $\frac{P^{H+}}{P^{L+}}\leqslant\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}$, we impose the following condition:[3]

Condition 1. $\left[\alpha^{-\frac{1}{1-\sigma}}\left(\frac{A_S}{A_P}\right)^{\frac{\sigma}{1-\sigma}}-1\right]^{\frac{1-\sigma}{\sigma}}\left(\frac{\alpha}{1-\alpha}\eta^{\sigma}\right)^{\frac{1}{\sigma}}>1.$

Under Condition 1, we have $\omega_1\left(\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}\right)>1$ and $\omega_2\left(\left(\frac{P^{H+}}{P^{L+}}\right)\right)\leq 1$. Figure 2 provides a graphical representation of Condition 1.[4]

Figure 2: Sectoral Mobility and Condition 1



2.2 Households

All agents have perfect foresight and live as an adult for one period. When the parent exits the economy at the end of a period, the agent begins adulthood in the location according to personal residency in the following period.[5] During the one-period adult lifetime, agents take various actions at different points in time, including making decisions on migration, work, fertility choice, and consumption-bequest choice. In this dynamic dynasty setting – via births, bequests, and locational choices – parental decisions affect children's optimization. As mentioned previously, agents are born with different skill levels and preferences toward children and inherit skill levels and residency status from parents (the assumptions will later be relaxed in Sections 4.1 and 4.3). Below, the lifetime of rural households is described first, followed by the description for urban households.

[3]See Appendix II (Section A.0) for derivation details.

[4]Figure 2 highlights the intuition that $\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}$ must locate to the left of the intersection of ω_1 and ω_2.

[5]Formally speaking, there is an overlapping instant between the parent and the child at the end of a period (after the fertility choice is made), which is the childhood of an agent. Nevertheless, this overlapping moment does not impact any economic variables in the model.

2.2.1 Rural agents

Due to its small magnitude in developing countries at the stage of relatively low urbanization, reverse migration from urban to rural areas is ruled out in our baseline model but allowed in the generalization in Section 4.1.[6] Agents born in rural areas choose whether to immediately migrate to cities after they were born. If the expected value of staying in the rural area is higher than the expected value of migrating to cities, agents will choose to stay in the rural area, and vice versa. After the migration decision, agents work in the chosen location throughout their lifetime. If the agent has moved to cities, urban residency may be granted after a certain duration, and that entitles the agent to enjoy urban benefits. Before the end of life, agents give birth to children subject to the fertility quota or the subsidy imposed by the government – if any exists. After giving birth to children, agents consume, incur child rearing costs, pay the above-quota penalty or receive a per child subsidy, and then exit the market.[7] Notably, by inheriting parents' urban residency status, the realization of parents' urban residency is by construction granting urban benefits to the children. Figure 3 depicts the timeline of rural agents.

Denote c as the adulthood consumption of an agent, n as the number of children to have, and b as the bequest left for each child, which is then completely consumed in the childhood. The utility of a rural worker staying in rural areas is:

$$u^R(c, b, n; \beta) = \min[\theta c, (1-\theta) nb] + \beta n^{\varepsilon}, \quad \varepsilon \in (0,1), \tag{10}$$

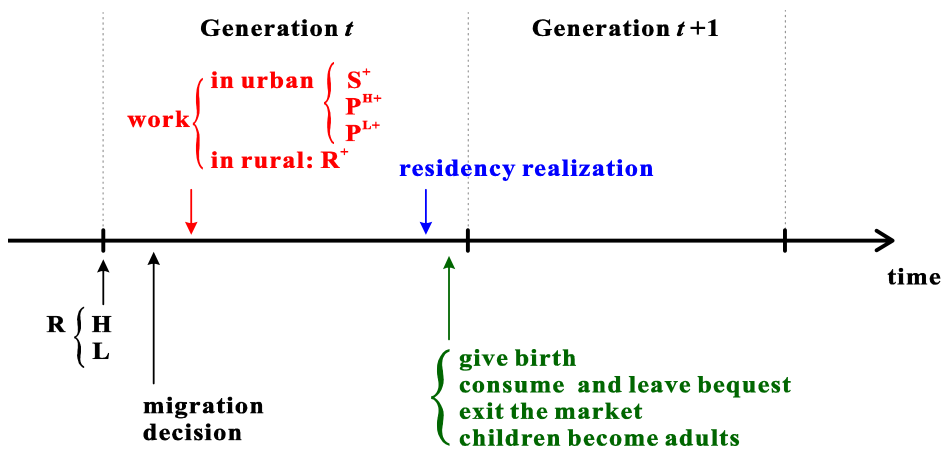
where $\theta \in (0,1)$ is the quality-based altruistic factor, which is proportional to the net income allocating to total bequests; $\beta \in \{\underline{\beta}, \bar{\beta}\}$, with $0 < \underline{\beta} < \bar{\beta}$, captures an agent's preference toward the number of children and is the quantity-based altruistic factor. β is re-drawn at birth for every generation: With a probability ζ, an agent gets less enjoyment from having children ($\beta = \underline{\beta}$), and with a probability $1-\zeta$, an agent gets more enjoyment from having children ($\beta = \bar{\beta}$). Given the agent's one-period lifetime, the Leontief setting in c and nb is meant to capture the consumption-saving and the intergenerational reallocation decision, with nb as *saving in kind*. Given the one-period setting of the model, we do not connect b directly to the productivity of the child but instead consider b as the cost of establishing good attributes of children – a quality dimension under this

[6] In the quantitative analysis below, we document that rural-urban migration is more than 12 times larger than reverse migration (Scharping, 1997). Moreover, without allowing for complicated heterogeneous locational preferences, only net migration flows matter. Banister (1997) found that the net migration rate is positive for each age group. Thus, focusing exclusively on rural-urban migration is viewed as a good benchmark.

[7] In Appendix IB.3, we provide evidence to justify our model setting that puts births as occurring after migration.

bequest setting with b capturing the investment in each child.

Figure 3: Timeline



Having children is costly. Assume that the child-rearing cost is ϕ_R^0 per child in rural areas. Two types of fertility policies are under our consideration. One is related to direct population control, whereas the other is related to the childrearing subsidy. For simplicity, both cases are modeled in the quantity aspect, which enables us to use a single variable $\bar{n}_R$ to capture either the fertility quota under control or the minimum number of children qualified for childrearing subsidies. Under direct population control, the government imposes an above-quota fine of $\bar{\phi}_R > 0$ per child, so the total penalty payment is $(n - \bar{n}_R) \cdot \bar{\phi}_R$. Under the childrearing subsidy, the government provides $-\bar{\phi}_R > 0$ of subsidy per child, giving a rural worker a total subsidy of $(n - \bar{n}_R) \cdot \left(-\bar{\phi}_R\right)$. With an income x from farming, the budget constraint of a rural worker under either scenario is then:[8]

$$c + nb + n\phi_R^0 + \max\left\{n - \bar{n}_R, 0\right\}\bar{\phi}_R = x. \tag{11}$$

A rural agent's problem is thus to choose $\{c, b, n\}$ to maximize lifetime utility (10) subject to the budget constraint (11).

2.2.2 Urban agents

Similar to rural agents, urban agents derive utility from consumption (c), quality of children (b), and number of children (n). Urban residents may also enjoy urban benefits (B) such as urban amenities, medical services, public transportation subsidies, allowances for the elderly, and other benefits. However, only residents with urban residency qualify to receive these benefits, and at the same time, such residents must pay an embarked tax τ to finance these benefits. For rural migrant workers arriving in the cities without urban residency, the workers have a probability ρ to obtain urban residency, which entitles them to urban benefits for μ fraction of their lifetime upon paying the urban embarked tax τ. The consideration of both ρ and μ not only enables us to better fit with

[8] Because the bequest received by this adult agent has already been consumed in her childhood, it does not show up in (11).

the *hukou* institutions in our quantitative analysis but also permits an extension of our baseline framework to allow for reverse migration based on ρ (see Section 4.1). As migrant workers may start their careers in cities with or without urban residency, and some may only later obtain urban residency, we need two indicator functions I^F and I^T to indicate the urban residency status of an urban worker:

$$
\begin{aligned}
I^F &= \begin{cases} 1, & \text{if the agent holds urban residency when starting to work,} \\ 0, & \text{if the agent does not hold urban residency when starting to work.} \end{cases} \\
I^T &= \begin{cases} 1, & \text{if the agent successfully obtains urban residency despite } I^F = 0 \text{ initially,} \\ 0, & \text{if the agent fails to obtain urban residency despite } I^F = 0 \text{ initially.} \end{cases}
\end{aligned}
$$

Depending on the status of urban residency, an urban worker enjoys (pays) urban benefits (taxes) of B (τ), μB $(\mu\tau)$ or nothing.

Let superscript U denote urban workers regardless of their urban residency status. An urban worker has the following utility function:

$$
u^U(c, b, n, B; \beta)\,|_{I^F, I^T} = \min\left[\theta c, (1-\theta)\, nb\right] + \beta n^{\varepsilon} + \left[I^F + \left(1 - I^F\right) I^T \mu\right] B. \tag{12}
$$

Denote $\bar{n}_U$ as the fertility quota or the minimum number of children to be eligible for urban childrearing subsidies for agents with urban residency. Two remarks are in order. First, as will be clarified below, we will consider a nondegenerate case where the fertility quota is binding for some but not all the agents. Second, urban childcare is more expensive than rural childcare, whereas urban workers get penalized more than rural when giving birth to an above-quota number of children. Let ϕ_U^0 be the childrearing cost in urban areas and $\bar{\phi}_U$ be the urban above-quota penalty or per child subsidy. In other words, we restrict our attention to the case of $\phi_U^0 > \phi_R^0$ and $\bar{\phi}_U \geq \bar{\phi}_R$.

Denote $w \in \{w_S, w_P^H, w_P^L\}$ as workers' wage income from working in either the SOE or the private sector based on skill level. An urban worker's budget constraint is:

$$
\begin{aligned}
& c + nb + n\phi_U^0 + \left[I^F + \left(1 - I^F\right) I^T\right] \max\{n - \bar{n}_U, 0\}\, \bar{\phi}_U \\
& + \left[1 - \left[I^F + \left(1 - I^F\right) I^T\right]\right] \max\{n - \bar{n}_R, 0\}\, \bar{\phi}_R \\
= \; & w - \left[I^F + \left(1 - I^F\right) I^T \mu\right] \tau \equiv \tilde{w},
\end{aligned} \tag{13}
$$

where $\tilde{w}$ denotes the urban net (of tax) wage that depends on urban residency status. All urban workers face the same childrearing costs regardless of residency status, albeit the benefits and obligations are associated with one's residency. Thus, an urban worker with urban residency $\left(\left(I^F = 1, I^T = 0\right) \text{ or } \left(I^F = 0, I^T = 1\right)\right)$ faces a budget constraint that can be rewritten as follows:

$$
c + nb + n\phi_U^0 + \max\{n - \bar{n}_U, 0\}\, \bar{\phi}_U = \tilde{w}. \tag{14}
$$

The budget constraint of an urban worker with rural residency $\left(\left(I^F, I^T\right) = (0, 0)\right)$ is:

$$
c + nb + n\phi_U^0 + \max\{n - \bar{n}_R, 0\}\, \bar{\phi}_R = w. \tag{15}
$$

Before moving to the migration decision, we identify the link between the quantity-based altruistic factor (β) and fertility constraint ($\bar{n}_j$). Denote n_j^* as the optimal fertility choice of an agent in location j ($j = U, R$). We are interested in the nondegenerate case: Agents with a high quantity-based altruistic factor ($\beta = \bar{\beta}$) will have $n_j^* > \bar{n}_j$ in which the fertility constraint is binding; those with a low quantity-based altruistic factor ($\beta = \underline{\beta}$) will have $n_j^* < \bar{n}_j$, and the fertility constraint is not binding. We summarize this relationship by using an indicator function I^β ($\beta = \bar{\beta}$) below, where

$$I^\beta = \begin{cases} 1, & \text{for agents with a high quantity-based altruistic factor } (\beta = \bar{\beta}), \\ 0, & \text{for agents with a low quantity-based altruistic factor } (\beta = \underline{\beta}). \end{cases}$$

2.3 Rural-Urban Migration Decision

Rural-born workers decide whether to migrate to cities by comparing the value of staying in rural areas to that of migrating to urban areas. Based on endowed skills and quantity-based altruistic factors, the workers' expected values of migrating to urban areas differ. High-skilled workers have a chance to work in the SOE sector, which immediately grants the workers urban residency. Workers who fail to obtain a job in the SOE sector then work as high-skilled workers in the private sector with rural residency; after staying in cities for a certain duration, the workers will have a chance to obtain urban residency. Low-skilled workers are only competent for low-skilled jobs in the private sector but also have a chance to obtain urban residency. Those who successfully obtain urban residency enjoy urban benefits and are obligated to regulations that come with urban residency.

In the following sections, we characterize rural workers' migration decisions by first defining the value functions of staying in rural areas, working in the SOE sector, and working in the urban private sector for both high- and low-skilled workers. We will then establish the conditions under which a rural worker decides to migrate to cities.

2.3.1 Value function of staying in rural areas

The value function for rural agents to stay in rural areas is independent of their skill levels:

$$\begin{aligned} V^R(\beta) &= \max_{c,b,n} u^R(c,b,n;\beta) \\ &= \max_{c,b,n} \{\min[\theta c, (1-\theta) nb] + \beta n^\varepsilon\} \\ s.t. \quad & c + nb + n\phi_R^0 + \max\{n - \bar{n}_R, 0\} \bar{\phi}_R = x. \end{aligned}$$

To solve $V^R(\beta)$, from the Leontief preference in c and nb, we substitute $c = \frac{(1-\theta)}{\theta} nb$ into the budget constraint, where the quality-based altruistic factor θ is clearly the fraction of income allocated to bequest. The maximization problem becomes:

$$\begin{aligned} V^R(\beta) &= \max_{b,n} \ (1-\theta) nb + \beta n^\varepsilon \\ s.t. \quad & \frac{1}{\theta} nb + n\phi_R^0 + \max\{n - \bar{n}_R, 0\} \bar{\phi}_R = x, \end{aligned}$$

where $\varepsilon < 1$. The first-order conditions for n and b are:

$$(1-\theta) b + \varepsilon \beta n^{\varepsilon - 1} = \lambda \left(\frac{1}{\theta} b + \phi_R^0 + I^\beta \bar{\phi}_R \right), \tag{16}$$

$$(1-\theta) n = \frac{\lambda}{\theta} n, \tag{17}$$

where λ as the Lagrangian multiplier associated with the budget constraint. Using (17) and (16) to substitute out $\lambda = \theta (1-\theta)$ with the fact that the marginal cost (MC) and marginal benefit (MB) from bequests exactly cancel out with each other under Leontief preferences, we obtain a key expression for fertility decisions:

$$\underbrace{\frac{\varepsilon \beta}{n^{1-\varepsilon}}}_{\text{MB of fertility}} = \underbrace{\theta (1-\theta) \left[\phi_R^0 + I^\beta \bar{\phi}_R \right]}_{\text{MC of fertility}}, \tag{18}$$

where the left-hand side and the right-hand side of the above equation are the MB and MC from having an extra child, respectively. With normality, an increase in the quality-based altruistic factor that raises bequest should suppress fertility. We thus impose the following condition:

Condition 2. $\theta < 1/2$.

From (18), it is thus straightforward to derive:

Proposition 1. (Rural Fertility Choice) *Under Condition 2, the optimal rural fertility rate (n_R^*) is increasing with the quantity-based altruistic factor (β), decreasing with the quality-based altruistic factor (θ) or the unit childrearing cost (ϕ_R^0), and independent of income (x). Moreover, while an above-quota fine ($\bar{\phi}_R > 0$) discourages fertility, a childrearing subsidy ($\bar{\phi}_R < 0$) encourages it.*

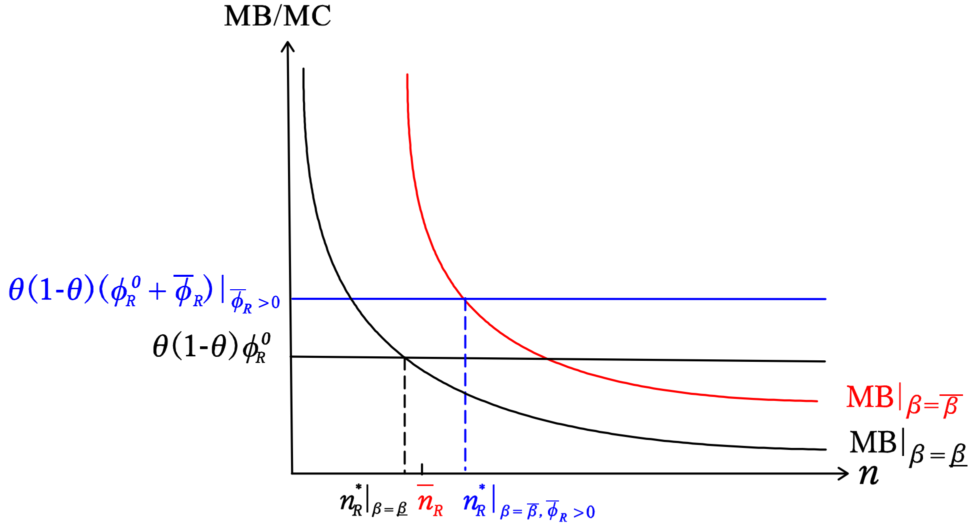
Similar to quasi-linear preferences, a nice property associated with Leontief preferences is that the optimal number of children n_R^* is independent of agents' income and wealth. This enables us to focus exclusively on the typically stronger substitution effect.

Figure 4 offers an informative look at plots of the MB and the MC curves of the fertility decision. Based on the results in Proposition 1, the MB curve of agents with $\bar{\beta}$ lies to the right of the MB curve of agents with $\underline{\beta}$, meaning that agents with stronger fertility preferences will choose to have more children than those with weaker preferences. As previously mentioned, we will focus below on the relevant case where the government's population policy is only binding for agents with stronger fertility preferences. This case is interesting because agents with $\underline{\beta}$ choose $n_R^*|_{\underline{\beta}} \leqslant \bar{n}_R$, whereas agents with $\bar{\beta}$ choose $n_R^*|_{\bar{\beta}} > \bar{n}_R$, by paying the above-quota penalty or receiving the childrearing subsidy. To guarantee that $n_R^*|_{\bar{\beta}} > \bar{n}_R$ under $\beta = \bar{\beta}$, we impose a regularity condition derived from (18):

Condition 3. $\bar{n}_R < \left[\frac{\varepsilon \bar{\beta}}{\theta(1-\theta)\left[\phi_R^0 + I^\beta \bar{\phi}_R\right]} \right]^{\frac{1}{1-\varepsilon}}$.

As a result, the case under consideration features $n_R^*|_{\bar{\beta}} > \bar{n}_R \geq n_R^*|_{\underline{\beta}}$, as drawn in Figure 4.

Figure 4: Determination of Fertility for Rural Workers



We can solve n_R^* analytically from (18) as:

$$n_R^* = \left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\phi_R^0 + I^\beta \bar{\phi}_R\right]} \right]^{\frac{1}{1-\varepsilon}}. \tag{19}$$

The optimal bequest or children's quality b_R^* can then be solved from the budget constraint (11) as:

$$b_R^* = \theta \left[\frac{x}{n_R^*} - \phi_R^0 - I^\beta \left(1 - \frac{\bar{n}_R}{n_R^*}\right) \bar{\phi}_R \right]. \tag{20}$$

Under our setting, total bequest is a constant share of "disposable income" at the end of one's lifetime after the government transfers (fine or subsidy) and childrearing spending (measured by $x - n\phi_R^0 - \max\{n - \bar{n}_R, 0\}\bar{\phi}_R$). Thus, the quantity-quality trade-off is obvious: A higher quantity of children n_R^* is associated with a lower quality measured by b_R^*. Although not the emphasis of our paper, this dimension of quantity-quality trade-off can be combined with Proposition 1 to imply:

Proposition 2. (Rural Child-Quality Investment Choice) *Optimal investment in the quality of children (b_R^*) rises with income (x) and is negatively associated with the quantity of children (n_R^*). The quantity-quality trade-off is stronger in the presence of an above-quota fine ($\bar{\phi}_R > 0$) but weaker in the presence of a childrearing subsidy ($\bar{\phi}_R < 0$).*

In other words, population controls tend to induce a stronger quantity-quality trade-off in fertility choice, but childrearing incentives tend to weaken it.

With n_R^* and b_R^* being pinned down by (18) and (20), the value function $V^R(\beta)$ is solved as:

$$V^R(\beta) = (1-\theta)\theta \left\{ x - n_R^*|_\beta \left[\phi_R^0 + I^\beta \left(1 - \frac{\bar{n}_R}{n_R^*|_\beta}\right) \bar{\phi}_R \right] \right\} + \beta \left(n_R^*|_\beta\right)^\varepsilon.$$

2.3.2 Value functions of urban workers

We start with the urban private sector where the main actions occur, followed by the urban SOE sector. In all cases, fertility choice retains a similar functional form to the rural case, with rural income being replaced by after-tax wage ($\tilde{w}$), childrearing cost being updated to the urban levels ϕ_U^0, and above-quota penalty or subsidy being revised to $\bar{\phi}_U$. To save space, we shall not report these solutions but note that the properties in Propositions 1 and 2 will be carried out for all urban workers of different types in different sectors.

Value function of workers in the private sector

1. High-skilled workers

For high-skilled rural migrants working in the private sector, after staying in urban areas for $(1-\mu)$ of their lifetime, the workers obtain urban residency with a probability ρ ($I^F = 0$, $I^T = 1$). With a probability $(1-\rho)$, the workers fail to do so and hold rural residency throughout life ($I^F = 0$, $I^T = 0$). Denote $V^{P,H}(\beta)$ as the value function of a high-skilled migrant worker in the private sector. $V^{P,H}(\beta)$ can thus be written as:

$$V^{P,H}(\beta) = \rho \left\{ \begin{array}{c} \max\limits_{c,b,n} u^U(c,b,n;\beta)\,|_{I^F=0,I^T=1} \\ s.t.\ c + nb + n\phi_U^0 + \max\{n - \bar{n}_U, 0\}\,\bar{\phi}_U = w_P^H - \mu\tau \end{array} \right\}$$
$$+ (1-\rho) \left\{ \begin{array}{c} \max\limits_{c,b,n} u^U(c,b,n;\beta)\,|_{I^F=0,I^T=0} \\ s.t.\ \ c + nb + n\phi_U^0 + \max\{n - \bar{n}_R, 0\}\,\bar{\phi}_R = w_P^H \end{array} \right\}.$$

Because urban workers' fertility decisions depend on urban residency, we denote separately $n_F^*|_\beta$ and $n_I^*|_\beta$ as the number of children chosen by private-sector workers with and without urban residency (i.e., by formal and informal private-sector workers, respectively). The same regularity condition used in the rural problem is required, and hence, Condition 3 is modified as:

Condition 3′. $\bar{n}_j < \left[\frac{\varepsilon\bar{\beta}}{\theta(1-\theta)\left[\phi_j^0 + I^\beta\bar{\phi}_j\right]}\right]^{\frac{1}{1-\varepsilon}}$, $j = R, U$.

Then, following the same steps as in the rural worker's optimization problem, we can solve:

$$n_F^*|_\beta = \left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\phi_U^0 + I^\beta\bar{\phi}_U\right]}\right]^{\frac{1}{1-\varepsilon}}; n_I^*|_\beta = \left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\phi_U^0 + I^\beta\bar{\phi}_R\right]}\right]^{\frac{1}{1-\varepsilon}}$$

Recall $n_R^*|_\beta$ from (19) and that $\phi_U^0 > \phi_R^0$ and $\bar{\phi}_U \geq \bar{\phi}_R$. It is straightforward to show that the urban fertility choices are always below the rural counterpart, with formal urban workers' fertility lower than that of informal urban workers:

Proposition 3. (Comparison of Optimal Fertility Choices) $n_F^*|_\beta \leq n_I^*|_\beta < n_R^*|_\beta$.

Substitute the solutions of the maximization problem into the value function above. $V^{P,H}(\beta)$,

$\beta = \{\underline{\beta}, \bar{\beta}\}$ can then be written as:

$$V^{P,H}(\beta) = \rho \left\{ \begin{array}{c} (1-\theta)\,\theta \left\{ w_P^H - \mu\tau - n_F^*|_\beta \left[\phi_U^0 + I^\beta \left(1 - \frac{\bar{n}_U}{n_F^*|_\beta} \right) \bar{\phi}_U \right] \right\} \\ +\beta \left(n_F^*|_\beta \right)^\varepsilon + \mu B \end{array} \right\}$$

$$+ (1-\rho) \left\{ \begin{array}{c} (1-\theta)\,\theta \left\{ w_P^H - n_I^*|_\beta \left[\phi_U^0 + I^\beta \left(1 - \frac{\bar{n}_R}{n_I^*|_\beta} \right) \bar{\phi}_R \right] \right\} \\ +\beta \left(n_I^*|_\beta \right)^\varepsilon \end{array} \right\}.$$

2. Low-skilled workers

Denote $V^{P,L}(\beta)$ as the value function for low-skilled migrant workers in the urban private sector, $\beta = \{\underline{\beta}, \bar{\beta}\}$. By applying the same procedure used to derive $V^{P,H}(\beta)$, we write $V^{P,L}(\beta)$ as:

$$V^{P,L}(\beta) = \rho \left\{ \begin{array}{c} (1-\theta)\,\theta \left\{ w_P^L - \mu\tau - n_F^*|_\beta \left[\phi_U^0 + I^\beta \left(1 - \frac{\bar{n}_U}{n_F^*|_\beta} \right) \bar{\phi}_U \right] \right\} \\ +\beta \left(n_F^*|_\beta \right)^\varepsilon + \mu B \end{array} \right\}$$

$$+ (1-\rho) \left\{ \begin{array}{c} (1-\theta)\,\theta \left\{ w_P^L - n_I^*|_\beta \left[\phi_U^0 + I^\beta \left(1 - \frac{\bar{n}_R}{n_I^*|_\beta} \right) \bar{\phi}_R \right] \right\} \\ +\beta \left(n_I^*|_\beta \right)^\varepsilon \end{array} \right\}.$$

Due to the absence of the income effect, low-skilled migrant workers' fertility choices resemble those of high-skilled migrant workers.

Value function of workers in the SOE sector

In many countries, working in the SOE sector requires a higher qualification, e.g. at least holding a high school diploma or passing certain government employee exams. We thus assume that only high-skilled workers can work in the SOE sector. Rural high-skilled agents migrating to cities, with a probability π, are recruited as SOE workers and immediately granted urban residency. These workers then enjoy complete urban benefits and pay full urban taxes. A SOE worker has the following value function ($I^F = 1$):

$$V^S(\beta) = \max_{c,b,n} \left\{ \min\left[\theta c, (1-\theta)\, nb \right] + \beta n^\varepsilon + B \right\}$$

$$s.t. \quad c + nb + n\phi_U^0 + \max\{n - \bar{n}_U, 0\}\, \bar{\phi}_U = w_S - \tau.$$

By substituting in the number of children chosen and the investment in children's quality, we obtain a SOE worker's value function $V^S(\beta)$, $\beta = \{\underline{\beta}, \bar{\beta}\}$:

$$V^S(\beta) = (1-\theta)\,\theta \left\{ w_S - \tau - n_F^*|_\beta \left[\phi_U^0 + I^\beta \left(1 - \frac{\bar{n}_U}{n_F^*|_\beta} \right) \bar{\phi}_U \right] \right\} + \beta \left(n_F^*|_\beta \right)^\varepsilon + B.$$

2.3.3 Migration decisions

We are ready to discuss agents' migration decisions. Migration is costly. To simplify the analysis, we assume that migration cost comes in the form of utility. Depending on skill levels and preferences, rural workers may have different abilities to adapt to urban life. Hence, we assume that moving from rural to urban areas involves a migration cost, $\psi \in \{\psi^L, \psi^H\}$, which is measured in utils, for low- and high-skilled agents, respectively.

Rural agents migrate to cities only if the expected value of migrating to urban areas is higher than the value of staying in the rural. Since high-skilled migrant workers have a probability π to obtain a job in the SOE sector and a probability $1-\pi$ to work in the private sector, these workers will migrate if the following inequality holds true:

$$\Delta V_H(\beta) \equiv \pi V^S(\beta) + (1-\pi) V^{P,H}(\beta) - \psi^H - V^R(\beta) \geq 0, \text{ for } \beta \in \{\underline{\beta}, \bar{\beta}\}. \tag{21}$$

Low-skilled rural workers can only work in the private sector after migrating to urban areas. Hence, these workers will migrate to urban areas if the following inequality is met:

$$\Delta V_L(\beta) \equiv V^{P,L}(\beta) - \psi^L - V^R(\beta) \geq 0, \text{ for } \beta \in \{\underline{\beta}, \bar{\beta}\}. \tag{22}$$

When the above two equations are held with equality, rural agents are indifferent between migrating to urban areas and staying in rural areas. The migration decision thus depends on the relative magnitudes of rural income (x, which is a function of z, q, R^+, and δ) and urban incomes (w_S, w_P^H, w_P^L), the relative childrearing costs in urban and rural areas (ϕ_R^0 and ϕ_U^0), population policies (governed by $\bar{n}_R$, $\bar{n}_U$, $\bar{\phi}_R$, and $\bar{\phi}_U$), urban benefits (B), urban tax (τ), and easiness of obtaining urban residency and urban benefits (π, ρ, and μ).[9] Define an indicator function with $I^m = 1$ if rural-urban migration takes place, and $I^m = 0$ otherwise. Then we have:

$$I^m = \begin{cases} 1, & \text{iff } \Delta V_i(\beta) \geq 0, i = H, L, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

There are four types of agents in the model: type-$\{H, \underline{\beta}\}$, type-$\{H, \bar{\beta}\}$, type-$\{L, \underline{\beta}\}$, and type-$\{L, \bar{\beta}\}$. Agents of different types will have different migration decisions depending on the relevant quantity-based altruistic factors and agent income levels, and the model thus has several possible equilibrium outcomes. As a result, the evolution of workers and the supply of urban amenities under each of these equilibrium outcomes will also differ, which we now turn to discuss.

3 Equilibrium

Prior to defining the equilibrium, we delineate the concept of a mixed migration equilibrium that restricts our attention to the most plausible migration pattern, based on which we specify the evolution of workers and the supply of urban amenities. A dynamic competitive migration equilibrium is subsequently defined, followed by a steady-state migration equilibrium.

[9] The probability π captures the likelihood of instant urban residency after migration. Otherwise, getting urban residency becomes uncertain, forcing migrants to wait. The probability ρ captures the likelihood of delayed urban residency after migration with the fraction of delay time captured by $1-\mu$. These possible outcomes on urban residency and benefits are summarized by the parameters (π, ρ, μ).

3.1 Mixed Migration Equilibrium

Generally, high-skilled agents have higher motivation to migrate to cities than low-skilled workers as skills are more rewarding in urban areas. For low-skilled workers in developing countries, wage incomes in cities are usually higher than the incomes from farming. However, the costs of raising children in terms of housing, spacing, and tuition are usually higher in cities, and agents who prefer to have more children may be more prone to stay in rural areas. Therefore, we choose to confine our attention to a specific migration equilibrium: Type-$\{H, \underline{\beta}\}$ workers always choose to migrate to cities, type-$\{L, \bar{\beta}\}$ workers always decide to stay in rural areas, and type-$\{H, \bar{\beta}\}$ and type-$\{L, \underline{\beta}\}$ agents are indifferent between migrating and staying so that some stay in rural areas while others migrate to cities. This is the most relevant case to study in developing countries at the stage of relatively low urbanization.

Denote Γ_H as the fraction of rural high-skilled workers with the high quantity-based altruistic factor (i.e., type-$\{H, \bar{\beta}\}$) being indifferent between migrating and staying but ultimately moving to cities, and $(1 - \Gamma_H)$ as the fraction of such workers staying in their rural hometowns. Similarly, denote Γ_L as the fraction of the rural low-skilled workers with low quantity-based altruistic factor (i.e., type-$\{L, \underline{\beta}\}$) being indifferent between migrating to cities and staying in rural areas, but ultimately moving to cities, and $(1 - \Gamma_L)$ as the fraction of such workers staying in rural areas. The migration patterns for rural high- and low-skilled workers of the equilibrium on which we focus are shown below (with formal definition and equilibrium conditions being relegated to Section 3.4):

	$\underline{\beta}$ ζ of workers	$\bar{\beta}$ $(1-\zeta)$ of workers
H	move	Γ_H move $(1 - \Gamma_H)$ stay
L	Γ_L move $(1 - \Gamma_L)$ stay	stay

Note that, if $0 < \Gamma_H < 1$ and $0 < \Gamma_L < 1$, we have the ***mixed migration equilibrium (MME)*** – "mixed" in the sense that both a positive fraction of high- and low-skilled workers migrate to cities in the equilibrium. If $\Gamma_H = 1$ and $0 < \Gamma_L < 1$, we have a skilled-based ***segregated migration equilibrium (SME)***: All high-skilled workers migrate, while low-skilled workers with low quantity-based altruistic factor $\underline{\beta}$ are indifferent between migrating and staying, and low-skilled workers with high quantity-based altruistic factor $\bar{\beta}$ always stay. Another fertility-based SME exists with $\Gamma_H = 0$ and $0 < \Gamma_L < 1$: In this case, all agents with high quantity-based altruistic factor (type-$\bar{\beta}$) stay in rural areas, while low-skilled workers with quantity-based altruistic factor $\underline{\beta}$ are indifferent between migrating and staying, and high-skilled workers with quantity-based altruistic factor $\underline{\beta}$ always move. In a special case with $\Gamma_H = 1$ and $\Gamma_L = 0$, we have a pure skilled-based SME with all high-skilled workers migrating to cities and all low-skilled workers remaining as farmers. In another polar case with $\Gamma_H = 0$ and $\Gamma_L = 1$, we have a pure fertility-based SME with all workers with quantity-based altruistic factor $\underline{\beta}$ migrating to cities and all with quantity-based altruistic factor $\bar{\beta}$ remaining as farmers. Different types of SMEs are readily summarized below:

	Γ_H	Γ_L
Pure skill-based SME	1	0
Skill-based SME	1	$\in (0,1)$
Fertility-based SME	0	$\in (0,1)$
Pure fertility-based SME	0	1

We regard the last case as theoretically possible but not realistically likely – given that the typical observation is that cities are more attractive than farmland to rural high-skilled workers. For the sake of brevity, we will thus not include the case in our theoretical analysis.

3.2 Evolution of Workers

As mentioned at the beginning of Section 2, agents have one-period lifetime and make migration decisions within that period, and hence, the beginning-of-the-period population and the actual-working population stocks differ, with the differences resulting from the inflows of workers. Because children inherit their parents' status of residency, if rural migrant workers do not successfully obtain urban residency before exiting the market, in the next period their children will start life from rural areas with rural residency. Similarly, if starting life in urban areas, workers hold urban residency bequeathed by their parents.

Denote U (R) as the beginning-of-the-period workers with urban (rural) residency, then the beginning-of-the-period population identity equations in urban and rural areas are:

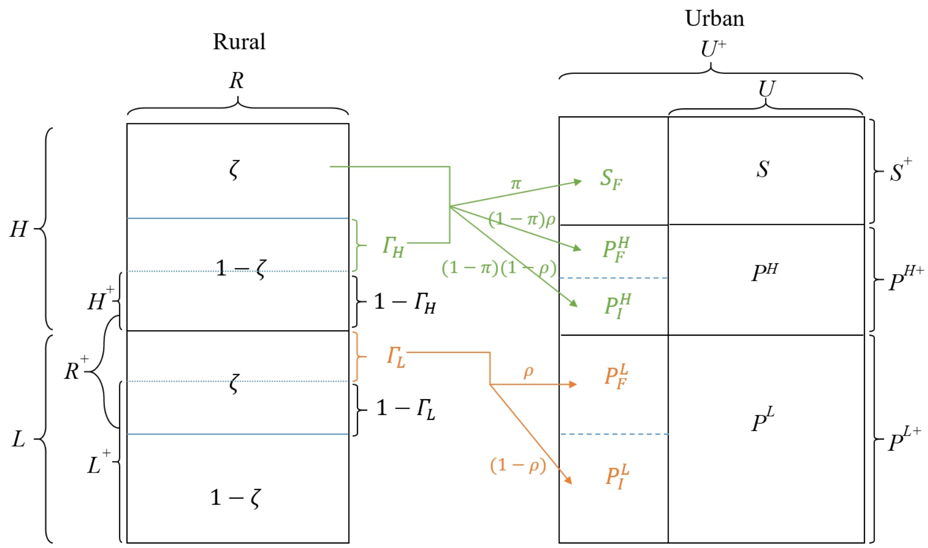
$$U = S + P^H + P^L, \tag{24}$$

$$R = H + L. \tag{25}$$

whereas S, P^H, and $P^L$are the beginning-of-the-period SOE, private-sector high- and low-skilled workers, respectively; H and L are the beginning-of-the-period rural high- and low-skilled workers, respectively. We use two subscripts, F and I, for these workers to indicate new migrants in a period. The former (F) represents *formal* workers whose residency matches the location, whereas the latter (I) identifies *informal* workers whose residency does not match the location. Thus, denote S_F, P_F^H, and P_F^L as the new migrants who successfully obtain urban residency, working as SOE, private-sector high- and low-skilled workers, respectively. Likewise, denote P_I^H and P_I^L as the new comers working as private-sector high- and low-skilled workers but failing to obtain urban residency. The detailed notations of population flows are summarized in Table 1. Figure 5 provides the population flow chart in the model.

[Insert Table 1 about here]

Figure 5: Population Flow Chart



The actual-working populations in urban areas under the equilibrium we examine are:

$$S^+ = S + S_F = S + \underbrace{\left[\zeta + (1-\zeta)\,\Gamma_H\right]\pi H}_{\text{obtain urban residency immediately upon arrival}}, \tag{26}$$

$$\begin{aligned} P^{H+} &= P^H + P_F^H + P_I^H \\ &= P^H + \underbrace{\left[\zeta + (1-\zeta)\,\Gamma_H\right](1-\pi)\,\rho H}_{\text{obtain urban residency}} + \underbrace{\left[\zeta + (1-\zeta)\,\Gamma_H\right](1-\pi)\,(1-\rho)\,H}_{\text{still hold rural residency}} \\ &= P^H + \left[\zeta + (1-\zeta)\,\Gamma_H\right](1-\pi)\,H, \end{aligned} \tag{27}$$

$$\begin{aligned} P^{L+} &= P^L + P_F^L + P_I^L \\ &= P^L + \underbrace{\zeta\Gamma_L\rho L}_{\text{obtain urban residency}} + \underbrace{\zeta\Gamma_L\,(1-\rho)\,L}_{\text{still hold rural residency}} \\ &= P^L + \zeta\Gamma_L L, \end{aligned} \tag{28}$$

where the reader is reminded that $Z_J\,(J = F, I)$ denotes the inflow of Z within the period. Total number of agents working in urban areas after migration inflows is thus equal to:

$$\begin{aligned} U^+ &= S^+ + P^{H+} + P^{L+} \\ &= S + P^H + P^L + \left[\zeta + (1-\zeta)\,\Gamma_H\right]H + \zeta\Gamma_L L \\ &= U + \left[\zeta + (1-\zeta)\,\Gamma_H\right]H + \zeta\Gamma_L L. \end{aligned} \tag{29}$$

Similarly, the actual-working populations in rural areas after migration outflows are:

$$R^+ = H^+ + L^+ = (1-\zeta)\,(1-\Gamma_H)\,H + (1-\zeta\Gamma_L)\,L. \tag{30}$$

As children inherit the skill levels (i.e., jobs) and residency status directly from parents and all agents live for one period, the evolutions of workers in the SOE and the private sectors are:

$$\begin{aligned}
S' &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] S + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, \Gamma_H n_F^*|_{\bar{\beta}}\right] \pi H, && (31)\\
P^{H'} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] P_F^H + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, \Gamma_H n_F^*|_{\bar{\beta}}\right] (1-\pi)\, \rho H, && (32)\\
P^{L'} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] P_F^L + \zeta \Gamma_L n_F^*|_{\underline{\beta}} \rho L, && (33)
\end{aligned}$$

where Z' denotes the next period value of Z. The evolution equation for U can be written as follows:

$$U' = S' + P^{H'} + P^{L'}. \tag{34}$$

The evolution equations for rural high- and low-skilled workers and total rural workers with rural residency can be written accordingly:

$$\begin{aligned}
H' &= H\left\{(1-\rho)(1-\pi)\left[\zeta n_I^*|_{\underline{\beta}} + (1-\zeta)\, \Gamma_H n_I^*|_{\bar{\beta}}\right] + (1-\zeta)(1-\Gamma_H)\, n_R^*|_{\bar{\beta}}\right\}, && (35)\\
L' &= L\left\{(1-\rho)\, \zeta \Gamma_L n_I^*|_{\underline{\beta}} + \zeta (1-\Gamma_L)\, n_R^*|_{\underline{\beta}} + (1-\zeta)\, n_R^*|_{\bar{\beta}}\right\}, && (36)\\
R' &= H' + L'. && (37)
\end{aligned}$$

3.3 Urban Benefits

Urban benefits, inclusive of urban amenities (e.g., parks, schools, museums, libraries, medical services, childcare and old age allowances, and other public services) are assumed to be financed by urban embarked taxes. Total urban taxes collected by the government are:

$$\begin{aligned}
T &= \left\{S + P^H + P^L + [\zeta + (1-\zeta)\, \Gamma_H]\, \pi H\right\} \tau && (38)\\
&\quad + \left\{[\zeta + (1-\zeta)\, \Gamma_H]\, (1-\pi)\, \rho H + \zeta \Gamma_L \rho L\right\} \mu\tau.
\end{aligned}$$

The government provides the urban benefits B uniformly to all residents with urban status:

$$B = B_0 G, \tag{39}$$

where G is the per capita budget for amenities and benefits, and B_0 is the government's technology scaling factor in the provision of urban amenities and benefits. Assume that the government runs a balanced budget in every period. Then the periodic balanced government budget implies:

$$G = \frac{T}{U + S_F + P_F^H + P_F^L}.$$

3.4 Migration Equilibrium

In equilibrium, all urban labor markets clear with labor demands given by (5)-(7):

$$S^d = S^+,\ P^{H,d} = P^{H+},\ P^{L,d} = P^{L+}, \tag{40}$$

where S^d, $P^{H,d}$ and $P^{L,d}$ are the labor demands in specific sectors. The rural labor market clears under the rural farming income given by (2):

$$R^d = R^+. \tag{41}$$

where R^d is the demand for rural labor.

We define the competitive equilibrium of the model below.

Definition. *A **dynamic competitive migration equilibrium** (DCME) of the model consists of migration decisions, rural farming income x, and urban wage rates $\{w_S, w_P^H, w_P^L\}$, such that*

(i) *(Optimization) given rural farming income x and urban wage rates $\{w_S, w_P^H, w_P^L\}$, based on their residency status, agents choose numbers of children according to (18) and (20); furthermore, rural high- and low-skilled agents make migration decisions according to (23);*

(ii) *(Market clearing) rural farming income satisfies (2), urban wage rates $\{w_S, w_P^H, w_P^L\}$ satisfy (5), (6), and (7), and labor markets clear according to (40) and (41);*

(iii) *(Urban amenities) the amenities in urban areas are supplied according to (39);*

(iv) *(Workers laws of motion) given the initial population $\{H^0, L^0, S^0, P^{H,0}, P^{L,0}\}$, high- and low-skilled workers in rural, SOE, and urban private sectors evolve according to (31)-(37), with workers actually devoted to production given by (26)-(30).*

We next define the balanced-growth equilibrium for the rest of our analysis.

Definition. *A **balanced-growth migration equilibrium (BGME)** of the model is a DCME when the growth rates of population variables, Z'/Z, are all constant ($Z = S, R, P^H$,P^L).*

Along a BGP, from (31)-(33), we get:

$$\frac{S'}{S} = \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \left[\zeta\pi n_F^*|_{\underline{\beta}} + (1-\zeta)\,\pi\Gamma_H n_F^*|_{\bar{\beta}}\right]\frac{H}{S}, \tag{42}$$

$$\frac{P^{H\prime}}{P^H} = \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \left[\rho\zeta\,(1-\pi)\, n_F^*|_{\underline{\beta}} + \rho\,(1-\zeta)\,(1-\pi)\,\Gamma_H n_F^*|_{\bar{\beta}}\right]\frac{H}{P^H}, \tag{43}$$

$$\frac{P^{L\prime}}{P^L} = \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \left(\rho\zeta\Gamma_L n_F^*|_{\underline{\beta}}\right)\frac{L}{P^L}. \tag{44}$$

Thus, along a BGP, we have a constant ratio of SOE workers to rural high-skilled workers (S/H) and constant ratios for workers of different skill types in the urban private sector (P^H/H and P^L/L). Next, from (35)-(36), we have constant growth rates of different skill-type workers. These rates, together with the fact that the growth rate of R is constant, give us a constant ratio of rural high- to low-skilled workers (H/L) from (37). This finding in turn implies that the ratio of high-skilled to low-skilled workers in the urban private sector is constant (P^H/P^L). Finally, from (24), we can show that the growth rate of U, and its ratio to high-skilled workers are both constant.

We are now ready to establish the property of common growth and, most importantly, the existence of the properties in a mixed migration equilibrium.

Proposition 4. (Common Growth) *At the BGME, the model exhibits a common growth rate property:* $\frac{Z'}{Z} = g$ *with* $Z = S, U, R, H, L, P^H$ *and* P^L.

We now present the first main finding that concerns the existence of a MME.

Theorem 1. (Mixed Migration Equilibrium) *Under proper ranges of migration costs and urban-rural income differentials, a mixed migration equilibrium arises along a balanced growth path. Moreover, a better urban amenity, a larger childrearing subsidy, a lower urban-rural childrearing cost differential, a lower urban fertility penalty, or a more secured rural land entitlement tends to result in a higher fraction of type-*$\{H, \bar{\beta}\}$ *and type-*$\{L, \underline{\beta}\}$ *agents choosing to migrate from rural to urban.*

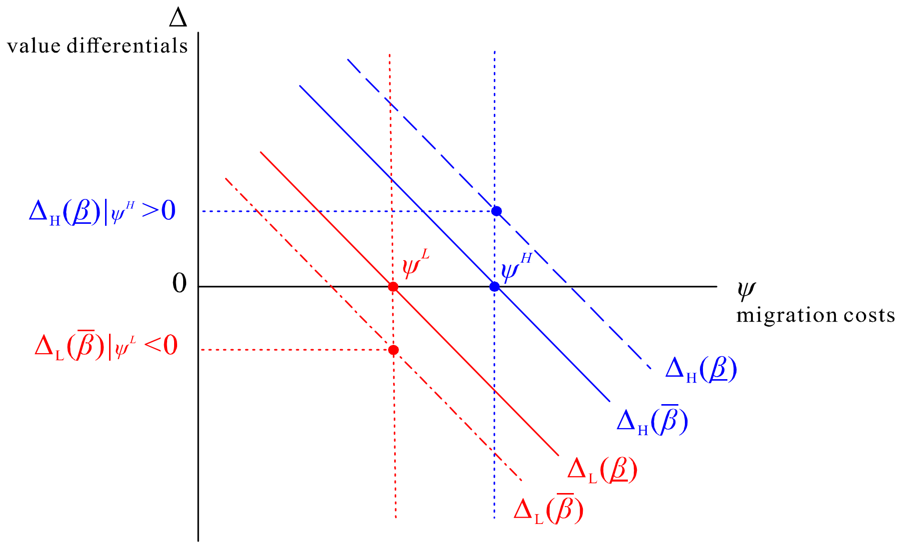
While the formal proof is relegated to Appendix IIA.1, Theorem 1 is established by means of proof by construction. To do so, we define a general class of indifference boundaries for the rural high- and low-skilled agents with type-β facing a migration cost ψ:

$$\begin{aligned} \Delta V_H(\beta, \psi) &\equiv \pi V^S(\beta) + (1-\pi) V^{P,H}(\beta) - \psi - V^R(\beta) = 0 \\ \Delta V_L(\beta, \psi) &\equiv V^{P,L}(\beta) - \psi - V^R(\beta) = 0 \end{aligned}$$

which are both linear and decreasing in the migration cost. To begin, we use (21) and (22) to identify conditions under which type-$\{H, \underline{\beta}\}$ workers always choose to migrate to cities, whereas type-$\{L, \bar{\beta}\}$ always decide to stay in rural areas, i.e., $\Delta V_H\left(\underline{\beta}, \psi^H\right) > 0 > \Delta V_L\left(\bar{\beta}, \psi^L\right)$. This task is trivial because, for any pair of quantity-based altruistic factors $\left\{\underline{\beta}, \bar{\beta}\right\}$, we can always adjust the migration costs of migration $\left\{\psi^L, \psi^H\right\}$ for the two inequalities to hold. Basically, this requires that ψ^L is sufficiently high but ψ^H is sufficiently low.

The major task is therefore to check the indifference boundaries of type-$\{H, \bar{\beta}\}$ and type-$\{L, \underline{\beta}\}$ agents that may lead to an interior solution so that $\Delta V_H\left(\bar{\beta}, \psi^H\right) = 0 = \Delta V_L\left(\underline{\beta}, \psi^L\right)$. To do so, we first write all the costs (migration, childrearing, and penalties or subsidies) in proportion to wage incomes. Then it is convenient to define urban-rural income differentials measured by the ratios of urban net wage to farmer's income: $\varsigma_H(\beta) \equiv \frac{\tilde{w}_P^H(\beta)}{x}$, $\varsigma_S(\beta) \equiv \frac{\tilde{w}_S(\beta)}{x}$, and $\varsigma_L(\beta) \equiv \frac{\tilde{w}_P^L(\beta)}{x}$. A straightforward examination of the value functions suggests that an increase in $\varsigma_H(\beta)$ or $\varsigma_S(\beta)$ tends to shift up the high-skilled indifference boundary $\Delta V_H\left(\beta, \psi^H\right) = 0$, whereas an increase in $\varsigma_L(\beta)$ tends to shift up the low-skilled indifference boundary $\Delta V_L\left(\beta, \psi^L\right) = 0$. Thus, with proper urban-rural income differentials, we can assure $\Delta V_H\left(\bar{\beta}, \psi^H\right) = 0 = \Delta V_L\left(\underline{\beta}, \psi^L\right)$, under which a fraction of both high- and low-skilled workers $(\Gamma_H, \Gamma_L) \in (0,1) \times (0,1)$ migrates to cities. This proof by construction is illustrated by Figure 6 given $\psi^L < \psi^H$ (which is for illustrative purposes only and otherwise inessential for the proof). Importantly, the parametric space supporting a mixed migration equilibrium is dense and hence nonempty. Using similar arguments, we can perform a comparative static analysis to establish that urban amenities and childrearing subsidies serve as positive forces for migration (shifting up indifference boundaries), whereas childrearing cost differentials, fertility penalties, and rural land entitlement requirements serve as negative forces (shifting down indifference boundaries).

Figure 6: Indifference Boundaries in MME



From the indifference boundaries, we get a novel locational quantity-quality trade-off for rural-urban migration, reflected by the presence of an expected locational quality gain and an expected fertility loss upon migrating to cities:

$$\Delta V_H(\beta,\psi) = \underbrace{\begin{array}{c}(1-\theta)\,\theta\left\{\pi w_S + (1-\pi)\,w_P^H - [\pi + (1-\pi)\,\rho\mu]\,\tau - x\right\} \\ + [\pi + (1-\pi)\,\rho\mu]\,B - \psi\end{array}}_{\text{expected locational quality gain } (+)}$$
$$\underbrace{+(1-\varepsilon)\,\beta\left\{\begin{array}{c}\pi\,(n_F^*|_\beta)^\varepsilon - (n_R^*|_\beta)^\varepsilon \\ + (1-\pi)\left[\rho\,(n_F^*|_\beta)^\varepsilon + (1-\rho)\,(n_I^*|_\beta)^\varepsilon\right]\end{array}\right\}}_{\text{expected fertility loss (-)}},$$

$$\Delta V_L(\beta,\psi) = \underbrace{(1-\theta)\,\theta\left[\rho\left(w_P^L - \mu\tau\right) + (1-\rho)w_P^L - x\right] + \rho\mu B - \psi}_{\text{expected locational quality gain } (+)}$$
$$+\underbrace{(1-\varepsilon)\,\beta\left[\rho\,(n_F^*|_\beta)^\varepsilon + (1-\rho)\,(n_I^*|_\beta)^\varepsilon - (n_R^*|_\beta)^\varepsilon\right]}_{\text{expected fertility loss (-)}}.$$

Again, the formal proof is relegated to Appendix IIA.2, but this property is readily summarized as:

Theorem 2. (Locational quantity-quality trade-off) *At the BGME, the rural-to-urban migration decision features a locational quantity-quality trade-off between payoffs from locational quality and fertility.*

The intuition underlying Theorem 2 is not difficult to get. Urban areas provide better job opportunities and amenities but have higher rearing costs per child. An urban migrant tends to reduce the number of children but instead invest more in the quality of children and bring the children to better careers and living environments.

4 Generalization

Before turning to quantitative analyses, however, we would like to generalize the baseline framework in four dimensions of particular interest: reverse migration, left-behind children, upward skill mobility, and directed urban benefits.

4.1 Reverse Migration

In this subsection, we allow for reverse migration (RM) by introducing a "general" urban residency shock. Specifically, the urban residency shock leads to relocating a constant fraction of newborns of generation-t parents from urban to rural occurring at the beginning of $t+1$. Denote the fraction of reverse migration as Λ. The beginning-of-the-period population identity equations after the reverse migration shock in urban areas are:

$$\begin{aligned} S^{RM} &= (1-\Lambda) S, \\ P^{RM,H} &= (1-\Lambda) P^{H}, \\ P^{RM,L} &= (1-\Lambda) P^{L}. \end{aligned}$$

Therefore, the beginning-of-the-period urban and rural population are given by:

$$\begin{aligned} U^{RM} &= S^{RM} + P^{RM,H} + P^{RM,L} = (1-\Lambda)\left(S + P^{H} + P^{L}\right), \\ R^{RM} &= R + \Lambda U = H + L + \Lambda\left(S + P^{H} + P^{L}\right). \end{aligned}$$

In Appendix IIB.1, we update the actual-working populations in both urban and rural areas after migration occurs and the evolutions of workers in all sectors, from which we obtain, along a BGP,

$$\begin{aligned} \frac{S^{RM\prime}}{S^{RM}} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \frac{\pi}{1-\Lambda}\left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right]\left[\Lambda + \left(1 + \Lambda\frac{P^H}{H}\right)\frac{H}{S}\right] \\ \frac{P^{RM,H\prime}}{P^{RM,H}} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \frac{(1-\pi)\,\rho}{1-\Lambda}\left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right]\left[\Lambda + \left(1 + \Lambda\frac{S}{H}\right)\frac{H}{P^H}\right] \\ \frac{P^{RM,L\prime}}{P^{RM,L}} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \frac{1}{1-\Lambda}\left(\rho\zeta\Gamma_L n_F^*|_{\underline{\beta}}\right)\left(\frac{L}{P^L} + \Lambda\right) \end{aligned}$$

Thus, at the BGME, we obtain common growth when we have a constant ratio of SOE workers to high-skilled workers in the urban private sector (S/P^H). Under this proportionality condition, the common growth property established in Proposition 4 remains valid. Moreover, it can be easily verified that the indifference boundaries remain unchanged. Finally, the fertility and migration decisions characterized in Propositions 1-3, which reflect the locational quantity-quality trade-off at the BGME, continue to hold true.

4.2 Left-Behind Children

We next study the implications of the left-behind children (LBC) of migrant workers. Specifically, we examine a case close to the migration pattern observed where all high-skilled migrants have children with urban residency and all low-skilled migrants have children without urban residency. That case is depicted as follows:

$$\rho^{LBC,i} = \begin{cases} 1, & \text{for all } i = H, \\ 0, & \text{for all } i = L. \end{cases}$$

In particular, in (27), (32), (35), $V^{P,H}(\beta)$, and $\Delta V_H(\beta,\psi)$, we set $\rho^{LBC,H} = 1$, while in (28), (33), (36), $V^{P,L}(\beta)$, and $\Delta V_L(\beta,\psi)$, we have $\rho^{LBC,L} = 0$.[10]

Because children of low-skilled migrants are left in rural areas, the childrearing cost must be adjusted. Although the rural childrearing cost is lower, remote parenting is costly. It is thus realistic to set the childrearing cost of left-behind children as $\Xi\phi_R^0$ where $\Xi > 1$ reflects a cost markup with $\Xi\phi_R^0 < \phi_U^0$ so that a net cost savings exists. The fertility decision of low-skilled migrants is simply given by:

$$n_I^{LBC*}|_\beta = \left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\Xi\phi_R^0 + I^\beta\bar{\phi}_R\right]}\right]^{\frac{1}{1-\varepsilon}} > n_I^*|_\beta.$$

In other words, the fertility rate of low-skilled migrants is higher than that in the baseline case. Notably, when offered urban residency, high-skilled migrants choose to take it.

We again leave the details in Appendix IIB.2. Because the effect fully applies to migrants, it is trivial that the fertility choice and the value function facing a rural stayer, denoted $n_R^{LBC*}|_\beta$ and $V^{LBC,R}(\beta)$, remain the same: $V^{LBC,R}(\beta) = V^R(\beta)$ and $n_R^{LBC*}|_\beta = n_R^*|_\beta$. In addition, due to being immediately granted urban residency, SOE workers' value functions remain the same. The main effects of left-behind children fall on the value functions of urban workers in the private sector.

For high-skilled rural migrants working in the private sector, after staying in urban areas for $(1-\mu)$ of their lifetime, the migrants obtain urban residency ($\rho^{LBC,H} = 1$). Thus, the workers' fertility choice is simply $n_F^{LBC*}|_\beta = n_F^*|_\beta$. This implies:

Proposition 5a. (Comparison of Fertility Choices) $n_F^{LBC*}|_\beta = n_F^*|_\beta \leq n_I^*|_\beta < n_R^*|_\beta, \beta = \{\underline{\beta}, \bar{\beta}\}$.

In Appendix IIB.2, we compare the values $V^{LBC,P,H}(\beta)$ and $V^{P,H}(\beta)$ to obtain:

[10] Previous studies have found that migrant household family income in the destination is significantly associated with migrant parents' arrangements for their children. Therefore, in the generalization for left-behind children, we assume that all low-skilled migrant parents cannot obtain urban *hukou* and leave their children behind. See Appendix IC for more detailed information.

$$V^{LBC,P,H}(\beta) - V^{P,H}(\beta)$$
$$= (1\text{-}\rho)\left\{ \underbrace{\left[\begin{array}{c} \mu\left[B\text{-}(1\text{-}\theta)\,\theta\tau\right] \\ +(1\text{-}\theta)\,\theta \left\{ \begin{array}{c} n_I^*|_\beta \left[\phi_U^0 + I^\beta \left(1\text{-}\frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right] \\ \text{-}n_F^*|_\beta \left[\phi_U^0 + I^\beta \left(1\text{-}\frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right] \end{array} \right\} \end{array} \right]}_{\text{fertility quality gain}} + \underbrace{\beta\left[(n_F^*|_\beta)^\varepsilon\text{-}(n_I^*|_\beta)^\varepsilon\right]}_{\text{fertility quantity loss}} \right\}$$

since $B > \tau$ and $n_F^*|_\beta < n_I^*|_\beta$. Thus, the net outcome features a locational quantity-quality trade-off.

In addition, it is straightforward to show:

Proposition 5b. (Comparison of Fertility Choices) Under $1 < \Xi < \phi_U^0/\phi_R^0$, $n_F^*|_\beta \le n_I^*|_\beta < n_I^{LBC*}|_\beta < n_R^*|_\beta$.

In other words, because of net childrearing cost savings, their fertility rates are higher than the counterpart of migrants in the informal sector under the baseline setting.

By the same procedure for deriving $V^{LBC,P,H}(\beta)$, together with $\rho^{LBC,L} = 0$ and $\Xi > 1$, we can compare the value functions to conclude:

$$\begin{aligned} V^{LBC,P,L}(\beta) - V^{P,L}(\beta) &= -\rho\mu\left[B - (1-\theta)\,\theta\tau\right] \\ &+ (1-\theta)\,\theta \left\{ \begin{array}{c} -n_I^{LBC*}|_\beta \left[\Xi\phi_R^0 + I^\beta\left(1 - \frac{\bar{n}_R}{n_I^{LCB*}|_\beta}\right)\bar{\phi}_R\right] \\ +\rho n_F^*|_\beta\left[\phi_U^0 + I^\beta\left(1 - \frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right] \\ +(1-\rho)n_I^*|_\beta\left[\phi_U^0 + I^\beta\left(1 - \frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right] \end{array} \right\} \\ &+\beta\left[\left(n_I^{LBC*}|_\beta\right)^\varepsilon - \rho\left(n_F^*|_\beta\right)^\varepsilon - (1-\rho)\left(n_I^*|_\beta\right)^\varepsilon\right] \end{aligned}$$

Under $1 < \Xi < \phi_U^0/\phi_R^0$ and the fertility ranking established in Proposition 5b, the second term is negative and the third positive. Thus, there is a fertility quality loss (sum of the first two terms) and a fertility quantity gain (the third term). Thus, the ranking depends, again, on the locational quality-quantity trade-off.

From the above value functions under left-behind children, we first summarize:

Proposition 5c. (Comparison of Value Functions) If the locational quality effect outweighs the quantity effect, then $V^{LBC,P,H}(\beta) > V^{P,H}(\beta)$ and $V^{LBC,P,L}(\beta) < V^{P,L}(\beta)$.

Intuitively, based on the revealed preference argument, urban benefits must exceed childrearing cost savings. Given this requirement, one would expect the locational quality effect to outweigh the quantity effect. Whether the locational quality effect is strong enough in reality is nonetheless a quantitative question to be addressed in Section 5.

From Proposition 5c, the indifference boundaries under left-behind children satisfy

$$\Delta V_H^{LBC}\left(\beta,\psi^H\right) \equiv \pi V^S(\beta) + (1-\pi)\,V^{LBC,P,H}(\beta) - \psi^H - V^R(\beta) > \Delta V_H\left(\beta,\psi^H\right), \text{ for } \beta \in \{\underline{\beta},\bar{\beta}\}.$$

Low-skilled rural workers will migrate to urban areas if the following inequality is met:

$$\Delta V_L^{LBC}\left(\beta,\psi^L\right) \equiv V^{LBC,P,L}\left(\beta\right) - \psi^L - V^R\left(\beta\right) < \Delta V_L\left(\beta,\psi^L\right), \text{ for } \beta \in \{\underline{\beta},\bar{\beta}\}.$$

So we arrive at:

Proposition 5d. (The Indifference Boundaries) If the locational quality effect outweighs the quantity effect, then the $\Delta V_H^{LBC}\left(\beta,\psi^H\right) = 0$ locus shifts up and the $\Delta V_L^{LBC}\left(\beta,\psi^L\right) = 0$ locus shifts down.

In general, the migration decision depends on the relative magnitudes of rural incomes (x, a function of z, q, R^+ and δ) and urban incomes (w_S, w_P^H, w_P^L), the relative childrearing costs in urban and rural areas (ϕ_R^0 and ϕ_U^0), population policies (governed by $\bar{n}_R$, $\bar{n}_U$, $\bar{\phi}_R$, and $\bar{\phi}_U$), urban benefits (B), urban tax (τ), and easiness of obtaining urban residency and urban benefits (π and μ).

Finally, as shown in Appendix IIB.2, along a BGP, we have:

$$\begin{aligned}
\frac{S^{LBC\prime}}{S} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right]\pi\frac{H}{S}, \\
\frac{P^{LBC,H\prime}}{P^H} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right](1-\pi)\frac{H}{P^H}, \\
\frac{P^{LBC,L\prime}}{P^L} &= \left[\zeta n_I^{LBC*}|_{\underline{\beta}} + (1-\zeta)\, n_I^{LBC*}|_{\bar{\beta}}\right] + \zeta\Gamma_L n_I^{LBC*}|_{\underline{\beta}}\frac{L}{P^L}.
\end{aligned}$$

Thus, at the BGME, we obtain common growth under the left-behind children scenario.

4.3 Upward Skill Mobility

In the baseline framework, we assume children inherit the skill levels of parents. One may ask what will happen if we allow for intergenerational upward skill mobility (USM) in the sense that children of high-skilled parents brought up to urban areas are high-skilled, while children of urban low-skilled parents have a probability ϑ to become high-skilled.

In this scenario, while S' remains unchanged, the actual-working populations in urban areas under the equilibrium we examine are:

$$\begin{aligned}
P^{USM,H+} &= \underbrace{P^H + \vartheta P^L}_{P^{USM,H}} + \underbrace{\left[\zeta + (1-\zeta)\,\Gamma_H\right](1-\pi)\,H}_{P_F^H + P_I^H}, \\
P^{USM,L+} &= \underbrace{(1-\vartheta)\, P^L}_{P^{USM,L}} + \underbrace{\zeta\Gamma_L L}_{P_F^L + P_I^L}.
\end{aligned}$$

In other words, a share of low-skilled children (ϑP^L) is upgraded to being high-skilled. From the evolutions of workers in the SOE and private sectors derived in Appendix IIB.3, we get, along a

BGP,

$$
\begin{aligned}
\frac{S^{USM\prime}}{S} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right] \pi \frac{H}{S} \\
\frac{P^{USM,H\prime}}{P^{USM,H}} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right] (1-\pi) \frac{H/P^H}{1+\vartheta\left(P^L/P^H\right)} \\
&\quad + \left(\zeta \Gamma_L n_F^*|_{\underline{\beta}} \rho \frac{L}{P^L}\right) \frac{\vartheta\left(P^L/P^H\right)}{1+\vartheta\left(P^L/P^H\right)} \\
\frac{P^{USM,L\prime}}{P^{USM,L}} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right] + \zeta \Gamma_L n_F^*|_{\underline{\beta}} \rho \frac{L}{P^L}
\end{aligned}
$$

Thus, at the BGME, we obtain common growth under upward skill mobility.

Although the common growth property remains valid under upward skill mobility, a general equilibrium wage effect shifts the indifference boundaries of migration. Since $P^{USM,H+} = P^{H+} + \vartheta P^L$ and $P^{USM,L+} = P^{L+} - \vartheta P^L$, it is straightforward to see that a negative general equilibrium labor endowment affects high-skilled wages and a positive effect impacts low-skilled wages:

$$
\begin{aligned}
w_P^{USM,H} &= \alpha \eta^\sigma A_P \left[\alpha \eta^\sigma + (1-\alpha)\left(\frac{P^{L+} - \vartheta P^L}{P^{H+} + \vartheta P^L}\right)^\sigma\right]^{\frac{1-\sigma}{\sigma}} < w_P^H, \\
w_P^{USM,L} &= (1-\alpha)\, A_P \left[\alpha\left(\eta \frac{P^{H+} + \vartheta P^L}{P^{L+} - \vartheta P^L}\right)^\sigma + (1-\alpha)\right]^{\frac{1-\sigma}{\sigma}} > w_P^L
\end{aligned}
$$

Because of changes in wages, a dynamic migration effect encourages more rural low-skilled but fewer rural high-skilled agents to migrate. Given the Leontief preferences, the fertility choices are not affected under upward skill mobility. However, the value functions of the urban private-sector workers are altered due to the changes in wages. Specifically, we have

$$
\begin{aligned}
V^{USM,P,H}(\beta) - V^{P,H}(\beta) &= (1-\theta)\,\theta\left(w_P^{USM,H} - w_P^H\right) < 0 \\
V^{USM,P,L}(\beta) - V^{P,L}(\beta) &= (1-\theta)\,\theta\left(w_P^{USM,L} - w_P^L\right) > 0
\end{aligned}
$$

We first summarize:

Proposition 6a. (Comparison of Value Functions) $V^{USM,P,H}(\beta) < V^{P,H}(\beta)$ and $V^{USM,P,L}(\beta) > V^{P,L}(\beta)$.

As a result, the indifference boundaries of the urban private-sector workers are affected as follows:

$$
\begin{aligned}
\Delta V_H^{USM}\left(\beta, \psi^H\right) &\equiv \pi V^S(\beta) + (1\text{-}\pi)\, V^{USM,P,H}(\beta)\text{-}\psi^H\text{-}V^R(\beta) < \Delta V_H\left(\beta, \psi^H\right), \text{ for } \beta \in \{\underline{\beta}, \bar{\beta}\}, \\
\Delta V_L^{USM}\left(\beta, \psi^L\right) &\equiv V^{USM,P,L}(\beta)\text{-}\psi^L\text{-}V^R(\beta) > \Delta V_L\left(\beta, \psi^L\right), \text{ for } \beta \in \{\underline{\beta}, \bar{\beta}\}.
\end{aligned}
$$

We thus conclude:

Proposition 6b. (The Indifference Boundaries) Under upward skill mobility, the $\Delta V_H^{USM}\left(\beta, \psi^H\right) = 0$ locus shifts down and the $\Delta V_L^{USM}\left(\beta, \psi^L\right) = 0$ locus shifts up.

Thus, the (indirect) wage effect of migration incentive lowers the urban skill composition, which works against the direct positive effect of upward skill mobility. In sum, the net effect of upward skill mobility on urban skill composition is ambiguous.

4.4 Directed Urban Benefits

In the baseline framework, all urban benefits are summarized by an additional utility term B. That is, it can be viewed as public goods and services (net of any negative externalities).[11] One may, however, argue that some urban benefits can be directed (DB): (i) to human capital at work, denoted as B_W, that affects both w_S and w_P^H in the budget constraints of formal urban high-skilled workers; (ii) to child education, B_E, that affects nb in the preferences of all formal urban workers; and (iii) to child care, B_C, that affects ϕ_U^0 in the budget constraints of all formal urban workers. We denote pure amenities by B_A, just like with B in the baseline case.

The government technologies are captured by $B_m = B_{0,m} G_m$, where $B_{0,m}$ is the government's technology scaling factor in the provision of urban amenities with the benefits of type $m = W, E, C, A$, and G_m is the per capita budget for amenities such that $\sum_m G_m = G$, where G is the per capita budget for all urban amenities and benefits.

An urban worker maximizes

$$\begin{aligned} u^{DB,U}(c,b,n;\beta)|_{I^F,I^T} &= \min\left\{\theta c, (1-\theta) nb \left[1 + \left(I^F + \left(1-I^F\right) I^T \mu\right) B_E\right]\right\} \\ &\quad + \beta n^{\varepsilon} + \left[I^F + \left(1-I^F\right) I^T \mu\right] B_A. \end{aligned}$$

subject to:

$$\begin{aligned} & c + nb + n\left[\phi_U^0 - \left(I^F + \left(1-I^F\right) I^T \mu\right) B_C\right] + \left[I^F + \left(1-I^F\right) I^T\right] \max\left\{n - \bar{n}_U, 0\right\} \bar{\phi}_U \\ & + \left\{1 - \left[I^F + \left(1-I^F\right) I^T\right]\right\} \max\left\{n - \bar{n}_R, 0\right\} \bar{\phi}_R \\ = \; & w + I^H \left[I^F + \left(1-I^F\right) I^T \mu\right] B_W - \left[I^F + \left(1-I^F\right) I^T \mu\right] \tau, \end{aligned}$$

where $I^H = 1$ for high-skilled workers (0 otherwise) and $w \in \{w_S, w_P^H, w_P^L\}$. All urban workers face the same childrearing costs regardless of residency status, albeit the benefits and obligations are associated with one's residency.

As shown in Appendix IIB.4, for SOE workers $\left(I^F = 1, I^T = 0\right)$, the fertility choice is:

$$n_F^{DB,S*}|_\beta = \left[\frac{1 + (1-\theta) B_E}{1 + B_E} \frac{\varepsilon \beta}{\theta (1-\theta) \left(\phi_U^0 - B_C + I^\beta \bar{\phi}_U\right)}\right]^{\frac{1}{1-\varepsilon}}$$

As a result, we have:

$$\begin{aligned} n_F^{DB,S*}|_\beta &> n_F^*|_\beta, B_C > 0 = B_E \\ n_F^{DB,S*}|_\beta &< n_F^*|_\beta, B_E > 0 = B_C \end{aligned}$$

This is intuitive because B_E promotes child quality but B_C encourages child quantity. Moreover,

[11] More discussions are relegated to Appendix IC.

we compare value functions with the baseline to obtain:

$$
\begin{aligned}
& V^{DB,S}(\beta) - V^{S}(\beta) \\
= \; & \theta(1-\theta)\frac{1+B_E}{1+(1-\theta)B_E}\left\{w_S + B_W - \tau - n_F^{DB,S*}|_\beta\left[\phi_U^0 - B_C + I^\beta\left(1-\frac{\bar{n}_U}{n_F^{DB,S*}|_\beta}\right)\bar{\phi}_U\right]\right\} \\
& -(1-\theta)\theta\left\{w_S - \tau - n_F^*|_\beta\left[\phi_U^0 + I^\beta\left(1-\frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\} \\
& +\beta\left[\left(n_F^{DB,S*}|_\beta\right)^\varepsilon - \left(n_F^*|_\beta\right)^\varepsilon\right] \\
\geq \; & 0, \text{ for } n_F^{DB,S*}|_\beta \approx n_F^*|_\beta
\end{aligned}
$$

In sum, for SOE workers, directed urban benefits are likely to bring in a higher locational quality gain (the difference of the first two terms in the second equality). Their effects on fertility quantity, however, are ambiguous (the last difference term).

Similarly, for private-sector high-skilled workers with urban residency $\left(I^F = 0, I^T = 1\right)$, we get

$$
n_F^{DB,P,H*}|_\beta = \left[\frac{1+(1-\theta)\mu B_E}{1+\mu B_E}\frac{\varepsilon\beta}{\theta(1-\theta)\left(\phi_U^0 - \mu B_C + I^\beta\bar{\phi}_U\right)}\right]^{\frac{1}{1-\varepsilon}},
$$

but for private-sector high-skilled workers with rural residency $\left(I^F = 0, I^T = 0\right)$, the value function is identical to the baseline case and hence $n_I^{DB,P,H*}|_\beta = n_I^*|_\beta$. We can then conclude:

Proposition 7a. (Comparison of Fertility Choices)

$$
\begin{aligned}
n_F^{DB,S*}|_\beta &> n_F^{DB,P,H*}|_\beta > n_I^*|_\beta \geq n_F^*|_\beta, B_C > I^\beta\left(\bar{\phi}_U - \bar{\phi}_R\right) \geq 0 = B_E \\
n_F^{DB,S*}|_\beta &< n_F^{DB,P,H*}|_\beta < n_F^*|_\beta \leq n_I^*|_\beta, B_E > 0 = B_C
\end{aligned}
$$

On the one hand, if the urban benefit directed to child care B_C is large enough to cover the gap between the locational above-quota fines, then the fertility choices are quantitatively larger than the benchmark case. On the other hand, in the presence of the directed urban benefit to child education B_E, the locational quality-quantity trade-off is magnified; the fertility choices are also quantitatively smaller than the baseline case.

We can also compare the value functions to obtain (details in Appendix IIB.4):

$$
\begin{aligned}
& V^{DB,P,H}(\beta) - V^{P,H}(\beta) \\
= \; & \rho\theta(1-\theta)\frac{1+\mu B_E}{1+(1-\theta)\mu B_E}\cdot \\
& \left\{w_P^H + \mu B_W - \mu\tau - n_F^{DB,P,H*}|_\beta\left[\phi_U^0 - \mu B_C + I^\beta\left(1-\frac{\bar{n}_U}{n_F^{DB,P,H*}|_\beta}\right)\bar{\phi}_U\right]\right\} \\
& -\rho(1-\theta)\theta\left\{w_P^H - \mu\tau - n_F^*|_\beta\left[\phi_U^0 + I^\beta\left(1-\frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\} \\
& +\beta\left[\left(n_F^{DB,P,H*}|_\beta\right)^\varepsilon - \left(n_F^*|_\beta\right)^\varepsilon\right] \\
\geq \; & 0, \text{ for } n_F^{DB,P,H*}|_\beta \approx n_F^*|_\beta
\end{aligned}
$$

In sum, for private-sector high-skilled workers, directed urban benefits are likely to bring in a higher locational quality gain but have an ambiguous effect on fertility quantity.

Next, for private-sector low-skilled workers with urban residency $\left(I^F = 0, I^T = 1\right)$, we get:

$$n_F^{DB,P,L*}|_\beta = \left[\frac{1 + (1-\theta)\,\mu B_E}{1 + \mu B_E} \frac{\varepsilon \beta}{\theta (1-\theta) \left(\phi_U^0 - \mu B_C + I^\beta \bar{\phi}_U\right)}\right]^{\frac{1}{1-\varepsilon}} = n_F^{DB,P,H*}|_\beta$$

For private-sector low-skilled workers with rural residency $\left(I^F = 0, I^T = 0\right)$, the value function is identical to the baseline case. We can thus expand Proposition 7a to:

Proposition 7b. (Comparison of Fertility Choices)

$$\begin{aligned} n_F^{DB,S*}|_\beta &> n_F^{DB,P,H*}|_\beta = n_F^{DB,P,L*}|_\beta > n_I^*|_\beta \geq n_F^*|_\beta, B_C > I^\beta \left(\bar{\phi}_U - \bar{\phi}_R\right) \geq 0 = B_E \\ n_F^{DB,S*}|_\beta &< n_F^{DB,P,H*}|_\beta = n_F^{DB,P,L*}|_\beta < n_F^*|_\beta \leq n_I^*|_\beta, B_E > 0 = B_C \end{aligned}$$

Accordingly, we show the following in Appendix IIB.4:

$$\begin{aligned} V^{DB,P,L}(\beta) - V^{P,L}(\beta) &\propto \theta(1-\theta)\frac{1+\mu B_E}{1+(1-\theta)\,\mu B_E} \cdot \\ &\quad \left\{ w_P^L + \mu B_W - \mu\tau - n_F^{DB,P,L*}|_\beta \left[\phi_U^0 - \mu B_C + I^\beta \left(1 - \frac{\bar{n}_U}{n_F^{DB,P,L*}|_\beta}\right)\bar{\phi}_U\right]\right\} \\ &\quad - (1-\theta)\,\theta \left\{ w_P^L - \mu\tau - n_F^*|_\beta \left[\phi_U^0 + I^\beta\left(1 - \frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\} \\ &\quad + \beta\left[\left(n_F^{DB,P,L*}|_\beta\right)^\varepsilon - \left(n_F^*|_\beta\right)^\varepsilon\right] \\ &\geq 0, \text{ for } n_F^{DB,P,L*}|_\beta \approx n_F^*|_\beta \end{aligned}$$

Similar to the case of high-skilled workers, for private-sector low-skilled workers, directed urban benefits are likely to bring in a higher locational quality gain but have an ambiguous effect on fertility quantity.

Remark 1. ($B_C = 0$) In the absence of directed benefit to childrearing, all other directed benefits generate the locational quality-quantity trade-off by lowering fertility quantity but raising fertility quality.

As a result, under $n_F^{DB,P,H*}|_\beta = n_F^{DB,P,L*}|_\beta \approx n_F^*|_\beta$, the indifference boundaries of the urban private-sector workers are affected as follows:

$$\begin{aligned} \Delta V_H^{DB}\left(\beta, \psi^H\right) &\equiv \pi V^{DB,S}(\beta) + (1-\pi)\,V^{DB,P,H}(\beta) - \psi^H - V^R(\beta) \geq \Delta V_H\left(\beta, \psi^H\right), \\ \Delta V_L^{DB}\left(\beta, \psi^L\right) &\equiv V^{DB,P,L}(\beta) - \psi^L - V^R(\beta) \geq \Delta V_L\left(\beta, \psi^L\right), \end{aligned}$$

for $\beta \in \{\underline{\beta}, \bar{\beta}\}$. This implies:

Proposition 7c. (The Indifference Boundaries) If the effects of directed benefits fall mainly on the locational quality instead of fertility quantity, i.e., $n_F^{DB,P,H*}|_\beta = n_F^{DB,P,L*}|_\beta \approx n_F^*|_\beta$, then the $\Delta V_i^{DB}\left(\beta, \psi^i\right) = 0$ locus ($i = H, L$) shifts up so that more migration takes place from rural to urban.

In sum, the effects of directed benefits on rural-to-urban migration are likely to be positive.

Remark 2. (Evolution of Workers) The evolution equations of population are affected quantitatively but not qualitatively because the effects of directed benefits only come from a change in fertility choices. Thus, at the BGME, we continue to obtain common growth.

Remark 3. (Institutional Variation in Directed Benefits) Our general principle of modelling directed benefits focuses on the agent's residency status. In practice, our setup may be modified when country-specific institutional restrictions are considered. For instance, when we adopt China data in the quantitative exercises below, the directed benefits to child care (B_C) only apply to SOE workers.[12] The obvious consequence of this modification is that the fertility choices of all urban workers are reduced to the level of the benchmark informal workers:

$$n_J^{DB,P,H*}|_\beta = n_J^{DB,P,L*}|_\beta = n_I^*|_\beta, \ J = F, I, \text{ when } B_C > I^\beta \left(\bar{\phi}_U - \bar{\phi}_R\right) \geq 0 = B_E \tag{45}$$

As a result, Propositions 7a and 7b continue to hold with minor modifications according to (45) for the case of $B_C > 0 = B_E$. Finally, this quantitative China case simply diminishes the positive effect of directed benefits to child care on fertility quantity (by reducing $n_J^{DB,P,H*}|_\beta$ and $n_J^{DB,P,L*}|_\beta$), so Proposition 7c is strengthened.

Remark 4. (Urban Status Effect) Recall that θ measures the quality-based altruistic factor, which is proportional to the net income allocating to total bequests. An urban status effect can potentially be captured by $\theta_U > \theta_R = \theta$. To see this, we recall Condition 2, under which we have the following: $\theta_U (1 - \theta_U) > \theta_R (1 - \theta_R) = \theta (1 - \theta)$. It is therefore straightforward to show that the urban status effect deepens the locational quality-quantity trade-off by raising the fertility quality (b) and reducing the fertility quantity (n). Moreover, if the urban status effect on fertility choice falls mainly on the locational quality instead of locational quantity, then the indifference boundary $\Delta V_i^{US} \left(\beta, \psi^i\right) = 0$ locus ($i = H, L$) shifts up so that more migration takes place from rural to urban. In other words, the implications of the urban status effect are qualitatively identical to the directed benefit to children education (B_E).

4.5 Taking Stock

In sum, urban areas provide better job opportunities and amenities, but at the expense of higher childrearing costs and rural land entitlement losses. In the baseline setting, we establish and characterize a dynamic competitive migration equilibrium of particular interest – the mixed migration equilibrium with a positive fraction of rural high- and a positive fraction of rural low-skilled workers migrating to cities in the equilibrium. We show that a better urban amenity, a larger childrearing subsidy, a smaller urban-rural childrearing cost differential, a lower urban fertility penalty, or a more secured rural land entitlement encourages rural-to-urban migration. We also show that the rural-to-urban migration decision features a locational quantity-quality trade-off along the BGP, along which an urban migrant tends to reduce the number of children and instead invest more in the quality of children for the sake of better education and better living.

[12] See Appendix IC for more in-depth details.

By generalizing the baseline framework to allow for reverse migration, left-behind children, upward skill mobility, or directed urban benefits, we show that the locational quantity-quality trade-off property is generally maintained. Such generalizations, however, do lead to richer but complicated new channels. With reverse migration, should the allocation of urban high-skilled workers to SOE and private sectors stay unchanged, the locational quality effect is expected to outweigh the quantity effect. With left-behind children, a similar dominant locational quality effect is also expected. With upward skill mobility, an indirect wage effect lowers the urban skill composition, thus working against the direct positive effect, so the net effects of upward skill mobility on urban skill composition and migration outcomes are generally ambiguous. Finally, the four different channels of directed benefits under consideration are all likely to induce rural-to-urban migration under standard regularity conditions.

Just how important the locational quantity-quality trade-off would be, how strong each of the underlying drivers is, and whether any of the generalized channels beyond the baseline framework may play a nonnegligible role are, nonetheless, quantitative questions to which we now turn.

5 Quantitative Analysis

In this section, we quantify our theoretical model, taking China as an example. China offers an interesting case due to the country's tight migration regulation and population control in past decades. These policies - the *hukou* regulation system, the migration-related land policy, and the one-child policy - not only provide costs and benefits to decision-making, but also interconnect to jointly shape an individual's choices on fertility and rural-urban migration. Our task is thus to investigate and quantify how these policies influence the macroeconomic performance in China.

Below, we first provide a brief overview of these policies while relegating the details to Appendix IB. We then calibrate our model to fit the data from China during 1980-2007. Based on the calibrated benchmark economy, we conduct various counterfactual policy experiments to study how they impact fertility decisions, rural-urban migration patterns, and economic development.

5.1 Institutional Background

China's migration control is based on the household registration system, *hukou*, which has been in effect since its implementation in 1950. Under the *hukou* regulation, permission is required for formal rural-urban migration, and annual quotas on migrants are controlled by the government. While those joining the SOE sector (via the channel of *zhaogong* and *zhaogan*) immediately obtain urban *hukou*, some private-sector migrants move without urban residency (temporary migrants) or are later granted urban *hukou*. Labor markets therefore become locationally segmented. Notably, another channel exists via attending college (*zhaosheng*, as studied in Liao, Wang, Wang, and Yip 2022), which is not the focus of this paper.

A migration-related policy is the reallocation of rural land. In rural areas, people are bound to land, and the distribution of land use rights, or land entitlement, is mainly based on household size. Land, however, is officially owned by village collectives (communal land). As a result, rural-urban migrants cannot sell the entitled land, which is then left idle. In addition, village collectives have the power to reallocate the land use rights according to changes in household size. Rural-urban migrants thus face risk losing their land entitlement. Even if land tenure security reform formally sets the effective duration of a land contract at 30 years and land transfers are allowed, the transfer contracts are still largely informal and the problem of land expropriation remains – the problem is indeed exacerbated due to urban expansion and infrastructure development.

China's well-known one-child policy also interplays with the *hukou* system, thereby affecting migration decisions. As the one-child policy was implemented by local governments, the population control is differential across locations and sectors. For example, the penalties on above-quota births in rural areas could be about 10-20 percent of family income lasting for 3-14 years, while SOE workers may lose the eligibility for promotion, be demoted, or be forced to quit their jobs. In addition, rural-urban migrants granted urban *hukou* are subject to the fertility restrictions of the migrants' destinations, but temporary migrants are regulated by the rules of the migrants' place of origin. As a result, those with higher fertility preferences may have a lower incentive to migrate from rural to urban areas and are less inclined to work in the SOE sector. These preferences create an interactive channel through which migration and fertility decisions are interconnected – the channel we explore in this paper.

5.2 Calibration

The period under examination is 1980-2007. In other words, we focus on the period after China's economic reform in 1980 but before the financial tsunami. Most rural workers migrate to cities at young ages. To calibrate the theoretical model, we therefore assume that agents enter the economy at age 18, remain economically active for 36 years, and exit the economy (or retire) at age 54.[13] Thus, the model period in the calibration is set at 36 years. We categorize those with a senior high school degree or above as high-skilled workers and others as low-skilled workers. Besides, to better capture various above-quota penalties implemented by local governments, we assume low-skilled workers bear the basic locational above-quota fines ($\bar{\phi}_U$ and $\bar{\phi}_R$), while high-skilled workers in different sectors and locations bear the extra above-quota fines ($\bar{\phi}_H$, $\bar{\phi}_S$, and $\bar{\phi}_P^H$). The details of above-quota penalties will be discussed later. In what follows, we first summarize the parameters and variables that we will calibrate and solve. Second, we describe the population ratios required for calibration. Third, we discuss the procedure for determining the parameters and variables, which occurs either via calibration or via being taken directly from the literature or data. Detailed information on data sources and the methods used to impute our targets are relegated to Appendix IC.

There are 31 parameters and variables to be calibrated and determined, including (i) preference parameters: the quality-based altruistic factor θ, the quantity-based altruistic factor $\underline{\beta}$ and $\bar{\beta}$, the utility concavity in the quantity of children ε, and the migration costs for high- and low-skilled workers, ψ^H and ψ^L; (ii) the proportion of agents who get less enjoyment from having children, ζ;

[13] China is a country in which workers retire at an early age. See Appendix IC for further discussion on the pattern of early retirement in China.

(iii) childrearing costs, above-quota penalties, and fertility constraint parameters: ϕ_U^0, ϕ_R^0, $\bar{\phi}_U$, $\bar{\phi}_R$, $\bar{\phi}_H$, $\bar{\phi}_S$, $\bar{\phi}_P^H$, $\bar{n}_U$, and $\bar{n}_R$; (iv) parameters related to urban benefits and urban embarked tax – μ, B_0, and τ; (v) production technology parameters: A_S, A_P, α, σ, η, z, δ, and q; (vi) probabilities of obtaining urban *hukou* and position in the SOE sector, ρ and π; and (vii) fractions of type-$\{H, \bar{\beta}\}$ and type-$\{L, \underline{\beta}\}$ workers migrating to urban areas, Γ_H and Γ_L.

Two groups of population ratios are needed for calibration, including the beginning-of-the-period population ($\frac{U}{R}$, $\frac{H}{L}$, $\frac{S}{P}$, $\frac{S}{U}$, and $\frac{P^H}{P^L}$) and the workers who actually work within the period ($\frac{U}{U^+}$, $\frac{P^{H+}}{P^{L+}}$, $\frac{S^+}{P^{H+}}$, and $\frac{R^+}{R}$). We compute the ratio of $\frac{U}{R}$ using population data for urban and rural residence, and obtain an average of 0.4579 during 1980-2007. To compute the ratio of $\frac{H}{L}$, we need information on rural workers' education levels. As the data became available since 1985, we back out the ratios in 1980-1984 using the growth rate of the $\frac{H}{L}$ ratios from 1985 to 2007. The average of the $\frac{H}{L}$ ratio is 0.1076 during 1980-2007. The $\frac{S}{P}$ and $\frac{S}{U}$ ratios are calculated using the data from *China Statistical Yearbook*. We define employees in state-owned units to be workers in the SOE sector and employees in other units in cities and towns to be workers in urban private sectors. The average of the $\frac{S}{P}$ ratio during 1980-2007 is 1.5235 and $\frac{S}{U}$ is 0.6094. The final beginning-of-the-period population ratio needed is the $\frac{P^H}{P^L}$ ratio. To compute this ratio, we need workers' information on two dimensions: employed sectors and education levels. As no suitable ready-for-use data exist, we resort to data from the Urban Household Survey (UHS). We first distinguish private workers from those working in the SOE sector. Second, those who have already retired but re-entered the workforce are excluded. Then, using our definition of high-skilled workers, we obtain a series of the $\frac{P^H}{P^L}$ ratios during 1987-2007. Third, we back out the $\frac{P^H}{P^L}$ ratios for 1980-1986 based on the geometric growth rate of the $\frac{P^H}{P^L}$ ratios over the period of 1987-2007. The average ratio of $\frac{P^H}{P^L}$ during 1980-2007 is 0.8159.

The second group of population ratios consists of workers who actually work within the period. The first ratio needed is $\frac{U}{U^+}$. The urban population (U) is directly obtained from the 2010 *China Population and Employment Statistics Yearbook*. The difference between U and U^+ is the rural-urban floating population. As the term "*mangliu*", or the so-called "blind flow", refers to massive migrants fluxing from rural areas to cities, rural-urban migrants account for most of China's internal migration. Thus, it is reasonable to assume the entire floating population in data is rural-urban migrants. The floating population data is taken from the Department of Population and Employment Statistics, National Bureau of Statistics of China. Using the data, we first compute the ratios of floating to urban population. For years without data, intrapolation is implemented. Then, the ratios of floating to urban population are converted to the ratios of $\frac{U}{U^+}$, and the average of $\frac{U}{U^+}$ during 1980-2007 is 0.8764. To compute the number of workers who actually work in each sector, we must know in which sectors the floating population actually works. However, the information is very limited, especially for early years. According to the Rural-Urban Migrant Survey (RUMS), the proportion of migrant workers employed as production workers, service workers, private enterprise owners, or self-employed reached 93 percent in 2007, while 81 percent of them were employed in private enterprises or self-employed. Based on this information, we can infer that roughly 7 percent of migrants worked in the SOE sector in 2007, while at most 12 percent were private business owners. As it was relatively inflexible for the SOE sector to hire migrants in early years and our data period spans 1980-2007, the 7 percent of migrants working in the SOE sector clearly seems a natural upper bound. We thus set 5 percent of the entire floating population ending up with jobs in the SOE sector, and set 20 percent and 75 percent of the floating population were employed as high- and low-skilled workers in private sectors, respectively. With this assumption, the beginning-of-the-period population ratios and the population identity equations, we can compute the ratios of $\frac{P^{H+}}{P^{L+}}$, $\frac{S^+}{P^{H+}}$, and $\frac{R^+}{R}$. The average of $\frac{P^{H+}}{P^{L+}}$ during 1980-2007 is 0.6348, the average of $\frac{S^+}{P^{H+}}$ becomes 3.0262, and the average of $\frac{R^+}{R}$ is 0.9354.

We are now ready to calibrate the model to data from China. The rural per capita income y_R, which is simply equal to x in equation (2), is normalized to 1 during 1980-2007. In rural China, land use rights are distributed based on family size and reallocated every 5-10 years. Hence, the average duration for land reallocation is 7.5 years. Adjusted by the model period, we obtain $\delta = 0.2083$. *China Statistical Yearbook* reports the average rural land per person. We use the average of 1980-2007 in the calibration and $q = 2.2564$ (mou). As shown in the figure in Appendix IC, average land per person q was stable over the period of 1983-2007, assuring that our use of an average measure is reasonable.[14] Accordingly, the farming technology z can be computed from equation (2) and is equal to 0.4202.

To pin down the parameters in urban production sectors, we need three income ratios from data: $\frac{w_P^H}{w_S}$ (private-sector premium), $\frac{w_P^H}{w_P^L}$ (skill premium), and $\frac{y_U}{y_R}$ (urban premium).[15] We use the UHS to calculate the first two wage ratios and the data from *China Statistical Yearbook* to compute the urban premium. For the years without data, intrapolation is implemented to impute the corresponding values. Thus, we obtain $\frac{w_P^H}{w_P^L} = 1.3944$, $\frac{w_P^H}{w_S} = 1.1346$, and $\frac{y_U}{y_R} = 1.9641$ for the averages of 1980-2007.[16] We set $\sigma = 0.8333$, which corresponds to a value of 6 for the elasticity of substitution (EIS) between high- and low-skilled workers. This EIS value falls within the range of EIS estimates for East Asian countries.[17] By setting the CES production share of high-skilled workers α at 0.5, η is calibrated to match the skill premium, $\frac{w_P^H}{w_P^L}$, and is equal to 1.3608. The technological scaling factors of the SOE sector and private firms, A_S and A_P, are jointly calibrated by matching the urban premium $\frac{y_U}{y_R}$ and private-sector premium $\frac{w_P^H}{w_S}$. A_S and A_P are equal to 1.4854 and 3.3287, respectively.

We next turn to decide the four proportions or probabilities, ζ, π, Γ_H and Γ_L. First, we compute ζ based on the survey data of family size preference by categorizing agents preferring less than or equal to one child as $\underline{\beta}$-type agents and those desiring more than one child as $\bar{\beta}$-type agents. We obtain $\zeta = 0.15$.[18] The remaining three probabilities or proportions of migration (π, Γ_H, and Γ_L)

[14] Due to conversion of land to the household responsibility system (HRS), a big jump in the time series of land per person occurred in 1983. Lin (1992) indicates that the reform was completed in 1983 and that the proportion of production teams that had adopted the HRS stayed at nearly 100 percent after 1983. This fact supports our assumption on stable q in the calibration. See Appendix IC for further details.

[15] Urban per capita income y_U is defined as the average output per worker in urban areas, i.e., $y_U = \frac{Y_S + Y_P}{U^+}$, while rural per capita income is simply $y_R = x = 1$.

[16] This implies the average private-sector premium (w_P^H / w_S) is about 13 percent in our calibration. See Appendix IC for a more detailed discussion.

[17] See Appendix IC for a more detailed discussion.

[18] See Appendix IC for the detailed description on the survey data of family size preference.

are jointly calibrated to match the ratios of $\frac{P^{H+}}{P^{L+}}$, $\frac{S^{+}}{P^{H+}}$, and $1-\frac{U}{U^{+}}$ over 1980-2007. The calibrated values for π, Γ_H, and Γ_L are 0.2, 0.019, and 0.3576, respectively. The implied fraction of high-skilled movers is 16.6 percent, three times as high as that of low-skilled movers (5.36 percent). Thus, skill sorting in our calibrated MME is consistent with Gollin, Lagakos, and Waugh (2014), who find robustly urban workers to have higher skills measured by years of schooling using data from 151 countries.

Regarding the parameters for the institution of the *hukou* system, μ and ρ, we determine their values from the literature. Prior to 1994, it was very difficult for rural migrants to obtain urban *hukou*. After 1994, rural migrants have been able to get urban *hukou*, usually within 2-5 years, via the blue-stamp system. We thus assume that $\mu = 0$ prior to 1994 and that $\mu = \frac{2+5}{2} = 3.5$ years for a migrant to obtain urban residency and hence be qualified for urban benefits. Therefore, the average μ, the fraction of lifetime that a migrant worker successfully obtained urban residency to enjoy urban benefits, is $1 - \frac{0+3.5}{2}/36 = 0.4514$ for the period of 1980 to 2007.[19] As for ρ, based on field interviews, only about 11 percent of the interviewees from rural areas successfully obtained urban residency, so we set ρ at 0.11.[20]

We now turn to determine the fertility constraints in rural and urban areas, $\bar{n}_U$ and $\bar{n}_R$, and the basic above-quota fines, $\bar{\phi}_R$ and $\bar{\phi}_U$. According to Ebenstein (2010), the average fertility quotas per couple in urban and rural China were one child and 1.6 children, respectively. Since our theoretical framework is a unisex model, to link the data to the model, we follow the standard method in the endogenous fertility literature to divide the fertility quota per family by two. We thus set $\bar{n}_U$ and $\bar{n}_R$ to 0.5 and 0.8, respectively. The implementation of the one-child policy varied across regions and sectors. The above-quota penalty could be assessed in either monetary fines or non-monetary forms. For example, workers in the SOE sector could lose jobs for violating the fertility policy. As no perfect method exists to impute the total above-quota penalties paid by parents, we assume that low-skilled workers bear the basic above-quota fines, $\bar{\phi}_R$ for rural low-skilled workers and $\bar{\phi}_U$ for urban low-skilled workers. High-skilled workers bear extra sector-specific fines (in addition to basic fines), which are $\bar{\phi}_H$, $\bar{\phi}_S$, and $\bar{\phi}_P^H$ for rural high-skilled workers, urban employees in the SOE sector, and urban high-skilled workers in the private sector, respectively. The three extra fines born by high-skilled workers are discussed in the next paragraph. Here, we focus on the determination of the basic above-quota fines. We denote $\tilde{\phi}_R$ as the proportion of the basic above-quota fine to a rural agent's wage and $\tilde{\phi}_U$ as the proportion to an urban agent's wage. Using the penalty data provided by Ebenstein (2010), we obtain $\tilde{\phi}_R$ =1.5928 and $\tilde{\phi}_U$ =1.6654.[21]

The determinations of the two childrearing costs (ϕ_U^0 and ϕ_R^0) and the three extra above-quota penalties born by high-skilled workers ($\bar{\phi}_H$, $\bar{\phi}_S$, and $\bar{\phi}_P^H$) are summarized as follows. First, as with the basic above-quota penalties, we define $\tilde{\phi}_U^0$, $\tilde{\phi}_R^0$, $\tilde{\phi}_H$, $\tilde{\phi}_S$, and $\tilde{\phi}_P^H$ as the proportions of childrearing cost (or extra fine) to an agent's wage. Second, we set $\tilde{\phi}_U^0$ at 25 percent of an urban agent's income,

[19] The period we examine is 1980-2007, and 1980-1994 roughly amounts to half of the examined period.

[20] See Appendix IC for more information on the determination for μ and ρ.

[21] In other words, a rural low-skilled worker pays a total above-quota fine of $\tilde{\phi}_R$; a rural high-skilled worker pays $\tilde{\phi}_R + \tilde{\phi}_H$; an urban low-skilled worker pays the rate $\tilde{\phi}_U$; a worker in the SOE sector pays the rate $\tilde{\phi}_U + \tilde{\phi}_S$; and an urban high-skilled worker in the private sector pays the rate $\tilde{\phi}_U + \tilde{\phi}_P^H$.

a common value in urban China.[22] Third, the remaining four parameters ($\tilde{\phi}_R^0$, $\tilde{\phi}_H$, $\tilde{\phi}_S$, and $\tilde{\phi}_P^H$) are calibrated by matching four fertility targets: n_R^L, n_R^H, n_S, and n_P^H.[23] Thus, we have $\tilde{\phi}_R^0$=0.4386, $\tilde{\phi}_H$=5.3123, $\tilde{\phi}_S$=2.1439, and $\tilde{\phi}_P^H$=0.2382.

Four preference parameters remain (ε, θ, $\underline{\beta}$, and $\bar{\beta}$). The concavity in the utility function for the number of children ε is set at 0.1, which is roughly equal to the value used in Gobbi (2013). For the altruistic factor θ, the Leontief preference implies that $c = \frac{(1-\theta)nb}{\theta}$, and hence, $\frac{c}{nb} = \frac{1-\theta}{\theta}$. Therefore, $1-\theta$ represents the fraction of income being spent on own consumption, and θ is the fraction of income for bequest. Using the information on rural per capita income and consumption reported in *China Statistical Yearbook*, we compute a time series of θ, and the average of θ becomes 0.2124 during 1980-2007. We thus set θ at 0.21. Preference parameters are not observable, so we simply set $\underline{\beta}$ at 0.1. The last preference parameter, $\bar{\beta}$, is jointly calibrated with the prior rural childrearing cost and the three extra fines ($\tilde{\phi}_R^0$, $\tilde{\phi}_H$, $\tilde{\phi}_S$, and $\tilde{\phi}_P^H$). In addition to the four data moments for fertility described above, we also match n_P^L in the joint calibration. The calibrated $\bar{\beta}$ is 0.5346.

We are left with the calibration for the urban embarked tax τ, urban benefits B, the government's technology scaling factor in the provision of urban benefits B_0, and the migration costs of rural type-$\{H, \bar{\beta}\}$ and type-$\{L, \underline{\beta}\}$ agents, ψ^H and ψ^L. No perfect measure exists to proxy urban benefits. As urban workers usually enjoy government-provided pension benefits, we thus use urban pension benefits to proxy B. To compute B, we use the pension replacement rate in China. It is then multiplied by the urban average income during the examined period and adjusted by both the model period and the average number of years that an urban worker enjoys after retirement. Thus, B is 0.9803. Following Song, Storesletten, Wang and Zilibotti (2015), we set the urban social security tax τ at 20 percent of an urban agent's income. The total social security taxes are collected to finance urban benefits. With the technology that the government provides urban benefits, we can determine B_0, which is 2.4863. Finally, ψ^H and ψ^L are calibrated using the indifference equations for migration, (21) and (22), and are equal to 0.4230 and 0.1535, respectively. This result indicates that, compared to low-skilled migrants, rural high-skilled workers who get more enjoyment from having children suffer more from migrating to cities. The calibration result is summarized in Table 2.[24]

[Insert Table 2 about here]

[22] See Appendix IC for further discussion of the literature on urban childrearing cost.

[23] The method we use to impute the fertility targets and the data sources are relegated to Appendix IC. In the calibration, for example, we compute a weighted fertility of P^H and P_F^H to match n_P^H. Fertility in other sectors is computed accordingly. Besides, as our model is unisex, we match half of the imputed fertility in the calibration.

[24] In Appendix ID, two sets of robustness tests are performed. First, we consider different but reasonable assumptions regarding the fraction of floating populations in different sectors. By recalibrating the model and re-conducting all the policy experiments, which will be elaborated on below, we find that the results are robust for all policy experiments on migration decisions, fertility responses, skill premium, and urban premium. The second set provides sensitivity tests by varying α, σ, τ, and private-sector premium w_P^H/w_S. In the tests, we re-calibrate the model by changing the parameters or the target one by one. Again, we find that the results are robust.

In Table 3, we further report important ratios implied by the calibration results. Our results indicate that the private sector has much better technology than the SOE and rural agricultural sectors. Among high-skilled rural workers, more than 15 percent decide to migrate to cities, while about 95 percent of rural low-skilled workers choose to stay in the countryside. Compared to rural workers, SOE-sector workers and private high-skilled workers face higher childrearing costs. However, private low-skilled parents bear slightly lower childrearing costs than rural workers. This difference could be due to the higher implicit cost in rural areas, (e.g., living far away from schools) or to the relatively higher childrearing cost for rural high-skilled parents (an argument used by Cheung (2022) for examining demographic transition and structural transformation in the U.S. at the turn of the twentieth century). Regarding the above-quota penalty, workers in the SOE sector were penalized the most heavily, followed by rural high-skilled workers, while rural low-skilled workers were punished the least. The result is consistent with reality. In fact, rural high-skilled workers are more likely to be leaders in the village and tend to hold positions designated by the government. The result indicates that, no matter the working location, workers in the SOE sector are being seriously monitored by the government. Based on the calibrated parameters, we now proceed to study the interaction of migration and fertility decisions.

[Insert Table 3 about here]

5.3 Counterfactual Analysis

Rural-urban migration and fertility decisions interplay and interact with each other in the process of economic development. The tighter population control and higher childrearing costs in urban China deter rural workers from migrating to cities; cause migrant workers to have fewer children; and affect the urban labor composition, skill premium, and urban premium - which subsequently influence prospective migrants' migration and fertility decisions. To quantify the importance of these fertility-migrating on interactions, we conduct two counterfactual experiments by shutting down the migration decisions (i.e., shutting down the mixed migration equilibrium) or removing the fertility differentials from urban to rural areas. We are particularly interested in learning the migration patterns, the fertility behavior, and other important macroeconomic variables and ratios under each of the scenarios. To proceed, we first compute the equilibrium of the benchmark model based on the calibrated parameters. Next, we shut down the channels under examination to compute the new equilibrium and then calculate the relative changes to the benchmark model. The results related to migration decisions and population ratios are summarized in Table 4.[25] Table 5 reports the results of the total fertility rates, the income levels, and the relative income ratios.[26]

[25] In Tables 4, 6, 9, and 11, the fraction of H (L) movers refers to the fraction of the high-skilled (low-skilled) migrants among the total high-skilled (low-skilled) rural workers, and the fraction of movers refers to all migrants among total rural workers for each experiment.

[26] Notably, in the calibration, we match half of the total fertility rates. However, to be comparable with data, the fertility rates reported in Tables 5, 7, 10, and 12 have been multiplied by two.

[Insert Table 4 about here]
[Insert Table 5 about here]

5.3.1 Shutting down the migration decision

The first counterfactual experiment is to shut down mixed migration equilibrium by setting $\Gamma_H = \Gamma_L = 0$, under which locational choice is purely a matter of exogenously drawing types with no endogenous migration decisions: High-skilled workers with low fertility preference are in urban areas, whereas all other types are in rural areas. Compared with the mixed migration equilibrium benchmark, the counterfactual equilibrium features higher quantity but lower quality for fertility choice. High-skilled workers with high fertility preference (type-$\{H, \bar{\beta}\}$) and low-skilled workers with low fertility preference (type-$\{L, \underline{\beta}\}$) stay in rural areas, so the fertility rates of rural high-skilled workers and urban private-sector low-skilled workers increase by 5.39 percent and 0.28 percent, respectively. Given that the childrearing cost is proportional to income, when more rural stayers are splitting rural land and thereby dragging down rural per capita output and rural childrearing cost, the rural fertility rate increases. Conversely, because only high-skilled workers with low fertility preference migrate to cities, skill composition in urban areas largely improves by 47.16 percent, leading to a lower skill premium but a higher urban per capita output. The urban-rural income gap would be 5.8 percent higher than the benchmark if no endogenous migration is allowed.

5.3.2 Removing fertility differentials

The second counterfactual experiment considers shutting down all differential fertility channels. To conduct this experiment, we set the urban fertility constraint to be as loose as that for rural ($\bar{n}_U = \bar{n}_R$), the urban above-quota fine to equal that of rural ($\tilde{\phi}_U = \tilde{\phi}_R$), the extra above-quota fines for urban SOE and private high-skilled workers to equal the extra fine for rural high-skilled workers ($\tilde{\phi}_S = \tilde{\phi}_P^H = \tilde{\phi}_H$), and the urban and rural childrearing costs to be the same ($\tilde{\phi}_U^0 = \tilde{\phi}_R^0$). All fines and childrearing costs are proportional to workers' income and thus adjusted by the corresponding wage ratios ($\frac{w_S}{w_R}$, $\frac{w_P^H}{w_R}$, and $\frac{w_P^L}{w_R}$) in the benchmark economy. Such adjustments ensure that rural and urban workers face the same values for above-quota fines and childrearing costs.

As suggested by Table 4, compared with the mixed migration equilibrium benchmark with differential fertility, this counterfactual equilibrium results in more rural-to-urban migration. More specifically, removing the fertility differentials greatly increases the migration incentives and boosts the high-skilled migration by 62.5 percent, thereby lowering the skill premium in the private sector and slightly increasing urban per capita output. With fewer stayers, each rural worker is allocated with more land. Rural per capita output slightly increases, and the increase outweighs the increase in urban per capita output. The urban-rural income gap consequently shrinks by 0.78 percent. In addition, in response to a higher rural per capita output, rural fertility rates drop as above-quota penalties and childrearing costs are proportional to income. Further, as above-quota fines and childrearing cost in cities become lower and wages become higher, the SOE workers and private low-skilled workers can afford having more children and thereby increase fertility. In contrast, private high-skilled workers choose to have fewer children in response to the decrease in wage income.

To summarize, the results of our counterfactual analysis justify the importance of considering the joint decisions between rural-urban migration and fertility. Overlooking either one when evaluating policies will result in biases due to changes in migration incentives and the exclusion of their interactive effect. In what follows, we thereby proceed with experiments on population- and migration-related policies based on the benchmark economy.

5.4 Policy Experiments

We are now ready to conduct policy experiments. More specifically, we want to investigate the macroeconomic consequences of two public policies of particular interest: (i) population control policies and, (ii) migration policies, via *hukou* regulations and migration-related land policies (e.g., variations in land reallocation and land supply), which both affect migration incentives.

Aligning with the discussion in Section 3.1, we restrict our attention to a benchmark economy with the most plausible migration pattern: Type-$\{H, \underline{\beta}\}$ workers always migrate to cities, type-$\{L, \bar{\beta}\}$ workers always stay in the rural, and type-$\{H, \bar{\beta}\}$ and type-$\{L, \underline{\beta}\}$ agents are indifferent between migrating and staying in rural areas. As some experiments involve large changes in the institutions workers face, we may end up with corner solutions ($\Gamma_H = \Gamma_L = 0$), SMEs with pure skill-based or skill-based migration ($\Gamma_H = 1, \Gamma_L \in [0, 1)$), or pure fertility-based or fertility-based migration ($\Gamma_H = 0, \Gamma_L \in (0, 1]$). The fertility responses, the population ratios, and the relative income levels are thus led by the asymmetric structure that only high-skilled workers with high fertility preference (type-$\{H, \bar{\beta}\}$) and low-skilled workers with low fertility preference (type-$\{L, \underline{\beta}\}$) respond to the experiments.

What follows is a detailed discussion of the policy experiments. Similar to Section 5.3, with the benchmark model we change the value of the relevant parameter of the concerned experiment to compute the new equilibrium and then calculate the relative changes to the benchmark model. Table 6 reports the results on migration decisions and population ratios, Table 7 on total fertility rates, and Table 8 on income levels and relative income ratios.

[Insert Table 6 about here]
[Insert Table 7 about here]
[Insert Table 8 about here]

5.4.1 Experiments on population policy

Four experiments are conducted here, two considering scenarios with a relaxation of the one-child policy and two studying the realistic reforms introduced in recent years. In the first experiment, we consider a uniform fertility constraint in both urban and rural areas, with the constraint set at 0.8 ($\bar{n}_U = \bar{n}_R = 0.8$) – a value prevalent in rural China in the benchmark model in a unisex setting.[27] Urban fertility control is thus more relaxed under this experiment. In this scenario, we

[27] Again, the observed fertility rates are our normalized figures multiplied by two.

find that rural workers are more willing to migrate to cities: Total rural migrants as a percentage of rural residents increases by 12.59 percent. In addition, the high-skilled workers with high fertility preferences respond much more to such relaxation, therefore leading to a lower H^{+}/L^{+} ratio in rural areas. As there are more high- than low-skilled workers migrating to cities, P^{H+}/P^{L+} increases, dragging down wages for the high-skilled workers by 0.47 percent and slightly increasing wages for the low-skilled workers by 0.44 percent. The skill premium thus decreases by 0.92 percent. Interestingly, both urban and rural per capita income increase. The former is due to a larger stock of high-skilled workers in urban areas, while the latter is because of land reallocation owing to fewer farmers in rural areas. As the childrearing costs increase with income levels, rural stayers and private-sector low-skilled workers have fewer children. For private high-skilled workers in urban areas, even though the urban fertility control is relaxed, those workers who get more enjoyment from having children do not respond much to the policy by increasing fertility. This lack of increase occurs due to the higher childrearing costs in urban areas. As shown in Table 7, the magnitudes of the increases in the total fertility rates for SOE and private high-skilled workers are much less than that of the decreases in the total fertility rates for rural workers.

In the second and third experiments, we study two realistic population policy reforms recently implemented in China. The second experiment considers the two-child policy in both rural and urban areas, with the constraint being set at 1 ($\bar{n}_U = \bar{n}_R = 1$). The third experiment examines the three-child policy with $\bar{n}_U = \bar{n}_R = 1.5$. We find that the effects of these two realistic policy reforms mirror those in the first experiment, but the magnitudes are larger. In other words, the relaxation of fertility control encourages more high-skilled migration, increases overall per capita output, and lowers the urban-rural income gap.

In the fourth experiment, we consider a scenario where population control is completely lifted. Agents can have children up to the biological limit, which is set at 5.5 based on the total fertility rates in the pre-population-control period in China.[28] Total fertility rates increase in both rural and urban areas, rising from the range of 1.18 and 2.21 in the benchmark model to 1.96 to 2.77. The increase in fertility is higher for rural high-skilled than rural low-skilled workers because the staying type-$\{H, \bar{\beta}\}$ workers respond more to the policy than type-$\{L, \underline{\beta}\}$ workers. Along the line, rural-urban migration decreases by 2.74 percent, with the reduction in migration coming more from the high-skilled workers – indeed, all high-skilled workers with high fertility preference (type-$\{H, \bar{\beta}\}$) would stay in rural areas and a fertility-based SME would arises. As a result, fewer high-skilled workers are in urban areas, leading to a lower urban per capita output, a lower rural per capita output due to more workers staying in rural areas, a slightly higher skill premium, and a higher urban-rural income gap ($\frac{y_U}{y_R}$).

5.4.2 Experiments on migration policy

We now turn to migration policies, inclusive of *hukou* regulations, land reallocation, and land supply policies that directly alter migration incentives.

1. Experiments on hukou regulation

[28] The biological fertility limit is set at 5.5 per family. Thus, in our experiment, the limit is 5.5/2=2.75 per worker.

We consider three scenarios for experiments on *hukou*-related migration policy: (i) migrants can never obtain urban residency ($\rho = 0$), (ii) migrant workers are ineligible for urban benefits even if urban residency has been obtained ($\mu = 0$ and therefore no tax payments), and (iii) migrant workers who obtain urban residency enjoy full urban benefits as urban natives ($\mu = 1$ and pay full urban earmarked tax τ).

In the first and the second experiments, we encounter a corner solution – both type-$\{H, \bar{\beta}\}$ and type-$\{L, \underline{\beta}\}$ workers decide to stay in the rural areas. With zero low-skilled migration, rural-urban migration simply comprises type-$\{H, \underline{\beta}\}$ migration. The skill premium in urban areas and the H^+/L^+ ratio in rural areas both decrease as a result. Rural workers are now allocated with less arable land due to more stayers in rural areas, leading to a decrease of 4.07 percent in rural per capita income. Further, the relatively lower childrearing costs in rural areas cause an increase in the fertility rates of rural high- and low-skilled workers, but the rural high-skilled fertility rate increases more. These effects occur because, among high-skilled workers, only the type-$\{H, \bar{\beta}\}$ stay in rural areas, while the low-skilled stayers comprise both preference types. The main differences between the two counterfactual experiments only emerge with private high-skilled fertility and urban benefits. When migrants can never obtain urban residency in the first experiment, the private high-skilled fertility rate is purely attributed by urban natives. However, in the second experiment where migrants are ineligible for urban benefits but can obtain urban residency, the private high-skilled fertility rate is a weighted average of urban natives and new comers of purely type-$\{H, \underline{\beta}\}$ migrants. Thus, unsurprisingly, the private high-skilled workers have lower fertility in the second experiment where migrants enjoy no urban benefits.

The third experiment conducts the scenario where migrant workers who obtain urban residency enjoy full urban benefits as urban natives. We have an unusual corner solution that $\Gamma_H = 1$ and $\Gamma_L = 1$. Due to the strong incentive from being granted with full urban benefits, all rural high-skilled and type-$\{L, \underline{\beta}\}$ workers flood into cities. Therefore, P^{H+}/P^{L+} increases, skill premium declines by 1 percent, and urban per capita output drops. In contrast, now only type-$\{L, \bar{\beta}\}$ agents live in rural areas. The population outflow from the rural areas largely causes the rural per capita income to increase by 17.6 percent, reducing the urban-rural income gap by more than 15 percent. Moreover, rural fertility rates and the fertility of private low-skilled workers all decline as their income levels increase.[29] Our quantitative results on fertility contrast those of international migration, e.g., Azarnert (2019).

2. Experiments on land entitlement system

We next turn our attention to land reallocation policy that may change migration incentives. The land left idle by migrant workers appears to be a waste of productive resources in rural areas. To prevent the waste of productivity, people may propose to redistribute the idle land to rural stayers as soon as possible. However, land reallocation implies that rural-urban migrants would lose their land use rights. This implicit cost discourages rural-urban migration. We therefore perform two

[29] The corner solution implies that no high-skilled workers stay in rural areas. In Table 7, the total fertility rate for rural high-skilled workers refers to the fertility of those who migrate to cities but eventually fail to obtain urban residency.

counterfactual experiments to investigate the macroeconomic effects of land reallocation policies: (i) an immediate land reallocation policy featuring zero land tenure security ($\delta = 0$); and (ii) a more secured land tenure policy that guarantees 15 years of land tenure ($\delta = \frac{15}{36} = 0.4167$) – in other words, compared with the benchmark economy, the duration of land reallocation is doubled.

As shown in Table 7, our results show that the effects of land reallocation policies on total fertility rates are relatively minor compared to the experiments on population policies and *hukou*-related migration policies. However, in Table 6, we find that immediate loss of the entitlement of rural land deters rural-urban migration: Rural high-skilled workers who get more enjoyment from having children are no longer moving to cities. The fraction of rural low-skilled movers also drops by 13.94 percent. Thus, similar to the experiment involving no population control, a fertility-based SME arises. Because of the relatively large decrease in the low-skilled migration, the skill premium decreases by 0.56 percent, but urban per capita output still increases by 0.17 percent as fewer workers with lower incomes are present in urban areas. Despite more stayers in rural areas, due to more efficient land use, rural per capita income increases by 0.47 percent, and the urban-rural income gap is narrowed by 0.3 percent.

Notably, immediate land reallocation brings two opposing effects on the overall income level (y) in the economy: the direct effect to increase overall output per capita when land is being utilized more efficiently, and the migration discouragement effect to depress overall output per capita when there are more rural stayers. As rural per capita output is lower than urban per capita output, keeping urban and rural per capita output constant, the average overall per capita output will be lower if more workers are in rural areas. The sum of the two opposite effects of immediate land reallocation on overall output per capita is therefore ambiguous. Our quantitative results show that the migration discouragement effect outweighs the direct effect and that the overall output per capita slightly decreases by 0.1 percent.[30]

In contrast to the immediate land reallocation policy, a guaranteed land tenure of 15 years promotes more migration for both the high- and low-skilled workers, slightly increases urban high-skilled fertility, and discourages urban low-skilled fertility as more high-skilled parents who get more enjoyment from having children and low-skilled workers who get less enjoyment from having children move to cities. Due to longer idle duration of land, the rural per capita output is slightly depressed. The urban per capita output is lowered, however, due to more low-skilled workers in cities dragging down the average income levels. These dynamics then result in a higher skill premium. Despite the slightly lower urban and rural per capita output, the overall per capita output in the economy is still higher by 0.74 percent due to the migration encouragement effect.

Our results on land entitlement policy echo that of Ngai, Pissarides and Wang (2019): A less secure land tenure system is an obstacle to urbanization and industrialization in China. Our paper also corroborates with the findings of the occupational misallocation model by Gottlieb and Grobovšek (2019), which is calibrated to Ethiopia: Relaxing a "use it or lose it" land policy mod-

[30] The overall output per capita is $y = \frac{y_U U^+ + y_R R^+}{U^+ + R^+} = \frac{y_U \left\{ \frac{U}{R} + \frac{H}{R}[\zeta + (1-\zeta)\Gamma_H] + \frac{L}{R}\zeta\Gamma_L \right\} + y_R \left\{ \frac{H}{R}(1-\zeta)(1-\Gamma_H) + \frac{L}{R}[\zeta(1-\Gamma_L) + 1 - \zeta] \right\}}{\frac{U}{R} + 1}$. Keeping y_U and y_R constant and $y_U > y_R$, when Γ_H and Γ_L are lower, $\frac{H}{R}(1-\zeta)\Gamma_H(y_U - y_R)$ and $\frac{L}{R}\zeta\Gamma_L(y_U - y_R)$ are both lower, leading to lower y.

erately raises overall per capita income. But different from the prior researchers' sorting approach, our land reallocation policy effect on the urban-rural productivity gap is reversed owing to two novel channels that work through the change in arable land supply in rural areas and the skill composition (hence the wage premium) in urban production. Regarding how the land reallocation policy affects fertility rates, our results align with those of Almond, Li, and Zhang (2019): The effects of land reallocation policies on overall fertility rates are relatively small compared to the effects of population policies on fertility rates.

3. Experiments on land supply system

To solve the problem of low work incentives in rural areas, China changed its farming institution from a collective system to a household-responsibility system (HRS) in 1979. By the end of 1983, most production teams had adopted HRS. The implementation of HRS assigned collectively owned land to each household with the guaranteed use right for several years.[31] We thus observe a sharp increase in land per person in 1983, as shown in Figure IC.1.

The HRS reform enables us to conduct three experiments on changes in land supply: (i) full HRS from 1980 to 2007 ($q = 2.4876$, the average of 1983-2007), (ii) stable land supply since 1983 ($q = 2.2564$, which is equal to the average of 1980, 1981, 1982, and thereafter the average of 1983-2007), and (iii) no HRS throughout ($q = 0.23$, the value of 1980). In the first experiment, imposing full HRS throughout 1980-2007 increases land per person from the benchmark value 2.2564 to 2.4876, thus increasing the incentives to stay in rural areas. As Table 6 shows, we encounter a corner solution where only type-$\{H, \underline{\beta}\}$ workers migrate to cities. Thus, the effects on the economy are like those within the experiments focused on no urban residency and no urban benefits. The main difference from the two *hukou*-related migration policies is that rural per capita output increases by 5.8 percent in the scenario of full HRS due to more land per person. Therefore, this reform shrinks the urban-rural income gap by 4 percent and increases overall per capita output by 1.1 percent. Moreover, the higher rural per capita income leads to higher childrearing costs, thus lowering rural total fertility rates.

The second experiment considers a stable land supply after 1983. As the value of q is very close to the value in the benchmark economy, the result of the second experiment resembles what occurs in the benchmark economy. Then, the third experiment studies what would happen if China had no HRS reform. Without HRS reform, land per person sharply drops. The strong incentive to migrate to cities results in an unusual corner solution: Only the type-$\{L, \bar{\beta}\}$ workers stay in rural areas, and all rural high-skilled workers move to cities. The result resembles those in the scenario with full urban benefits. However, in the land supply experiment, because land per person drops, rural per capita income decreases substantially by around 88 percent. The cheaper childrearing cost results in a significant rebound of rural low-skilled workers' fertility (from 2.2 in the benchmark to 4.8). In addition, the urban-rural income gap is enlarged, and overall per capita output declines by 26.5 percent.

4. Summary

The above experiments all highlight the presence of a locational quantity-quality trade-off of

[31] See Lin (1992) for the more in-depth details.

children: More migration is accompanied by a lower fertility but higher overall per capita output. No population control leads to a higher fertility for both the high- and the low-skilled workers, accompanied by lower migration – especially a larger reduction in the high-skilled migration and thereby a lower overall output per capita. Relaxing urban population control to rural areas is better than entirely removing population control as the former induces more high-skilled migration, leading to higher urban output and overall output per capita. From a nationwide perspective, stricter population control policies in urban areas may not be ideal in lowering fertility rates because such a tightened fertility policy in urban areas deters high-skilled workers with high fertility preferences from migrating to cities. This effect is undesirable because, had high-skilled workers migrated to urban areas and faced a higher childrearing cost, the workers would have naturally adjusted their fertility down. An urban sprawl control policy (e.g., eliminating urban benefits to all rural migrants) will raise urban benefits enjoyed by workers with urban residency as planned but would not grant any incentive for rural-urban migration, which is undesirable for advancing the economy. An urban promotion policy (e.g., providing full benefits to all rural migrants) would induce more migration overall. With more generous rural land entitlement or a lower supply of rural land, migration incentives are strengthened. Rural per capita output is lower due to more land being left idle by migrants while urban per capita output decreases due to the influx of more low-skilled workers. Nonetheless, the overall per capita output is higher due to the urban wage premium – a result that is possible because of the endogenous population weights by virtue of rural-urban migration.

In sum, when formulating policies for developing countries, one shall not neglect policy impacts via migration and fertility responses. Overlooking the locational quantity-quality trade-off of children may lead to nonnegligible biases in assessing the consequences and effectiveness of government policies.

5.5 Experiments for the Generalized Model

We are now left with a quantitative assessment of the roles played by reverse migration, left-behind children, upward skill mobility, and directed urban benefits in the magnitude and the macroeconomic consequences of equilibrium migration flows from rural to urban areas. To understand the impacts, we conduct experiments by adding one at a time to the baseline model and then compare the new equilibrium to the benchmark economy. Tables 9 and 10 summarize the results of reverse migration, left-behind children, and upward skill mobility; Tables 11 and 12 report the results of directed urban benefits. Data descriptions and calculations are relegated to Appendix IC.

[Insert Table 9 about here]

[Insert Table 10 about here]

5.5.1 Experiment on reverse migration

As discussed in Section 4.1, the major departure of the generalization for reverse migration from the baseline model is the introduction of the urban residency shock (Λ) or the fraction of reverse migration to urban population. Setting Λ at zero returns to the baseline model. Our task here is thus to pin down Λ, compute the new equilibrium, and calculate the relative changes to the benchmark model.

By normalizing total population to one and using the data on rural-urban population ratio in 1980-2007, the ratio of rural-urban migration to rural population in our benchmark economy, and the data on rural-urban and urban-rural migration as a percentage of total migration reported by Scharping (1997), we can compute the ratio of urban-rural migration – the reverse migration – to urban population and obtain $\Lambda = 1.1516\%$.

Tables 9 and 10 suggest that reverse migration disproportionally discourages rural high- and low-skilled workers from migrating to cities by hurting the high-skilled migration more. Urbanization rate drops by nearly one percent. Despite only minor impacts on the ratio of urban private high- to low-skilled workers, the skill premium, and the fertility rates in urban and rural areas, the worse labor composition in urban areas and more stayers in rural areas lower overall per capita output by about 0.26%.

5.5.2 Experiment on left-behind children

To calibrate the childrearing cost markup of left-behind children (Ξ), we utilize remittance data with appropriate adjustments. We consider both aggregate and household-based measures. From the aggregate perspective, the remittance income for left-behind children is about 4.37 percent of total rural household income. As the childrearing cost in our benchmark economy is 43.86 percent of rural household income, the aggregate measure for the markup is $(0.0437+0.4386)/0.4386 = 1.0996$. Alternatively, the household-based imputed remittance for left-behind children is about 3.37 percent of rural household income. This gives a household-based markup of 1.0768. These two measures are comparable, and we simply average them to obtain $\Xi = 1.0882$. Finally, to be consistent with reality, the childrearing cost of left-behind children is adjusted by the wage ratio of rural to private low-skilled workers in the benchmark economy.

By considering left-behind children, we find that the fraction of movers reduces sharply by more than a quarter due largely to the shrinkage of low-skilled migration. In other words, the fraction of low-skilled movers declines by 36.3 percent, while that of high-skilled movers declines by less than 10 percent. With an improved skill composition of urban workers, both urban per capita output and the urban-rural income gap increase, but the skill premium in the urban private sector falls. Fewer migrants result in lower rural per capita income and a cheaper rural childrearing cost. Consequently, the average fertility rates of rural high- and low-skilled workers rise by about 1.5-2 percent. Overall per capita output falls by about 1.5 percent due to a lower urbanization rate and thereby shrinks urban activities.

5.5.3 Experiment on upward skill mobility

We now turn to the experiment on upward skill mobility. In Liao, Wang, Wang and Yip (2022), the upward skill mobility is computed at 0.04 percent and 1.17 percent, respectively, for the periods of 1980-1994 and 1995-2007, and the average upward mobility rate is 0.605 percent. We thus set $\vartheta = 0.605\%$ for this experiment. Others remain unchanged. While such mobility is quantitatively trivial, its implication remains intellectually intriguing.

We find that upward skill mobility has a direct skill enhancement effect in urban areas. However, such mobility discourages rural high-skilled workers but encourages rural low-skilled workers to migrate to cities. As shown in Table 9, compared with the benchmark economy, the fraction of high-skilled movers decreases by 2.7 percent, while that of low-skilled movers increases by 1.6 percent. This endogenous migration effect worsens the skill composition of the urban workforce. Therefore, with the direct and various general-equilibrium effects, the $P^{USM,H+}/P^{USM,L+}$ ratio slightly increases by 0.14 percent. The resulting effects on fertility, overall per capita output, wage ratios and income gaps are found to be limited.

5.5.4 Experiment on directed urban benefits

In the generalized framework in Section 4.4, we allow urban benefits to be directed to (i) human capital at work (B_W), (ii) child education (B_E), or (iii) child care (B_C), beyond pure amenities (B_A) considered as B in the baseline model. We also follow the general principle that all directed benefits come with urban residency status. However, not every urban resident in China is eligible for the four types of urban benefits. While all urban residents with formal *hukou* enjoy pure amenities B_A and better education for children B_E, only SOE workers enjoy childcare subsidies B_C, and only SOE and urban high-skilled workers are eligible for B_W. Thus, we must tailor the framework for directed urban benefits in Section 4.4 to the institutions in China. Specifically, we denote $\bar{G}_m$ as the government expenditure per person for urban residents receiving full urban benefit B_m, $m = A, W, E, C$. While the total urban tax collected by the government (T) remains the same as in (38), the government's balanced budget rule is modified as:

$$\begin{aligned} T &= \bar{G}_A \left[U + S_F + \mu \left(P_F^H + P_F^L\right)\right] + \bar{G}_W \left(S + S_F + P^H + \mu P_F^H\right) \\ &\quad + \bar{G}_E \left[U + S_F + \mu \left(P_F^H + P_F^L\right)\right] + \bar{G}_C \left(S + S_F\right). \end{aligned} \tag{46}$$

In addition, the government technologies are adjusted accordingly and become: $B_m = B_{0,m}\bar{G}_m$, with the benefit type $m = A, W, E, C$.

We first re-calibrate the modified model to data from China. Specifically, we calibrate the four directed urban benfits (B_A, B_W, B_E, and B_C) and $\bar{G}_W, \bar{G}_E$, and $\bar{G}_C$ using various data. Then, given $\tau = 20\%$ as in the benchmark economy and T computed therein, $\bar{G}_A$ is solved using the government's balanced budget rule (46). With all B_m and $\bar{G}_m$ being determined, we can calibrate the four types of $B_{0,m}$ using the government technologies. The detailed calibration strategy, data sources, and calibrated results are relegated to Appendix IC.

Based on the calibrated result, we then proceed to conduct separate experiments for each type of the directed urban benefits and study the pure effect of each benefit on migration and fertility decisions. To ensure comparability, we consider an exogenous one-percentage increment per urban resident tax revenue to be allocated to $\bar{G}_m$, thereby increasing the corresponding directed urban benefit. The results are summarized in Tables 11 and 12.

[Insert Table 11 about here]
[Insert Table 12 about here]

1. Directed urban benefits to human capital at work

As B_W is only granted to SOE and private high-skilled workers, the increase in $\bar{G}_W$ strengthens the migration incentives for rural high-skilled workers. The fraction of high-skilled movers increases by about 10 percent, while that of low-skilled movers only increases by 0.5 percent. This change improves the skill composition in cities, lowers the skill premium, and increases the urban per capita output. Fewer rural stayers result in slightly higher rural per capita output, thereby shrinking the urban-rural income gap. In addition, as B_W enhances urban high-skilled workers' income and more type-$\{H, \bar{\beta}\}$ workers migrate to cities, the fertility of the SOE and the urban private high-skilled workers increases, while the fertility of rural high-skilled workers declines.

2. Directed urban benefits to child education

All urban residents with formal *hukou* enjoy the directed urban benefits to child education (B_E). Therefore, an increase in $\bar{G}_E$ promotes rural-urban migration, and the effects of an increase in $\bar{G}_E$ on population ratios, output per capita, and income ratios resemble that of an increase in $\bar{G}_W$, but with a smaller magnitude. However, different from the experiment on $\bar{G}_W$, an increase in $\bar{G}_E$ shifts parents' choice from child quantity to child quality. We therefore observe that the total fertility rates of SOE and private high-skilled workers decline.

3. Directed urban benefits to child care

Only SOE workers are eligible to receive the directed urban benefits for child care (B_C). As only a small proportion of high-skilled migrants can be employed in the SOE sector, the strengthening effect of an increase in $\bar{G}_C$ on migration incentives is minor - the fraction of high-skilled movers increases by only 0.27 percent. Besides, as the benefits decrease the childrearing costs for SOE workers, their fertility slightly increases as a result.

4. Summary

The above three directed urban benefits all strengthen migration incentives, especially for rural high-skilled workers. Because the beneficiaries are not the same, the magnitudes of the respective impacts on migration decisions, output levels, income ratios, and population ratios differ, with the effects of an increase in directed urban benefits to human capital at work outweighing those of the other two benefits. The effects on fertility vary across types of directed benefits. While the benefit to child education shifts parents from child quantity to child quality, both the benefits directed to human capital at work and to child care lead to a fertility rebound. Finally, comparing with the above three benefits, the last two rows in Tables 11 and 12 suggest that an increase in $\bar{G}_A$ (therefore, an increase in pure amenities B_A) brings the largest effect on migration, fertility and the macroeconomic variables. This finding supports the setup of our baseline model where only pure amenities are considered.

5.5.5 Taking Stock

We once again verify the presence of a locational quantity-quality trade-off of children in all cases of generalizations of the baseline model. Most interestingly, reverse migration and left-behind children both weaken migration incentives and reduce rural-urban migration. Staying in rural areas results in a higher fertility rate for both rural high- and low-skilled workers. In contrast, allowing upward skill mobility encourages low-skilled migration but at the cost of a lower fertility rate. As for urban directed benefits, all benefits considered strengthen rural workers' migration incentives, especially for high-skilled workers, thereby enhancing the skill composition and lowering the skill premium in cities. Among all directed urban benefits, pure amenities still play the dominant role in affecting migration decisions and macroeconomic variables. However, the relative changes to the benchmark under all dimensions of generalizations, inclusive of reverse migration, left-behind children, upward skill mobility, and directed urban benefits, are quite limited, thereby justifying our baseline model.

To conclude, any policy that leads to a locational quantity-quality trade-off of children would be worthy because it improves children's quality and skill composition in cities, encourages rural-urban migration, and thereby promotes overall per capita output and reduces the urban-rural income gap.

6 Conclusions

We have constructed a dynamic competitive migration equilibrium framework with heterogeneous agents in skills and fertility preferences to explore the macroeconomic consequences of the dynamic interplay of migration and fertility decisions. We have established a locational quantity-quality trade-off of children and found conditions to support a mixed migration equilibrium. By calibrating the model to fit the data from China, we have identified interesting interactions between fertility choices and migration decisions in various counterfactual experiments with regard to changes in population control and migration-related policies. We have shown that all such changes induce a locational quantity-quality trade-off of children in which more rural-urban migration is accompanied by lower fertility but higher overall per capita output.

To achieve the main purposes mentioned above, we have imposed various simplifying assumptions. Before wrapping up this paper, it is thereby useful to acknowledge the limitation of our study, which may be viewed as potentially interesting lines of future research as well. First, one may quantify dynamic impulse response to time varying TFP advancement, particularly the continual widening of the urban-rural TFP gap captured by $A_P(1-\alpha)^{1/\sigma}/z$. Accordingly, one could examine whether a migration or fertility policy may better advance the economy as it progresses. Second, one may consider an exogenous labor composition change due to *xiagang* that lowers the private-SOE sector TFP gap, A_P/A_S. One may thus examine the potential short-run productivity loss and long-run productivity gain of such a privatization policy. Third, one may also allow for another dimension of heterogeneity via skill draws where high-skilled agents draw η following a Pareto distribution with the minimum skill $\eta_t^{\min}=1$ (i.e., par with the low-skilled workers). Under the thick tail distribution, the expected value of the skill premium computed using the relative productivity measure is higher than the simple average of the skill premium. Therefore, this is expected to amplify migration incentives of the high-skilled workers and strengthen locational quantity-quality trade-off. Fourth, one may further consider urban networks that can reduce migration cost ψ, increase the chances of obtaining SOE jobs π or urban residency ρ, or enhance urban learning per

Lucas (2004), as captured by continually rising $\eta_t^{\min}$. Either is expected to encourage migration and lead to a locational quantity-quality trade-off in favor of children's quality. Finally, our analysis is exclusively positive and hence potentially rewarding via being generalized to allow for normative analyses, such as to examine optimal taxation or subsidies related to urban labor markets and migration, with endogenous provision of urban amenities and city government financing in a Tiebout equilibrium context, as in Conley and Konishi (2002). Of course, while potentially of interest, these extensions are beyond the scope of the current paper.

To this end, we would like to highlight three policy prescriptions that are general for many developing countries. The first lesson is: Do not abuse the power of restricting migration, within a country or internationally, as the most likely consequences are detrimental. The second is: Population control, if required at the early development stage when facing poverty without concern regarding the ageing issue, should be non-discriminatory and be applied to all people at all locations. The last is: Any incentives toward enhancing the quality of children would be rewarding, including mandatory basic education, means-tested college tuition subsidies or low-interest student loans, and the ban of child labor, to name but a few.

References

[1] Aghion P. and P. Bolton, 1997, "A theory of trickle-down growth and development," *Review of Economic Studies*, 64(2), pp. 151-172.

[2] Almond, D., H. Li and S. Zhang, 2019, "Land reform and sex selection in China," *Journal of Political Economy*, 127(2), pp. 560-585.

[3] Andreoni, J., 1989, "Giving with impure altruism: Applications to charity and Ricardian equivalence," *The Journal of Political Economy*, 97(6), pp. 1447-1458.

[4] Azarnert, L., 2019, "Migration, congestion, and growth," *Macroeconomic Dynamics*, 23(8), pp. 3035-3064.

[5] Banister, J., 1997, "China: Internal and regional migration trends," Cologne China Studies Online - Working papers on Chinese Politics, Economy and Society, No. 14.

[6] Bernard, A., M. Bell and J. Cooper, 2018, "Internal migration and education: A cross-national comparison," Background Report to the 2019 Global Education Monitoring Report. Paris, France:UNESCO.

[7] Cheung, T, 2022, "Schooling, skill demand and differential fertility in the process of structural transformation," *American Economic Journal: Macroeconomics*, forthcoming.

[8] Conley, J.P. and H. Konishi, 2002, "Migration-proof Tiebout equilibrium: Existence and asymptotic efficiency," *Journal of Public Economics*, 86 (2), pp. 243-262.

[9] Ebenstein, A., 2010, "The 'missing girls' of China and the unintended consequences of the one child policy," *Journal of Human Resources*, 45(1), pp. 87-115.

[10] Gobbi, P.E., 2013, "A model of voluntary childlessness," *Journal of Population Economics*, 26(3), pp. 963-982.

[11] Gollin, D., D. Lagakos and M.E. Waugh, 2014, "The agricultural productivity gap," *Quarterly Journal of Economics*, 129 (2), pp. 939-993.

[12] Gottlieb, C. and J. Grobovšek, 2019, "Communal land and agricultural productivity," *Journal of Development Economics*, 138, pp. 135-152.

[13] Harris, J.R. and M.P. Todaro, 1970, "Migration, unemployment and development: A two-sector analysis," *The American Economic Review*, 60(1), pp. 126-142.

[14] Lagakos, D., 2020, "Urban-rural gaps in the developing world: Does internal migration offer opportunities?" *Journal of Economic perspectives*, 34(3), pp. 172-192.

[15] Laing, D., C. Park and P. Wang, 2005, "A modified Harris-Todaro model of rural-urban migration for China," in Critical Issues in China's Growth and Development in Honor of Gregory Chow (edited by F. Kwan and E. Yu), Ashgate Publishing Limited, pp. 245-264.

[16] Lewis, W.A., 1954, "Economic development with unlimited supplies of labour," *The Manchester School*, 22(2), pp. 139–191.

[17] Liao, P., 2013, "The one-child policy: A macroeconomic analysis," *Journal of Development Economics*, 101, pp. 49-62.

[18] Liao, P., P. Wang, Y. Wang and C. Yip, 2022, "Educational choice, rural-urban migration and economic development," *Economic Theory* 74, pp. 1-67.

[19] Lin, J.Y., 1992, "Rural reforms and agricultural growth in China," *The American Economic Review*, 82, pp. 34-51.

[20] Lucas, R.E., 2004, "Life earnings and rural-urban migration," *Journal of Political Economy*, 112(S1), pp. S29–S59.

[21] Ma, L. and Y. Tang, 2020, "Geography, trade, and internal migration in China," *Journal of Urban Economics*, 115, 103181.

[22] Ngai, L.R., C.A. Pissarides and J. Wang, 2019, "China's mobility barriers and employment allocations," *Journal of the European Economic Association*, 17(5), pp. 1617-1653.

[23] Sato, Y., 2007, "Economic geography, fertility and migration," *Journal of Urban Economics*, 61(2), pp. 372-387.

[24] Sato, Y. and K. Yamamoto, 2005, "Population concentration, urbanization, and demographic transition," *Journal of Urban Economics*, 58(1), pp. 45-61.

[25] Scharping, T., 1997, "Studying migration in contemporary China: models and methods, issues and evidence," *Cologne China Studies Online - Working Papers on Chinese Politics, Economy and Society.*

[26] Song, Z., K. Storesletten, Y. Wang and F. Zilibotti, 2015, "Sharing high growth across generations: pensions and demographic transition in China," *American Economic Journal: Macroeconomics*, 7(2), pp. 1–39.

[27] Todaro, M.P., 1969, "A model of labor migration and urban unemployment in less developed countries," *The American Economic Review,* 59, pp. 138-148.

Appendix (Online Publication Only)

In Appendix I, we document our motivating facts, institutional backgrounds, the data for quantitative analysis, and the results of the sensitivity and robustness tests. In Appendix II, we provide supplementary theoretical details, including the mathematical proofs for the baseline and for the generalized models.

Appendix I: Data, Institution, Quantitative Details and Robustness Tests

A. Motivating Facts

The first part of this subsection shows how we construct the measure of rural-to-urban migration intensity for country i in year t. In the second part, we use the migration intensity data in Bernard, Bell, and Cooper (2018) and present the patterns between initial relative income, migration intensity, and total fertility rates (TFRs).[32]

We start with population identity equations. Denote urban and rural employment for country i in year t as $L_{i,t}^{U}$ and $L_{i,t}^{R}$, respectively. The rural-urban migration flow $M_{i,t}$ is then given by:

$$\begin{aligned} L_{i,t+1}^{U} &= L_{i,t}^{U} + M_{i,t}; \\ L_{i,t+1}^{R} &= L_{i,t}^{R} - M_{i,t}. \end{aligned}$$

We further define rural-urban migration intensity as $MI_{i,t} = M_{i,t}/L_{i,t}^{R}$ and the urban-rural employment ratio as $UE_{i,t} = L_{i,t}^{U}/L_{i,t}^{R}$. Then we have:

$$UE_{i,t+1} = \frac{L_{i,t+1}^{U}}{L_{i,t+1}^{R}} = \frac{L_{i,t}^{U} + M_{i,t}}{L_{i,t}^{R} - M_{i,t}} = \frac{\frac{L_{i,t}^{U}}{L_{i,t}^{R}} + MI_{i,t}}{1 - MI_{i,t}}.$$

By rearranging the above equation, we obtain:

$$MI_{i,t} = \frac{\frac{L_{i,t+1}^{U}}{L_{i,t+1}^{R}} - \frac{L_{i,t}^{U}}{L_{i,t}^{R}}}{1 + \frac{L_{i,t+1}^{U}}{L_{i,t+1}^{R}}} = \frac{UE_{i,t+1} - UE_{i,t}}{1 + UE_{i,t+1}}.$$

The data for $L_{i,t}^{U}$ and $L_{i,t}^{R}$ are taken from Global Jobs Indicators Database (JOIN) of World Bank. $L_{i,t}^{U}$ is measured by urban employed population aged 15-64 and $L_{i,t}^{R}$ is rural employed population aged 15-64. Using the above $MI_{i,t}$ formula, we construct a time-series of rural-urban migration intensity for country i whenever the data is available. Then, the time-series of migration intensity is averaged from 1996 to 2015 to represent country i's rural-urban migration intensity. Besides, we take the data of GDP per capita (current US dollars) in 1996 from World Development Indicators (WDI) to calculate initial relative income for each country. Initial relative income is defined as country i's GDP per capita in 1996 relative to the United States' per capita GDP in 1996. We also obtain TFR data in 1996 for each country from WDI. Thus, we have the initial relative income in 1996, TFR in 1996, and rural-urban migration intensity for the period of 1996-2015 for each country.

[32] Bernard, A., M. Bell, and J. Cooper, 2018, "Internal migration and education: A cross-national comparison," Background Report to the 2019 Global Education Monitoring Report. Paris, France:UNESCO.

Next, all available countries are classified as developing and developed countries based on World Bank's classification in 1996. Developing countries include lower-middle and low-income countries, and developed countries are defined as upper-middle and high-income countries. Thus, we have 33 observations for developed countries and 83 observations for developing countries. The mean of migration intensity for developed countries is -0.0027 and for developing countries is 0.0108. The mean of TFR for developed countries is 1.754, while it is 4.173 for developing countries.

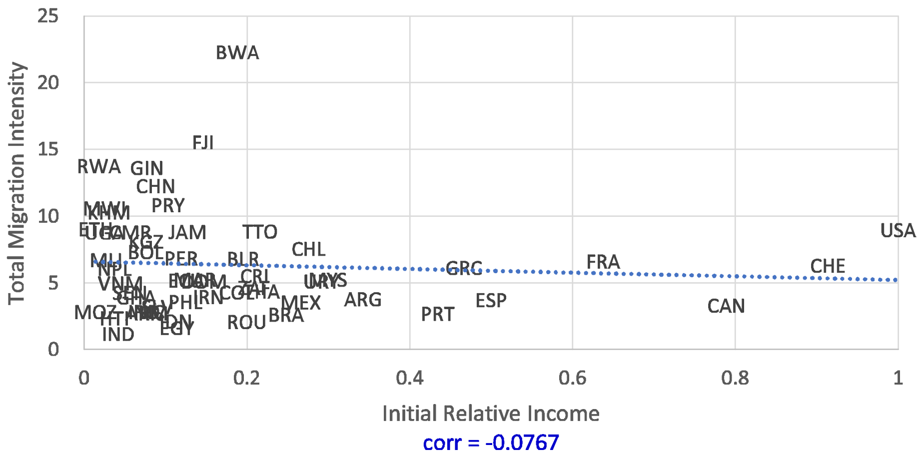
Taking one step forward, we are interested in whether developed and developing countries' migration intensities and TFRs are equal or not. We first test if the two samples are normally distributed. The results of Shapiro-Wilk W test suggest that neither migration intensity nor TFR in developed and developing countries follow a normal distribution. We thus choose Mann-Whitney test instead. As discussed in the introduction, the results of Mann-Whitney test suggest that both migration intensity and TFR in developing countries are significantly different from those in developed countries. In other words, rural-urban migration and fertility decisions are both related to the stage of economic development. This fact motivates the study of our paper.

Alternatively, we can use the migration intensity data in Bernard, Bell, and Cooper (2018) to be a measurement of migration. We merge the migration intensity data from Bernard, Bell, and Cooper (2018), the real GDP per capita data from Penn World Table (PWT 9.0), and the TFR data from World Development Indicators (WDI) to provide the two patterns again. As the migration intensity data in Bernard, Bell, and Cooper (2018) is based on census data from sample countries over the period of 1996-2011, the span of our data analysis is set from 1996 to 2011. Total migration intensity, now, is defined as total migrants as a ratio of total population aged 15 and above and is directly taken from Bernard, Bell, and Cooper (2018). Note that, different from our measure that is designed to capture rural-urban migration, the migration intensity data in Bernard, Bell, and Cooper (2018) includes total migrants, i.e., the sum of rural-rural, rural-urban, urban-rural and urban-urban migrants. Again, we choose the U.S. as our benchmark and compute the relative real GDP per capita to the U.S. in 1996. The results are plotted in Figure IA.1. While total migration intensity decreases with the relative income in 1996, with a correlation coefficient of -0.0767, TFR declines more sharply in initial relative income, with a correlation coefficient of -0.5965.[33] The pattern shown here is consistent with what we have shown in the main text.

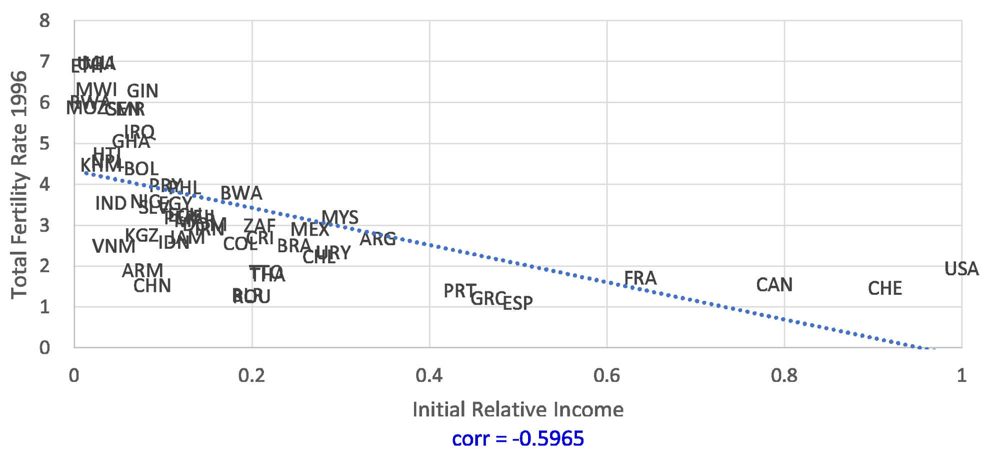
[33] By excluding the outlier, Boswana (BWA), the pattern in Figure IA.1(a) and IA.1(b) remains with the correlation coefficient improving slightly to -0.0886 and -0.5966, respectively.

Figure IA.1: Initial Relative Income, Migration Intensity, and Total Fertility Rate

(a) Migration Intensity



(b) Total Fertility Rate



B. Institutional Background in China

B.1 The *Hukou* System In the 1950s, China implemented an unique *hukou* system to solve the serious problem of rural-urban labor migration. The *hukou* system required everyone to register in one and only one place of his/her residence. A formal rural-urban migration required both appropriate documents and a migration permit (quota) issued by the government. The regular channels of rural-urban migration included recruitment by a state-owned enterprise (*zhaogong*), enrolment in an institution of higher education (*zhaosheng*), promotion to a senior administrative job (*zhaogan*), and displacement due to state-initiated land expropriation (*zhengdi*). Because the government tightly controlled the annual migration quotas and the *hukou* registration could not be changed freely, the *hukou* system formed a dual structure with no labor mobility but urban-rural distinctions.

The *hukou* system not only restricted internal migration itself, but also integrated with social and economic controls. In the rural areas, adults were bound to the land. They were belonged to communes and had to participate in agricultural production to obtain food rations for their households. For urban residents, work units were the primary administrative units to assign most social services. For example, a work unit determined grain rations, education for children, health benefits, and purchasing house. The above controls were effective because, in the pre-reform period, local governments predominated the distribution of daily necessities. Few of them were traded in the market. Besides, there were few urban jobs outside the state-owned enterprises. Therefore, it was difficult to survive outside one's *hukou* registration place without a work unit and a formal urban *hukou*.

Since 1980s, due to the economic reform, the growing market-oriented economy demanded more cheap labor. It became easier for people to find city jobs in private sectors and survive outside their *hukou* registration place. In addition, the economic reform also forced the government to relax the administrative control, such as the abolition of food rations and the commune system. These factors increased rural-urban migration and resulted in a significant temporary migration (floating population, who stayed outside their *hukou* registration place). The noticeable mobility forced the government to have its *hukou* regulation reforms. For example, temporary residence certificate was introduced to all urban areas in 1985 and extended to rural areas in 1995; local governments allowed "self-supplied food grain" urban *hukou*; and the "blue-stamp" urban *hukou* was conducted in the 1990s. Finally, in 2014, the central government announced to gradually implement unified urban-rural household registration system.

B.2 Land Reallocation and Migration As mentioned in Section B.1, rural adults were bound to lands and were belonged to communes. Because rural production could not be freely traded in the market, rural people had low incentives to work. To motivate rural production, household responsibility system (HRS) was introduced to replace the rural commune system in the early 1980s. Under the HRS, village collectives officially owned the land and distributed the use rights to households. Households only had fixed-term contracts to use land for production activities. Besides, the government gave each household a quota of goods to produce. Households received compensation for the quota. Above-quota production could be sold in the market at unregulated prices. This system motivated households to work harder and successfully increased rural production in a short period of time.

The distribution of land use rights was mainly based on population size within a household. After several years, village collectives had the right to reallocate land according to changes of household size. Therefore, the land rights for rural households were incomplete and the land tenure arrangement interacted with rural-urban migration decisions in the following ways. First, rural-urban migration resulted in a risk of land expropriation in the next land reallocation because rural-urban migration decreased household size, especially permanent migration. Second, it was unclear if households had the rights to rent the land they obtained or to transfer the contract. Since households did not have the property right of land, land could not be sold when households migrated to cities and might become idle land. This represented the opportunity cost of rural-urban migration.

To strengthen land tenure security, the Land Management Law and the Rural Land Contracting Law (RLCL) were implemented in 1998 and 2002, respectively. These laws focused on three reforms: (i) The duration of land contract should be 30 years; (ii) Large scale of land reallocations are not allowed. Only small-scale adjustments with clear conditions are agreed; (iii) Land transfer between households is permitted. However, despite the above land reforms, households remained facing the risk of land reallocation. For example, Deininger and Jin (2009) find that about 1/3 households out of their sample experienced land reallocations during 2002-2004. Tao and Xu (2007) find similar evidence.[34] On one hand, land expropriation is exacerbated due to urban expansion and infrastructure development. On the other hand, although land transfers are allowed after the reforms and the transfers do not affect the underlying contracts with the village collectives, the transfer market for land is still immature. Transfer contracts remain informal and are usually made with relatives.

B.3 The One-child Policy and Migration In the early 1970s, China introduced the "Later, Longer, and Fewer" family planning program to lower the fast population growth. However, the program did not successfully reach the ideal population growth rate. To quickly achieve the goal, the central government implemented the well-known one-child policy in 1979. The strict one-child policy successfully lowered China's total fertility rate to be less than two. However, when the one-child generation grew up, it was realized that the fast decline in fertility will result in a fast demographic change, shortage of labor force, and rapid population aging in the near future. Thus, the one-child policy was eventually abolished in 2015.

The principle of the one-child policy restricted each couple to have only one child. The policy was introduced by the central government but the implementation was conducted by local governments. Thus, the detailed rules differed between provinces, areas, and ethnic groups. In some provinces and ethnic minorities, more than one child was permitted under some special conditions. In principle, the implementation contained benefits for one-child families and penalties on the above-quota birth. The benefits and subsidies that one-child families could obtain included child allowance, priority access to schools, employments, health care, and housing. In rural areas, one-child families could even enjoy tax credit and receive more land. In contrast, the penalties on above-quota births would be 10-20 percent of family income lasting for 3-14 years. In particular, if parents worked in the SOE sector, which included the government sector and the state-owned enterprises, they would not eligible for promotion, be demoted or be forced to quit their jobs. Moreover, the above-quota children were not allowed to attend public schools.

The one-child policy also interacted with the *hukou* system and had impacts on migration decisions. Permanent migrants (formal rural-urban migration) had to follow the fertility policies of their destination cities but usually a transition period was allowed. The transition period ranged from two to five years. The rules in some cities were even stricter. For example, if the wife had not been pregnant when a couple obtained Beijing *hukou*, the certificate for having a second child they received in rural areas would be withdrawn. In contrast, fertility decisions of temporary migrants of course were restricted by their original places. However, the family planning officials who were responsible for temporary migrants' fertility had difficulty on tracking where the migrants move to. Thus, the growing amount of temporary migrants increased the difficulty of implementing the one-child policy in China. Under the one-child policy, it is a general practice that births happen after migration. Bernard, Bell and Cooper (2018) report that China's migration age at peak was aged 20.5. On the other hand, Chen (1991) suggests that the average age for the first birth was

[34] Deininger, K. and S. Jin, 2009, "Securing property rights in transition: Lessons from implementation of China's rural land contracting law," *Journal of Economic Behavior and Organization*, 70(1-2), pp. 22-38. Tao, R. and Z. Xu, 2007, "Urbanization, rural land system and social security for migrants in China," *Journal of Development Studies*, 43(7), pp. 1301-1320.

older than aged 23 during 1980-1987. It was even older during 2006-2016 (He et al., 2018). This justifies our model setting that births occur after migration.[35]

C. Data for Quantitative Analysis

- Retirement age

  The statutory retirement age in China is relatively young compared to other countries. The normal retirement age has been set at 60 years for men, 50 years for blue collar women and 55 years for white collar women in China since the 1950s. It is possible to claim a pension benefit from the age of 55 for men and 50 for women if the individual engaged in physical work in certain industries or posts (OECD, 2019).[36]

- Population

  High-skilled workers refer to those whose education level are senior high school or above. Those below senior high school level are classified as low-skilled workers.

  - Urban and rural population (U and R)

    The data of urban and rural population is directly obtained from *China Population and Employment Statistics Yearbook* 2010. We compute the urban-rural population ratios ($\frac{U}{R}$) for the years of 1980-2007. The average from 1980 to 2007 is 0.4579.

  - Rural high- and low-skilled population (H and L)

    China Rural Statistical Yearbook 1987-2008 report the statistics of rural workers' educational attainment. Using our definition of high-skilled workers, we thus compute the rural high-to-low-skilled ratio ($\frac{H}{L}$) for the years from 1985 to 2007. However, the data before 1985 is not available. We back out the ratios for 1980-1984 by the following steps. First, we compute the annual growth rate of the rural high-to-low-skilled ratio from 1985 to 2007 and the geometric average is 3.728 percent. Second, with the rural high-to-low-skilled ratio in 1985 and the average annual growth rate, we are able to calculate the rural high-to-low-skilled ratio from 1980 to 1984. The average of $\frac{H}{L}$ during 1980-2007 is then equal to 0.1076.

  - Urban SOE and private employment (S and P)

    The number of urban employments by sectors is available in *China Statistical Yearbook*. We define state-owned units as urban SOE sector and others to be urban private sectors (including collective-owned units, cooperative units, joint ownership units, limited liability corporations, share holding corporations Ltd., private enterprises, units with funds from Hong Kong, Macao and Taiwan, foreign funded units, and self-employed individuals). Then we obtain the urban SOE-private employment ratio ($\frac{S}{P}$) for the years of 1980, 1985, 1990, and 1995-2007. For the years without data (1981-1984, 1986-1989, and

[35] Chen, Y., 1991, "An analysis of Chinese women's changing age patterns of first marriage and first childbearing," *Chinese Journal of Population Science*, 5, pp. 39-45. He, D., X. Zhang, Y. Zhuang, Z. Wang and S. Yang, 2018, "China fertility status report, 2006-2016: An analysis based on 2017 China fertility survey," *Population Research*, 42(6), pp. 35-45.

[36] OECD, 2019, Pensions at a Glance 2019: OECD and G20 Indicators, OECD Publishing, Paris, https://doi.org/10.1787/b6d3dcfc-en.

1991-1994), we use linear interpolation to back out the corresponding urban SOE-private employment ratios. The average of $\frac{S}{P}$ during 1980-2007 is 1.5235. In addition, we also use the number of urban employments by sectors to compute the SOE-urban population ratio, $\frac{S}{U}$. According to the definition in our model, $U = S + P$. We thus use the number of employment in state-owned units as the numerator and the sum of employment in state-owned units and other units as the denominator to obtain $\frac{S}{U}$. Similarly, the years without data are backed up by linear interpolation. The average of $\frac{S}{U}$ during 1980-2007 is 0.6094.

– Urban private high- and low-skilled population (P^H and P^L)

  Urban Household Survey (UHS) provides the number of urban workers on two dimensions: workers' employment types and their education levels. We thus use UHS to compute the private high- to low-skilled worker ratio ($\frac{P^H}{P^L}$) with the following steps. First, we define workers in other economic units, private entrepreneurs in cities and towns, employees in private businesses, and other employments as private sectors, while those who already retired but re-entered the labor force are excluded. Second, given our definition of high-skilled workers, we compute the number of high- and low-skilled workers in private sectors and the ratio of high- to low-skilled workers for 1987-2007. Third, for the years without data (1980-1986), we calculate the annual growth rate of $\frac{P^H}{P^L}$ from 1987 to 2007 and have the geometric average growth rate. Then, linear interpolation is implemented to obtain the ratio of high- to low-skilled workers during 1980-1986. Finally, we obtain the average of $\frac{P^H}{P^L}$ during 1980-2007, which is 0.8159.

– Urban population (U) and the number of workers actually work in urban areas (U^+)

  Urban population (U) is directly obtained from *China Population and Employment Statistics Yearbook* 2010. The difference between U and U^+ is rural-urban floating population. As rural-urban migrants was the majority in China's internal migration, we assume the entire floating population reported in data is the rural-urban migrants. Using the floating population of 1982, 1987, 1990, 1995, 2000, and 2005 provided by Department of Population and Employment Statistics, National Bureau of Statistics of China (2008, appendix table 1), we first calculate the corresponding ratios of floating to urban population. Second, linear interpolation is implemented to compute the ratios for the years that the data is not available. For the year of 1980 and 1981, we calculate the two ratios based on the year of 1982 and the geometric average on annual growth rate of floating-to-urban population ratio during 1982-1987. Similarly, the ratios of 2006 and 2007 are computed based on the year of 2005 and the geometric average on annual growth rate of floating-to-urban population ratio during 2000-2005. Now we have a series of floating-to-urban population ratio from 1980 to 2007. Third, the floating-to-urban population ratios are converted to $\frac{U}{U^+}$ and the average of $\frac{U}{U^+}$ during 1980-2007 is 0.8764.

– Workers that actually work in urban areas (P^{H+}, P^{L+}, and S^+)

  To compute the workers that actually work in each sector, we need to know the floating population working in each sector. Unfortunately, the data is not observed. Using the Rural-Urban Migrant Survey (RUMS), Li (2010) estimates that the proportion of migrant workers employed as production workers, service workers, private enterprise owners, or self-employed reached 93 percent in 2007, while 81 percent of them were employed in private enterprises or self-employed.[37] Based on this information, we can infer that

[37] Li, S., 2010, "The economic situation of rural migrant workers in China," *China Perspectives*, [Online], DOI :

roughly 7 percent of migrants worked in the SOE sector in 2007, while at most 12 percent of migrants were private business owners. As it was relatively inflexible for the SOE sector to hire migrants in early years and our data period spans from 1980 to 2007, it is natural to see the 7 percent of migrants working in the SOE sector as an upper bound. We thus assume that 5 percent of the entire floating population ended up with jobs in the SOE sector, 20 percent worked as high-skilled workers in private sectors, and the rest (75 percent) found jobs as private low-skilled workers. Then we are able to compute the workers that actually work in urban areas ($\frac{P^{H+}}{P^{L+}}$ and $\frac{S^+}{P^{H+}}$) by the above population ratios and the population identity below:

$$\begin{aligned}\frac{P^{H+}}{P^{L+}} &= \frac{P^{H+}/U^+}{P^{L+}/U^+} = 0.6348,\\ \frac{S^+}{P^{H+}} &= \frac{S^+/U^+}{P^{H+}/U^+} = 3.0262,\end{aligned}$$

where

$$\begin{aligned}\frac{P^{H+}}{U^+} &= [\frac{P^H}{U} + \frac{\tilde{P}_F^H}{U} + \frac{\tilde{P}_I^H}{U}]\frac{U}{U^+} = \frac{P^H}{U}\frac{U}{U^+} + 0.2(1 - \frac{U}{U^+}),\\ \frac{P^{L+}}{U^+} &= [\frac{P^L}{U} + \frac{\tilde{P}_F^L}{U} + \frac{\tilde{P}_I^L}{U}]\frac{U}{U^+} = \frac{P^L}{U}\frac{U}{U^+} + 0.75(1 - \frac{U}{U^+}),\\ \frac{S^+}{U^+} &= [\frac{S}{U} + \frac{\tilde{S}_F}{U}]\frac{U}{U^+} = \frac{S}{U}\frac{U}{U^+} + 0.05(1 - \frac{U}{U^+}),\\ \frac{P^H}{U} &= (1 - \frac{S}{U})(\frac{P^H/P^L}{1 + P^H/P^L}),\\ \frac{P^L}{U} &= (1 - \frac{S}{U})(\frac{1}{1 + P^H/P^L}).\end{aligned}$$

- Workers that actually work in rural areas (R^+)

  To obtain the $\frac{R}{R^+}$, the following population identities are applied. First, we define $\tilde{U} = \tilde{S}_F + \tilde{P}_F^H + \tilde{P}_F^L + \tilde{P}_I^H + \tilde{P}_I^L$ and thus we have $R^+ = R - \tilde{U}$ and $\frac{R}{R^+} = \frac{1}{1-\frac{\tilde{U}}{R}}$. Similarly, in urban areas, we know that $\frac{U}{U^+} = \frac{U}{U+\tilde{U}} = \frac{1}{1+\frac{\tilde{U}}{U}}$. Rearrange to obtain $\frac{\tilde{U}}{U} = \frac{U^+}{U} - 1$. Therefore, $\frac{\tilde{U}}{R} = \frac{\tilde{U}}{U}\frac{U}{R} = (\frac{U^+}{U} - 1)\frac{U}{R}$. Finally, we have:

$$\frac{R}{R^+} = \frac{1}{1 - \frac{\tilde{U}}{R}} = \frac{1}{1 - (\frac{1}{U/U^+} - 1)\frac{U}{R}}$$

  With the ratios of $\frac{U}{U^+}$ and $\frac{U}{R}$ mentioned above, we have the average of $\frac{R}{R^+} = 1.069$ or $\frac{R^+}{R} = 0.9354$ during 1980-2007.

- Rural land

  - Duration of land reallocation (δ)

    According to Article 27 of *Notes to Law of the People's Republic of China on Land Contract in Rural Areas (Zhonghua Renmin Gongheguo nong cun tu di cheng bao fa shi*

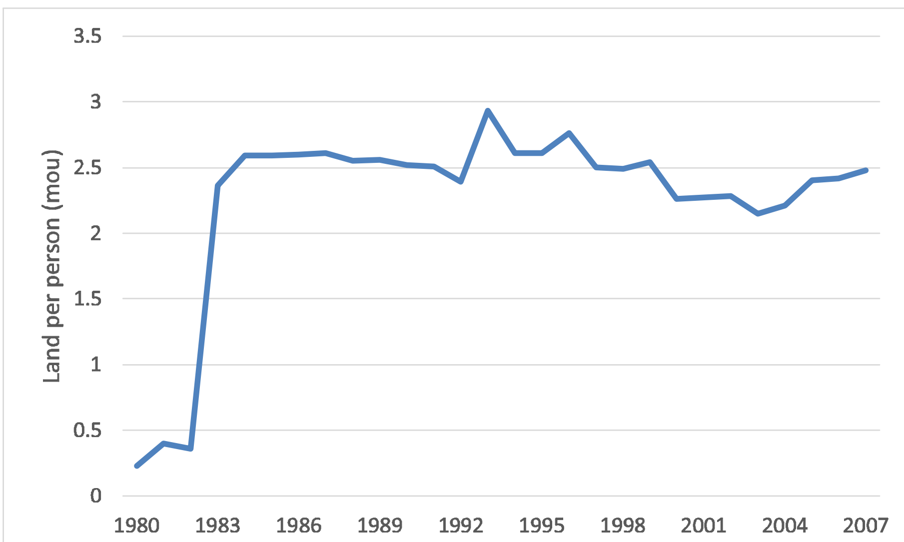
https://doi.org/10.4000/chinaperspectives.5332.

yi) implemented in 2003, the duration of land reallocation in general should be between 5 to 10 years. To have a representative value of the duration, we thus take the average and adjust by the model period to obtain $\delta = \frac{(5+10)/2}{36} = 0.2083$.

- Land per person (q)

  The data of average land per person is taken from *China Statistical Yearbook* 1987-2007 and *Area of Land Managed by Rural Households* in National Data provided by National Bureau of Statistics of China. Average land per person includes two items: area of cultivated land under management by rural households and hilly area under management by rural households. We use the average of land per person during 1980-2007 in our calibration. Thus, q is 2.2564 mou per person. The time series of the land per person from 1980 to 2007 is plotted in Figure IC.1. The big jump in 1983 is due to the land reform from the collective system to the household responsibility system. Lin (1992) indicates that such reform completed in 1983 and the proportion of production teams that had adopted the HRS maintained at close to 100 percent after 1983.[38] This supports the assumption of stable q in our calibration. For more discussion to China's urban land institutions, the reader is referred to Cai, Henderson, and Zhang (2013).[39]

Figure IC.1: Land per person



- Wages

  - Rural per capita income (y_R)

    It is normalized to one during 1980-2007.

[38] Lin, J. Y., 1992, "Rural reforms and agricultural growth in China," *The American Economic Review*, 82, pp. 34-51.

[39] Cai, H., J.V. Henderson, and Q. Zhang, 2013, "China's land market auctions: evidence of corruption?," *Rand Journal of Economics*, 44(3), 488–521.

- Urban wages (w_S, w_P^H, and w_P^L)

  Similar to the number of urban workers, UHS also provides information on workers' income. We thus use UHS to compute urban income in each sector with the following steps. First, we define workers in other economic units, private entrepreneurs in cities and towns, employees in private businesses, and other employments as private sectors, while those working in state-owned units and collective economic units in cities and towns are classified as the SOE sector. We exclude those who already retired but re-entered the labor market. Second, we only consider wage-related income to be a worker's wage here. Third, given our definition of high-skilled workers, we are able to compute the average wage of high-skilled workers from 1987 to 2007. Similarly, the average wage of low-skilled workers from 1987 to 2007 are calculated. Then we obtain a series of skill premium ($\frac{w_P^H}{w_P^L}$) during 1987-2007. Fourth, for the years without data (1980-1986), we calculate the annual growth rate of skill premium from 1987 to 2007 and have the geometric average growth rate. Interpolation is implemented to obtain the skill premium during 1980-1986. Thus the average skill premium during 1980-2007 is 1.3944. Finally, using the definition of the SOE sector, we obtain the average wage of workers in the SOE sector during 1987-2007. A series of $\frac{w_P^H}{w_S}$ during 1987-2007 is computed. Based on the data of 1987 and the geometric average growth rate, we obtain $\frac{w_P^H}{w_S}$ for the period between 1980-1986. Therefore, the average $\frac{w_P^H}{w_S}$ during 1980-2007 is 1.1346. This implies the average private-sector premium ($\frac{w_P^H}{w_S}$) is about 13 percent in our calibration. During the periods that we examine, such premium turned negative in two sub-periods, 1988-1991 and 2001-2007. The former sub-period is due to the 1984 reform of the Contract Responsibility System (*cheng bao zhi*) that led to higher productivity and hence a drop of the private sector premium. After Deng Xiao Ping's Southern Tour in 1992 that further opened the China market for domestic and foreign-owned private businesses, there was a *xiagang* act over the second half of 1990s, laying off 40 percent of the SOE workers, causing a drop in the private sector premium again (Lin, Lu, Zhang, and Zheng 2020).[40]

- Urban premium ($\frac{y_U}{y_R}$)

  Table 10-2 in *China Statistical Yearbook 2011* reported disposable income per capita in urban and rural areas for 1978, 1980, 1985, and 1990-2007. We obtain urban and rural real per capita income by considering price changes (the base year is 1978). Then, urban premium (the ratio of urban-rural real per capita income) of 1980, 1985, and 1990-2007 are obtained. For the years without data (1981-1984 and 1986-1989), linear interpolation is applied to compute the corresponding urban and rural real per capita income. Then, the urban premium is computed accordingly. With the entire series, we then compute the average of urban premium from 1980 to 2007, which is 1.9641.

- Elasticity of substitution between high- and low-skilled workers (σ)

  In the literature, the estimates for developed countries range from 1 to 3 (see, for example, Autor, Katz, and Krueger 1998; Acemoglu 2003; and Ciccone and Peri 2005), while the estimated values for Asian economies are larger, mostly falling between 2 and 7.[41] For example,

[40] Lin, K.J., X. Lu, J. Zhang, and Y. Zheng, 2020, "State-owned enterprises in China: A review of 40 years of research and practice," *China Journal of Accounting Research*, 13(1), pp. 31-55.

[41] Autor, D.H., L.F. Katz, and A.B. Krueger, 1998, "Computing inequality: Have computers changed the labor

Toh and Tat (2012) estimated that the value for Singapore is 4.249.[42] Te Velde and Morrissey (2004) used data from Singapore, Hong Kong, Korea, the Philippines, and Thailand and obtained a value of 2.78.[43] The results in Gindling and Sun (2002) imply that the value in Taiwan is between 2.3 and 7.4.[44] We thus set $\sigma = 0.8333$, which corresponds to an elasticity of substitution between high- and low-skilled workers to be 6. This value of the EIS between high- and low-skilled workers falls within the range of the estimates for East Asian countries.

- Quality-based altruistic factor (θ)

  We adopt Leontief preference in the theoretical model and obtain the relation $\frac{c}{nb} = \frac{1-\theta}{\theta}$. This implies that $\theta = \frac{1}{\frac{c}{nb}+1}$. To calibrate θ, we use the information on average rural per capita income and consumption in *China Statistical Yearbook*. The rural consumption-income ratio is obtained by average consumption expenditure minus average expenditure on educational and recreational activities and other expenditure over average rural income. The rural saving-income ratio is average rural income minus average consumption expenditure over average rural income. All are in per capita term. With the series of the rural consumption-income ratio and the rural saving-income ratio, we are able to compute the consumption-saving ratio ($\frac{c}{nb}$) and obtain a series of θ during 1980-2007. The average of θ from 1980 to 2007 becomes 0.2124.

- Imputed fertility (n_R^L, n_R^H, n_P^L, n_P^H, and n_S)

  For calibration, we need the fertility of parent with different skill types and working in different sectors. Unfortunately, the corresponding data is not available. We thus adopt the total fertility rates by education levels in Retherford, Choe, Chen, Li, and Cui (2005) to be the base and use population distribution of education level in rural, urban, or SOE sector as a weight to compute a weighted fertility for each type.[45] The population distributions of education level in rural areas are obtained from *China Rural Statistical Yearbook* 1987-1995 and 1997-2008. The population distributions of education level in urban areas are from *China Population Statistics Yearbook* 1988, *10 Percent Sampling Tabulation on the 1990 Population Census of the People's Republic of China*, *1995 China 1% Population Sample Survey*, and *The Tabulation on the 2000 Population Census of the People's Republic of China*. For the SOE workers, census data is not available. We thus use the information of SOE workers' education levels in UHS to compute the distribution. The definition of the SOE sector is the same as that in calculating w_S (those working in state-owned units and collective economic

market?" *The Quarterly Journal of Economics*, 113(4), pp. 1169-1213. Acemoglu, D., 2003, "Cross-country inequality trends," *The Economic Journal*, 113, F121-F149. Ciccone, A. and G. Peri, 2005, "Long-run substitutability between more and less educated workers: Evidence from U.S. states, 1950-1990," *The Review of Economics and Statistics*, 87(4), pp. 652-663.

[42] Toh, R. and H. Tat, 2012, "Trade liberalization, labor demand shifts and earnings inequality in Singapore," *Review of Urban and Regional Development Studies*, 24(3), pp. 65-82.

[43] Te Velde, D. and O. Morrissey, 2004, "Foreign direct investment, skills and wage inequality in East Asia," *Journal of the Asia Pacific Economy*, 9(3), pp. 348-369.

[44] Gindling, T.H. and W. Sun, 2002, "Higher education planning and the wages of workers with higher education in Taiwan," *Economics of Education Review*, 21(2), pp. 153-169.

[45] Retherford, R., M.K. Choe, J. Chen, X. Li and H. Cui, 2005, "How far has fertility in China really declined?" *Population and Development Review*, 31(1), pp. 57–84.

units in cities and towns). Specifically, to be consistent with the assumption in our theoretical model, we only consider those whose education level are senior high school or above when n_S is computed. The imputed fertility in various years are reported in Table IC.1. Our quantitative analysis studies the period of 1980-2007. However, the last total fertility rate that Retherford, Choe, Chen, Li, and Cui (2005) reported is that of 1996-2000. Thus, our imputed fertility also ends in 1996-2000. Besides, in our theoretical framework, we assume there is only one parent in a household. We therefore match a half of the imputed fertility in Table IC.1 in the calibration.

[Insert Table IC.1 about here]

- Basic above-quota penalty ($\bar{\phi}_R$ and $\bar{\phi}_U$)

  Because the above-quota penalty could be non-monetary and there is no suitable measurement to measure the overall above-quota penalty, we assume low-skilled workers bear the basic above-quota fine $\bar{\phi}_R$ in rural areas and $\bar{\phi}_U$ in urban areas. Define $\tilde{\phi}_R$ to be the proportion of basic above-quota fine to a rural agent's wage and $\tilde{\phi}_U$ is the proportion of basic above-quota fine to an urban agent. Then, we use the information provided by Ebenstein (2010) to impute $\tilde{\phi}_R$ and $\tilde{\phi}_U$ by the following steps.[46] First, Ebenstein (2010) provides above-quota fine rates and fertility policies for provinces in China during 1979-2000. If the fertility policy of a province in a year is equal to one, the province in that year is classified as urban areas; otherwise, a province is defined to be rural areas if its fertility policy is greater than one. Following this classification, we have two subgroups: rural and urban groups. Second, we compute the simple average of the above-quota fines in the urban group to obtain $\tilde{\phi}_U = 1.6654$. Similarly, the simple average of the above-quota fines in the rural group is $\tilde{\phi}_R = 1.5928$.

- Urban childrearing cost as a percentage of urban per capita income ($\tilde{\phi}_U^0$)

  As there is no nationwide survey on urban childrearing cost in China during the period of 1980-2007, we are forced to rely on estimates and surveys for selected cities in China. According to Ye and Ding (1998), the total raising cost of a child from age 0 to 16 accounts for 34 percent and 20 percent of family income for Ximen and Beijing during 1995-1996, respectively.[47] Zhang (2000) reports that for Zhengzhou in 1998, 56 percent of households spent less than 30 percent of family income on child.[48] We therefore set $\tilde{\phi}_U^0$ at 25 percent, a value falling within the range of the figures reported by Ye and Ding (1998) and Zhang (2000).

- Proportion of rural households getting less enjoyment from having children (ζ)

  Table 1 in Hermalin and Liu (1990) provides a distribution of family size preferences in surveys in selected areas of China with various years.[49] We define people who desire less than or just

[46] Ebenstein, A., 2010, "The 'missing girls' of China and the unintended consequences of the one child policy," *Journal of Human Resources*, 45(1), pp. 87–115.

[47] Ye, W. Z. and Y. Ding, 1998, "The cost of child care in Xiamen Special Economic Zonein China," *Population and Economics*, No.6, pp.24-57 (in Chinese).

[48] Zhang, Y.Z., 2000, " A Survey on the cost of raising children aged 0-16 in Zhengzhou City," *Marketing Research*, No.6, pp. 35-37 and 55 (in Chinese).

[49] Hermalin, A.I. and X. Liu, 1990, "Gauging the validity of responses to questions on family size preferences in China," *Population and Development Review*, 16(2), pp. 337-354.

one child as $\underline{\beta}$-type agents and focus on rural or suburban areas. Then, the proportion of rural households who gets less enjoyment from having children is around 15.37%. Besides, Table 19 in Scharping (2003) also reports surveys on family size preferences for 1980-1997. Following the same definition of $\underline{\beta}$-type agents and focus on rural or suburban areas, we obtain 15.29%. Based on these two results, we set $\zeta = 0.15$.[50]

- Probability of obtaining urban residency (ρ) and the fraction of lifetime to enjoy urban benefits/pay urban embarked tax (μ)

  Regarding the parameters for the institution of the *hukou* system, μ and ρ, we resort to the literature. Prior to 1994, it was very difficult for rural migrants to obtain urban *hukou*. The blue-stamp system, implemented in 1994, opened the door to rural migrants. According to Liu (2005), it took two to five years for rural migrants to get urban *hukou* via the blue-print system.[51] We therefore compute μ based on this information. Assume that rural migrants who worked in urban private sector were not qualified for urban benefits prior to 1994 because they could not obtain urban *hukou*, i.e. $\mu = 0$ prior to 1994. After 1994, due to the relaxation of the blue-print system, it took an average of $\frac{2+5}{2} = 3.5$ years for a migrant to obtain urban residency and hence to be qualified for urban benefits. The period we examine is 1980-2007, and 1980-1994 roughly amounts to half of the examined period. Therefore, the average μ, the fraction of lifetime that a migrant worker successfully obtained urban residency to enjoy urban benefits, is $1 - \frac{0+3.5}{2}/36 = 0.4514$ for the period of 1980 to 2007. As for ρ, according to Wu and Treiman (2004), only about 11 percent of the interviews from rural areas successfully obtained urban residency. We thus set ρ at 0.11.[52]

- Urban benefit (B)

  In our baseline model, all urban benefits are summarized by an additional utility term B. That is, it can be viewed as public goods and services (net of any negative externalities). For example, Au and Henderson (2006) find most Chinese cities are undersized, suggesting urban amenities exceeding urban disamenities from negative externalities.[53] However, there is no perfect measure to proxy urban benefit. We thus use retirement benefit to represent urban benefit. The retirement benefit is the sum of pension he or she receives during the years that a retired agent is still alive. During the periods that we study, the life expectancy was about 75 years and the average retirement age was 53 years old. Besides, the average replacement rate (for all retirees) in China's pension system from 1980-1995 was about 81.67%, as reported in Table 2 of West (1999).[54] Therefore, we obtain the urban benefit using the replacement rate multiplied by urban per capita income (model value) and the number of years alive after retirement. The urban benefit is adjusted by the model period and $B = 0.9803$.

[50] Scharping, T., 2003, *Birth Control in China 1949-2000: Population Policy and Demographic Development*, Routledge.

[51] Liu, Z., 2005, "Institution and inequality: the hukou system in China," *Journal of Comparative Economics*, 33(1), pp. 133–157.

[52] Wu, X. and D.J. Treiman, 2004, "The household registration system and social stratification in China: 1955–1996," *Demography*, 41, pp. 363–384.

[53] Au, C.C. and J.V. Henderson, 2006, "Are Chinese Cities Too Small?," *Review of Economic Studies*, 73(3), pp. 549-576.

[54] West, L.A., 1999, "Pension reform in China: Preparing for the future," *The Journal of Development Studies*, 35(3), pp. 153-183.

- Rate of reverse migration (Λ)

  To calibrate Λ, the following procedure is applied. First, the average rural-urban population ratio in 1980-2007 is 0.4579. Together with the normalization of total population to one, we have $R = 0.6859$ and $U = 0.3141$. Second, the ratio of rural-urban migration to rural population in our benchmark economy is 0.0646 and rural-urban migration as a percentage of total migration is 0.49 in Scharping (1997).[55] We thus obtain rural-urban migration as a percentage of total population is 0.0443. Third, Scharping (1997) also reported that urban-rural migration as a percentage of total migration is 0.04. This implies that the urban-rural migration as a percentage of total population is 0.0036. Finally, we are able to compute the ratio of urban-rural migration to urban population, $\Lambda = \frac{0.0036}{0.3141} = 1.1516\%$.

- Markup of childrearing cost for left-behind children (Ξ)

  Tong, Yan, and Kawachi (2019) find that migrant household family income in the destination is significantly associated with migrant parents' arrangement to their children.[56] In the generalization for left-behind children, we thus assume that all low-skilled migrant parents are not able to obtain urban *hukou* and leave their children behind. Low-skilled rural-urban migrant parents bear a markup of childrearing cost for left-behind children. We utilize remittance data to calibrate the markup. Two measures are considered: aggregate and household-based measures. From the aggregate perspective, Messinis and Cheng (2007) use the estimation by China's Ministry of Agriculture to obtain that about 98 million of migrant workers in 2003 providing a total of 370 billion RMB (Ren Min Bi) remittances, accounting for 40 percent of total rural household income in China.[57] Notably, remittances are for not only left-behind children but also other family members, particularly the elderly. As there were about 69.7 million children left behind in 2013 (Tong, Yan, and Kawachi, 2019) and total rural population amounted to 637.8 million in 2013 (from World Bank), left-behind children thus were about $69.7/637.8 = 0.1093$ of rural population. One may thus take 10.93 percent of the remittances as for left-behind children, which constitutes $0.1093 \cdot 0.4 = 0.0437$ of total rural household income. In our calibration, 0.4386 is childrearing cost per child out of rural household income. Hence, the aggregate markup is $\Xi = (0.0437 + 0.4386)/0.4386 = 1.0996$. Alternatively, we can measure the remittances from household-based perspective. Based on a large-sample (China Poverty Monitoring Survey during 1997-2001) and a relatively small dataset (China Rural Poverty Survey), Du, Park, and Wang (2005) estimate that remittances increase rural household income by 8.5-13.1 percent, averaging 10.8 percent.[58] Given that imputed fertility is 2.2 and average rural household size is about 3.2. One may take $2.2/3.2 = 0.687$ of the remittance as for left-behind children, which is about $0.108 \cdot 0.687/2.2 = 0.0337$ per child per dollar of rural household income. This gives a household-based markup of $\Xi = (0.0337 + 0.4386)/0.4386 = 1.0768$. The

[55] Scharping, T., 1997, "Studying migration in contemporary China: models and methods, issues and evidence," *Cologne China Studies Online - Working Papers on Chinese Politics, Economy and Society.*

[56] Tong, L., Q. Yan, and I. Kawachi, 2019, "The factors associated with being left-behind children in China: Multilevel analysis with nationally representative data," PLoS ONE, 14(11):e0224205.

[57] Messinis, G. and E. Cheng, 2007, "The value of education and job training in the developing world: New evidence from migrant workers in China," Working Paper Series No. 36, Centre for Strategic Economic Studies, Victoria University.

[58] Du, Y., A. Park, and S. Wang, 2005, "Migration and rural poverty in China," *Journal of Comparative Economics*, 33(4), pp. 688-709.

two measures are comparable, so we simply average them to obtain the markup $\Xi = 1.0882$. It is worthy to notice that in our quantitative analysis the childrearing cost is proportional to parent's income. Thus, the cost of raising a left-behind child is $\Xi\tilde{\phi}_R^0 w_P^L$, as the parent is a private low-skilled worker. To be consistent with reality that left-behind children actually live in rural areas, we use the wage ratio of rural to private low-skilled worker in the benchmark economy to adjust the cost of raising a left-behind child. That is, we use $(\frac{W_R}{w_P^L})\Xi\tilde{\phi}_R^0 w_P^L$ to be the cost of raising a left-behind child in the experiment.

- Directed urban benefits

Recall that in the generalized framework, we allow urban benefits to be directed to (i) human capital at work (B_W), (ii) child education (B_E), or (iii) child care (B_C), beyond pure amenities (B_A) that is considered as B in the baseline model. We now illustrate how to compute $\bar{G}_m$ and B_m, where $m = A, W, E, C$.

$\bar{G}_W$ is computed using the data, financing and expenditure of re-employment service in enterprises, in *China Labor Statistical Yearbook 2002-2005*. We first calculate the proportion of total funding that is from government. Second, we compute this proportion as a percentage of average GDP from 2001 to 2004. Then, we calculate the urban output share by assuming that urban output includes industry and services sectors (78.49 percent of aggregate output). Finally, the percentage is adjusted by urban output share to obtain $\bar{G}_W = 0.0043$. In addition, we obtain B_W by the following steps. First of all, using survey-based *xiagang* data in 2000, Bidani, Goh, and O'Leary (2009) find such government sponsored training improved job finding in a moderate unemployment city, Wuhan, by 8 percent, but reduced job finding in a high unemployment city, Shengyang, by 6 percent.[59] On balanced we set the improvement in job finding as 1 percent. In both cities, re-employed workers did not have significant wage gains. Thus, the human capital effect can be measured purely by re-employment with faster job finding rates. Second, based on China Urban Labor Survey, Giles, Park, and Cai (2006) show the half life of the unemployment spell for a 30-40 years old man and woman was about 48 and 56 months, respectively, averaging 52 months.[60] Thus, a 1 percent improvement in job finding probability translates to $(0.01/0.5)(52/12) = 0.0867$ year reduction in unemployment spell. Third, only unemployed workers enjoy such reduction in unemployment spell, so the unemployment spell is adjusted by urban unemployment rate 7.91 percent (Liao, Wang, Wang, and Yip 2022) to become 0.006858, which can be interpreted as a percentage of the average annual wage to obtain the value of the human capital effect.[61] Note that in the generalized model, only SOE and private high-skilled workers are qualified for the benefit. However, SOE and private high-skilled workers' benefits are different. Define $B_{W,S} = 0.006858 \cdot w_s$ to be the benefit obtained by a SOE worker and $B_{W,P^H} = 0.006858 \cdot w_P^H$ as the benefit for a private high-skilled worker. B_W is then computed as the weighted average of $B_{W,S}$ and B_{W,P^H} (using the population of each type as the weight) and $B_W = 0.0143$ in the calibration. Finally, using the government technology for directed urban benefits to human capital at work, we obtain $B_{0,W} = 3.3354$.

$\bar{G}_E$ is measured by relative government spending in education. Li and Luo (2010) estimate that

[59] Bidani, B., N. Blunch, C. Goh, and C.J. O'Leary, 2009, "Evaluating job training in two Chinese cities," *Journal of Chinese Economic and Business Studies*, 7(1), pp. 77-94.

[60] Giles, J., A. Park, and F. Cai, 2006, "Reemployment of dislocated workers in urban China: The roles of information and incentives," *Journal of Comparative Economics*, 34(3), pp. 582-607.

[61] Liao, P., P. Wang, Y. Wang, and C. Yip, 2022, "Educational choice, rural-urban migration and economic development," *Economic Theory* 74, pp. 1-67.

local government educational subsidy per capita in urban and rural areas in 2002 are about 482 and 282 RMB, respectively.[62] The corresponding urban and rural per capita incomes are 8,084 and 2,592 RMB, respectively. By normalizing rural subsidy to zero, the incremental urban subsidy is 200 RMB per capita, or $200/8084 = 0.02474$ of urban income. Therefore, in our calibration $\bar{G}_E = 0.02474 \cdot y_U = 0.0486$. To compute B_E, more works are required. Note that in the generalized model, $c = \frac{(1-\theta)B_E}{\theta} nb$. Because nb is the real dollar measure of child quality, B_E can be interpreted as improvement probability (pr) times income gap between high- and low-skilled (gap). The improvement probability pr can be measured as the advantage that a rural born child enjoys if he or she is moving to cities. That is, once a rural born child moves to cities, his or her chance of becoming high-skilled worker is higher. We thus define pr to be:

$$pr = \frac{\text{prob. of urban children going to college}}{\text{prob. of rural children going to college}} - 1.$$

Table 3 in Gou (2006) provides the probabilities of going to college for urban and rural children, respectively.[63] The simple average of pr during 1989-2005 is 0.2619. The income gap between high- and low-skilled (gap) is measured by Micerian returns to college education. Immediately after college expansion, Zhang et al. (2005) estimate the returns of college versus high school to be 37.3-38.7 percent over 1999-2001, averaging 38 percent.[64] Then we have $gap = 0.38 \cdot w_P^L$. We thus compute $B_E = pr \cdot gap = 0.14$. Finally, with urban benefit production technology, we obtain $B_{0,E} = 2.8817$.

Regarding the directed urban benefits to child care, Du and Dong (2013) discusses that childcare subsidies in China are only to SOE workers and government employees, not to private-sector workers.[65] Thus, in the quantitative analysis, we modify the budget constraints and optimization conditions, accordingly. Based on Chinese Enterprise Social Responsibility Survey conducted in 2006, Du and Dong (2013) find that, despite a cutback of government sponsored public child centers or kindergartens, there were still about 20 percent of kindergartens run by SOEs and charged nonemployees 1000-10000 RMB per year. That is, the subsidy is about $0.2 \cdot (1000 + 10000)/2 = 1100$ RMB, which is the nominal average subsidy to childcare from SOEs to their employees in 2006. In 2006, the average nominal disposable income of urban households is 11759.5 RMB. Thus, the childcare subsidy for SOE employees can be computed as $1100/11759.5 = 0.09354$ of urban household income and $B_C = 0.09354 \cdot w_S = 0.1891$ in the calibration. As B_C is measured by government subsidies, we have $\bar{G}_C = B_C = 0.1891$. Finally, from urban benefit production technology, we obtain $B_{0,C} = 1$.

The last directed urban benefit is B_A. We follow the same strategy that we used in the baseline calibration to compute B_A. Note that in the baseline calibration, we measured the pure amenities by the retirement benefit computed from replacement ratio. All urban residents with hukou enjoy

[62] Li, S. and C. Luo, 2010, "Re-estimating the income gap between urban and rural households in China," *Procedia Social and Behavioral Sciences*, 2, pp. 7151-7163.

[63] Gou, R., 2006, " Examing equality in higher education from the perspective of rural-urban access to higher education," Research in Educational Development, 5, pp. 29-31 (in Chinese).

[64] Zhang, J., Y. Zhao, A. Park and X. Song, 2005, "Economic returns to schooling in urban China1988-2001," *Journal of Comparative Economics*, 33, pp. 730-752.

[65] Du, F. and X. Dong, 2013, "Women's employment and child care choices in urban China during the economic transition," *Economic Development and Cultural Change*, 62(1), pp. 131-155.

the benefit of B_A, thus it is computed as

$$B_A = \text{replacement rate} \cdot y_U \cdot \frac{75 - 53}{\text{length of model period}},$$

where 75 refers to the expected life expectancy and age 53 is the average retirement age. We then have $B_A = 0.9803$. We further assume the government maintains a balanced budget to compute $\bar{G}_A$ and obtain $\bar{G}_A = 0.3263$. Finally, $B_{0,A} = 3.0042$, which is calibrated by the urban benefit production technology. The calibration result is provided in Table IC.2. As our calibration targets remain unchanged, the calibration result for the generalized model with directed urban benefits is very similar to the baseline economy.

[Insert Table IC.2 about here]
[Insert Table IC.3 about here]

We further summarize the proportion of per capita urban taxes that is allocated to each benefit in Table IC.3. The share of pure amenities is the largest (82.11 percent), followed by the share of benefits for child education (12.23 percent). Directed urban benefits to human capital at work take the smallest share (0.84 percent). This distribution justifies the setup in our baseline model that we focus on the main directed urban benefits (pure amenities).

D. Sensitivity and Robustness Tests

Two sets of robustness tests are provided here. The first set of robustness tests varies the assumption on the proportions of floating population that are distributed to be SOE, private high- and low-skilled workers. The second set provides sensitivity tests on α, σ, τ and the positive private-sector premium ($\frac{w_P^H}{w_S}$).

D.1 Testing the Assumption on the Proportions of Floating Population In Section 5.2, to compute workers that actually work in each sector, we have assumed that 5 percent of the entire floating population ended up with jobs in the SOE sector, 20 percent worked as high-skilled workers in the private sector, and the rest (75 percent) found jobs as private low-skilled workers. In this appendix we provide the sensitivity and robustness tests for this assumption.

Recall that the ratios of the stocks of the sectoral specific within-the-period workers change as the assumption on the fractions of floating population in each sector change. As we have targeted the ratios of the stocks of the within-the-period workers in calibration, altering the assumption means that we have to recalibrate the model. For easy communication, in the following, we will call the calibration in the main text the "benchmark calibration".

Denote φ_S, φ_P^H, and φ_P^L as the fractions of the floating population ending up working as SOE, private-sector high- and private-sector low-skilled workers, where $\varphi_S + \varphi_P^H + \varphi_P^L = 1$. Based on the data in Li (2010), we consider four plausible assumptions on φ_S, φ_P^H, and φ_P^L: (1) $\varphi_S = 5\%$, $\varphi_P^H = 22\%$, and $\varphi_P^L = 73\%$; (2) $\varphi_S = 5\%$, $\varphi_P^H = 18\%$, and $\varphi_P^L = 77\%$; (3) $\varphi_S = 7\%$, $\varphi_P^H = 20\%$, and $\varphi_P^L = 73\%$; and (4) $\varphi_S = 3\%$, $\varphi_P^H = 20\%$, and $\varphi_P^L = 77\%$. The first two robustness tests fix φ_S at the value in the benchmark calibration while alter φ_P^H and φ_P^L by 2 percent, and the latter two robustness tests fix φ_P^H at 20 percent as in the benchmark calibration, adjusting φ_S and φ_P^L by 2 percent. Table ID.1 reports the summary of the calibrated parameters and variables under the four assumptions. Table ID.2 shows the relative productivity, relative childrearing costs and relative above-quota penalties, fractions of moving, population ratios, total fertility rates, and income ratios for the four robustness tests. Table ID.3 presents the policy experiment results under the four assumptions. For easy comparison, we provide in each table the values of the benchmark calibration.

[Insert Table ID.1 about here]
[Insert Table ID.2 about here]
[Insert Table ID.3 about here]

Not surprisingly, the probability of being recruited as SOE workers (π) and the fractions of workers migrating to cities (Γ_H and Γ_L) change as the assumption on the fractions of floating population in each sector change; so do the ratios of the within-the-period worker stocks. As we have targeted the same urbanization rates, fertility rates, income ratios, and fraction of newly moved, the benchmark calibration and the four robustness tests display very similar relative productivity, relative childrearing costs and relative above-quota penalties, and the same model implied urbanization rates, fertility rates, fraction of moving, population ratios and income ratios.

Next we move to the results of the policy experiments under the four different assumptions. As shown in Table ID.3, rural and urban per capita output, relative income ratios such as urban wage premium and skill premium all exhibit the same direction of movements as what is observed in the policy experiments of the benchmark calibration. The variable that we have mixed signs is the overall output per capita under the experiment of immediate reallocation of land. As being discussed in Section 5.4, this is a result from endogenous fractions of migration: If the migration discouragement effect of more workers staying in rural areas outweighs the direct effect of the increases in both the rural and urban per capita output, the overall output per capita will decrease, and vice versa. To sum up, our results are quite robust.

D.2 Sensitivity Tests for Parameters and Private-Sector Premium This subsection provides sensitivity tests for the three parameters α, σ, τ, and the target of positive private-sector premium ($\frac{w_P^H}{w_S}$). To examine the robustness of our main results, we vary these parameters or the target one at a time and re-calibrate the baseline model. Other parameters and targets remain unchanged. The results are reported in Table ID.4 and ID.5.

[Insert Table ID.4 about here]
[Insert Table ID.5 about here]

First of all, the private firm in urban areas employs a CES production technology with the share parameter α. We preset $\alpha = 0.5$ in the baseline calibration. Here we vary α to be 0.45 or 0.55 for the robustness tests. Second, σ governs the elasticity of substitution between high- and low-skilled workers. In the baseline calibration, we follow the literature to set $\sigma = 0.8333$ so that the elasticity of substitution between high- and low-skilled workers is equal to 6. We thus provide sensitivity tests by assuming the elasticity of substitution to be 4 or 8, so that $\sigma = 0.75$ or 0.875, respectively. Third, we follow Song et al. (2015) to set $\tau = 0.2$ in the baseline calibration.[66] Here we examine the scenarios of a higher tax rate ($\tau = 0.25$) or a lower tax rate ($\tau = 0.15$). Finally, the private-sector premium is about 13 percent in our baseline economy ($\frac{w_P^H}{w_S} = 1.1346$). Due to

[66] Song, Z., K. Storesletten, Y. Wang, and F. Zilibotti, 2015, "Sharing high growth across generations: pensions and demographic transition in China," *American Economic Journal: Macroeconomics*, 7(2), pp. 1-39.

the SOE reforms of introducing enterprise responsibility system and *xiagang* policy in China, we consider a scenario that what if our calibration target is a negative private-sector premium. That is, we average the periods of 1988-1991 and 2001-2007 where $\frac{w_P^H}{w_S}$ is smaller than one to obtain a new target $\frac{w_P^H}{w_S} = 0.8983$. Then, we re-solve A_P to match the new target with the assumption that A_S is fixed at 1.4854. Alternatively, we re-solve A_P and A_S together to match $\frac{w_P^H}{w_S} = 0.8983$ and $\frac{y_U}{y_R} = 1.9641$.

Not surprisingly, the related parameters change as the above three parameters or the private-sector premium vary. For example, varying α affects the calibrated values of A_P, A_S, and η. However, because we still match the targets of urban-rural income gap, skill premium, population ratios, and total fertility rates, a change in α has no effects on migration decisions, childrearing costs, and above-quota fines. Similarly, when the target of private-sector premium becomes 0.8983, the calibrated values of A_P, A_S, $\bar{\beta}$, migration utility costs, and above-quota fines change accordingly, but the magnitudes are limited. Therefore, we conclude that our main results remain valid, due to our calibration methodology that uses moments of the so-called great ratios rather than levels.

Appendix II: Theoretical Details

In this appendix, we provide detailed proofs for the baseline as well as the generalized models.

A. Mathematical Proofs for the Baseline Model

A.0 Condition 1 (Restriction on Skill Premia) Recall the definition: $\omega_2\left(\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}\right) = 1$ and the fact that ω_2 is increasing in $\frac{P^{H+}}{P^{L+}}$, we must have $w_S/w_P^H \leq 1$. To ensure $w_P^H/w_P^L > 1$, we reproduce (8) below:

$$\frac{w_P^H}{w_P^L} = \frac{\alpha}{1-\alpha}\eta^{\sigma}\left(\frac{P^{H+}}{P^{L+}}\right)^{-(1-\sigma)} \equiv \omega_1\left(\frac{P^{H+}}{P^{L+}}\right), \tag{47}$$

where $\omega_1(0) = \infty$ and $\omega_1'\left(\frac{P^{H+}}{P^{L+}}\right) < 0$. We first get

$$\omega_1\left(\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}\right) = \left[\alpha^{-\frac{1}{1-\sigma}}\left(\frac{A_S}{A_P}\right)^{\frac{\sigma}{1-\sigma}} - 1\right]^{\frac{1-\sigma}{\sigma}}\left(\frac{\alpha}{1-\alpha}\eta^{\sigma}\right)^{\frac{1}{\sigma}}$$

To assure that $w_P^H/w_P^L > 1$ for all $\frac{P^{H+}}{P^{L+}} \leq \left(\frac{P^{H+}}{P^{L+}}\right)_{\max}$, we impose the following sufficient condition:

$$\begin{aligned}
1 &< \omega_1\left(\left(\frac{P^{H+}}{P^{L+}}\right)_{\max}\right) \\
&= \left[\alpha^{-\frac{1}{1-\sigma}}\left(\frac{A_S}{A_P}\right)^{\frac{\sigma}{1-\sigma}} - 1\right]^{\frac{1-\sigma}{\sigma}}\left(\frac{\alpha}{1-\alpha}\eta^{\sigma}\right)^{\frac{1}{\sigma}} \\
&\Leftrightarrow \left[\alpha^{-\frac{1}{1-\sigma}}\left(\frac{A_S}{A_P}\right)^{\frac{\sigma}{1-\sigma}} - 1\right]^{\frac{1-\sigma}{\sigma}}\left(\frac{\alpha}{1-\alpha}\eta^{\sigma}\right)^{\frac{1}{\sigma}} > 1
\end{aligned}$$

which yields Condition 1.

A.1 Proof of Theorem 1 (Existence of MME) Consider the locus $\Delta V_H\left(\bar{\beta},\psi^H\right)=0$ where the fertility loss and the quality gain of migration balance off. Suppose we reduce the fertility loss by lowering β from $\bar{\beta}$ to $\underline{\beta}$. We then compare the expected fertility loss between $\Delta V_H\left(\bar{\beta},\psi^H\right)=0$ and $\Delta V_H\left(\underline{\beta},\psi^H\right)$:

$$\begin{aligned}
& \Delta V_H\left(\bar{\beta},\psi^H\right)-\Delta V_H\left(\underline{\beta},\psi^H\right) \\
= & (1-\varepsilon)\,\bar{\beta}\left\{\pi\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\pi)\left[\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}\right]-\left(n_R^*|_{\bar{\beta}}\right)^{\varepsilon}\right\} \\
& -(1-\varepsilon)\,\underline{\beta}\left\{\pi\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}+(1-\pi)\left[\rho\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\underline{\beta}}\right)^{\varepsilon}\right]-\left(n_R^*|_{\underline{\beta}}\right)^{\varepsilon}\right\} \\
\propto & \;\bar{\beta}\left\{\begin{array}{c}\pi\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}-\pi\left[\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}\right] \\ +\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}-\left(n_R^*|_{\bar{\beta}}\right)^{\varepsilon}\end{array}\right\} \\
& -\underline{\beta}\left\{\begin{array}{c}\pi\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}-\pi\left[\rho\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\underline{\beta}}\right)^{\varepsilon}\right] \\ +\rho\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\underline{\beta}}\right)^{\varepsilon}-\left(n_R^*|_{\underline{\beta}}\right)^{\varepsilon}\end{array}\right\} \\
< & \;0
\end{aligned}$$

because $\bar{\beta}>\underline{\beta}$ and

$$\frac{\partial}{\partial\beta}\left\{\begin{array}{c}\pi\left(n_U^*\right)^{\varepsilon}-\pi\left[\rho\left(n_U^*\right)^{\varepsilon}+(1-\rho)\left(n_M^*\right)^{\varepsilon}\right] \\ +\rho\left(n_U^*\right)^{\varepsilon}+(1-\rho)\left(n_I^*|\right)^{\varepsilon}-\left(n_R^*\right)^{\varepsilon}\end{array}\right\}<0,$$

by Proposition 3. This yields $\Delta V_H\left(\underline{\beta},\psi^H\right)>\Delta V_H\left(\bar{\beta},\psi^H\right)=0$ so that the locus $\Delta V_H\left(\underline{\beta},\psi^H\right)$ lies above $\Delta V_H\left(\bar{\beta},\psi^H\right)=0$. So $\Delta V_H\left(\underline{\beta},\psi^L\right)=0$ lies on the right of $\Delta V_H\left(\bar{\beta},\psi^H\right)=0$ as in Figure 6. Likewise for the location comparison for the loci $\Delta V_L\left(\underline{\beta},\psi^L\right)=0$ and $\Delta V_L\left(\bar{\beta},\psi^L\right)$ in Figure 6.

For the location of the loci $\Delta V_L\left(\underline{\beta},\psi^L\right)=0$ and $\Delta V_H\left(\bar{\beta},\psi^H\right)=0$ in Figure 6, we can explain as follows. Consider the loci $\Delta V_H\left(\bar{\beta},\psi^H\right)=0$ and $\Delta V_L\left(\underline{\beta},\psi^H\right)$. Comparing the expected fertility loss between $\Delta V_H\left(\bar{\beta},\psi^H\right)=0$ and $\Delta V_L\left(\underline{\beta},\psi^H\right)$:

$$\begin{aligned}
& \Delta V_H\left(\bar{\beta},\psi^H\right)-\Delta V_L\left(\underline{\beta},\psi^H\right) \\
= & (1-\varepsilon)\,\bar{\beta}\left\{\pi\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\pi)\left[\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}\right]-\left(n_R^*|_{\bar{\beta}}\right)^{\varepsilon}\right\} \\
& -(1-\varepsilon)\,\underline{\beta}\left[\rho\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\underline{\beta}}\right)^{\varepsilon}-\left(n_R^*|_{\underline{\beta}}\right)^{\varepsilon}\right] \\
\propto & \;\bar{\beta}\left\{\pi\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}-\pi\left[\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}\right]+\left[\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}\right]-\left(n_R^*|_{\bar{\beta}}\right)^{\varepsilon}\right\} \\
& -\underline{\beta}\left[\rho\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\underline{\beta}}\right)^{\varepsilon}-\left(n_R^*|_{\underline{\beta}}\right)^{\varepsilon}\right] \\
= & \;\bar{\beta}\underbrace{\left\{\pi\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}-\pi\left[\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}\right]\right\}}_{-} \\
& +\bar{\beta}\left[\rho\left(n_F^*|_{\bar{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\bar{\beta}}\right)^{\varepsilon}-\left(n_R^*|_{\bar{\beta}}\right)^{\varepsilon}\right]-\underline{\beta}\left[\rho\left(n_F^*|_{\underline{\beta}}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\underline{\beta}}\right)^{\varepsilon}-\left(n_R^*|_{\underline{\beta}}\right)^{\varepsilon}\right] \\
< & \;0,
\end{aligned}$$

because $\bar{\beta}>\underline{\beta}$ and

$$\frac{\partial}{\partial\beta}\left[\rho\left(n_F^*|_{\beta}\right)^{\varepsilon}+(1-\rho)\left(n_I^*|_{\beta}\right)^{\varepsilon}-\left(n_R^*|_{\beta}\right)^{\varepsilon}\right]<0,$$

by Proposition 3. So the former has a larger fertility loss than the latter. Next, we compare the expected locational quality gain:

$$\begin{aligned}
& (1-\theta)\,\theta\left\{\pi w_S + (1-\pi)\,w_P^H - [\pi + (1-\pi)\,\rho\mu]\,\tau - x\right\} + [\pi + (1-\pi)\,\rho\mu]\,B - \psi^H \\
& - \left\{(1-\theta)\,\theta\left[\rho\left(w_P^L - \mu\tau\right) + (1-\rho)w_P^L - x\right] + \rho\mu B - \psi^H\right\} \\
= \; & (1-\theta)\,\theta\left\{\left[\pi w_S + (1-\pi)\,w_P^H\right] - \left[\rho w_P^L + (1-\rho)w_P^L\right]\right\} \\
& - (1-\theta)\,\theta\pi\,(1-\rho\mu)\,\tau + \pi\,(1-\rho\mu)\,B \\
= \; & (1-\theta)\,\theta\left\{\left[\pi w_S + (1-\pi)\,w_P^H\right] - w_P^L\right\} + \pi\,(1-\rho\mu)\,[B - (1-\theta)\,\theta\tau] \\
> \; & 0.
\end{aligned}$$

So the former has a larger locational quality gain than the latter.

Given that $\Delta V_H\left(\bar{\beta}, \psi^H\right) = 0$ has both a larger fertility loss and a larger locational quality gain than $\Delta V_L\left(\underline{\beta}, \psi^H\right)$. Two possible cases emerge. Firstly, if $\Delta V_L\left(\underline{\beta}, \psi^H\right) < 0$, then the locus $\Delta V_L\left(\underline{\beta}, \psi^L\right) = 0$ lies below the locus $\Delta V_H\left(\bar{\beta}, \psi^H\right) = 0$. This is the situation shown in Figure 6 where $\psi^H > \psi^L$. Secondly, if $\Delta V_L\left(\underline{\beta}, \psi^H\right) > 0$, then the locus $\Delta V_L\left(\underline{\beta}, \psi^L\right) = 0$ lies above the locus $\Delta V_H\left(\bar{\beta}, \psi^H\right) = 0$ and we have $\psi^L > \psi^H$. In either case, the parametric space supporting a mixed migration equilibrium is dense and hence nonempty.■

A.2 Proof of Theorem 2 (Locational Quantity-Quality Trade-off) Writing out the indifference boundaries, we have:

$$\begin{aligned}
\Delta V_H(\beta,\psi) \;\equiv\; & \pi V^S(\beta) + (1-\pi)\,V^{P,H}(\beta) - \psi - V^R(\beta) \\
= \; & \pi\left\{(1-\theta)\,\theta\left[w_S - \tau - n_F^*|_\beta\left[\phi_U^0 + \left(1 - \frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right] + \bar{\beta}\,(n_F^*|_\beta)^\varepsilon + B\right\} \\
& + (1-\pi)\,\rho\left\{\begin{array}{c} (1-\theta)\,\theta\left[w_P^H - \mu\tau - n_F^*|_\beta\left[\phi_U^0 + \left(1 - \frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right] \\ +\bar{\beta}\,(n_F^*|_\beta)^\varepsilon + \mu B \end{array}\right\} \\
& + (1-\pi)\,(1-\rho)\left\{\begin{array}{c} (1-\theta)\,\theta\left[w_P^H - n_I^*|_\beta\left[\phi_U^0 + \left(1 - \frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right] \\ +\bar{\beta}\,(n_I^*|_\beta)^\varepsilon \end{array}\right\} \\
& - \left\{(1-\theta)\,\theta\left[x - n_R^*|_\beta\left[\phi_R^0 + \left(1 - \frac{\bar{n}_R}{n_R^*|_\beta}\right)\bar{\phi}_R\right]\right] + \bar{\beta}\,(n_R^*|_\beta)^\varepsilon\right\} - \psi.
\end{aligned}$$

Using the optimal fertility condition (18), we obtain:

$$\begin{aligned}
\Delta V_H(\beta,\psi) \;=\; & \pi\left[(1-\theta)\,\theta\,(w_S - \tau) + (1-\varepsilon)\,\beta\,(n_F^*|_\beta)^\varepsilon + B\right] \\
& + (1-\pi)\,\rho\left[(1-\theta)\,\theta\left(w_P^H - \mu\tau\right) + (1-\varepsilon)\,\bar{\beta}\,(n_F^*|_\beta)^\varepsilon + \mu B\right] \\
& + (1-\pi)\,(1-\rho)\left[(1-\theta)\,\theta w_P^H + (1-\varepsilon)\,\bar{\beta}\,(n_I^*|_\beta)^\varepsilon\right] \\
& - \left[(1-\theta)\,\theta x + (1-\varepsilon)\,\bar{\beta}\,(n_R^*|_\beta)^\varepsilon\right] - \psi \\
= \; & \underbrace{\begin{array}{c} (1-\theta)\,\theta\left\{\pi w_S + (1-\pi)\,w_P^H - [\pi + (1-\pi)\,\rho\mu]\,\tau - x\right\} \\ + [\pi + (1-\pi)\,\rho\mu]\,B - \psi \end{array}}_{\text{expected locational quality gain (+)}} \\
& + \underbrace{(1-\varepsilon)\,\beta\left\{\pi\,(n_F^*|_\beta)^\varepsilon + (1-\pi)\left[\rho\,(n_F^*|_\beta)^\varepsilon + (1-\rho)\,(n_I^*|_\beta)^\varepsilon\right] - (n_R^*|_\beta)^\varepsilon\right\}}_{\text{expected fertility loss (-)}}.
\end{aligned}$$

Likewise, we have:

$$
\begin{aligned}
\Delta V_L(\beta,\psi) &= V^{P,L}(\beta) - \psi - V^R(\beta) \\
&= \underbrace{(1-\theta)\,\theta\left[\rho\left(w_P^L - \mu\tau\right) + (1-\rho)w_P^L - x\right] + \rho\mu B - \psi}_{\text{expected locational quality gain (+)}} \\
&\quad + \underbrace{(1-\varepsilon)\,\beta\left[\rho\left(n_F^*|_\beta\right)^\varepsilon + (1-\rho)\left(n_I^*|_\beta\right)^\varepsilon - \left(n_R^*|_\beta\right)^\varepsilon\right]}_{\text{expected fertility loss (-)}}.
\end{aligned}
$$

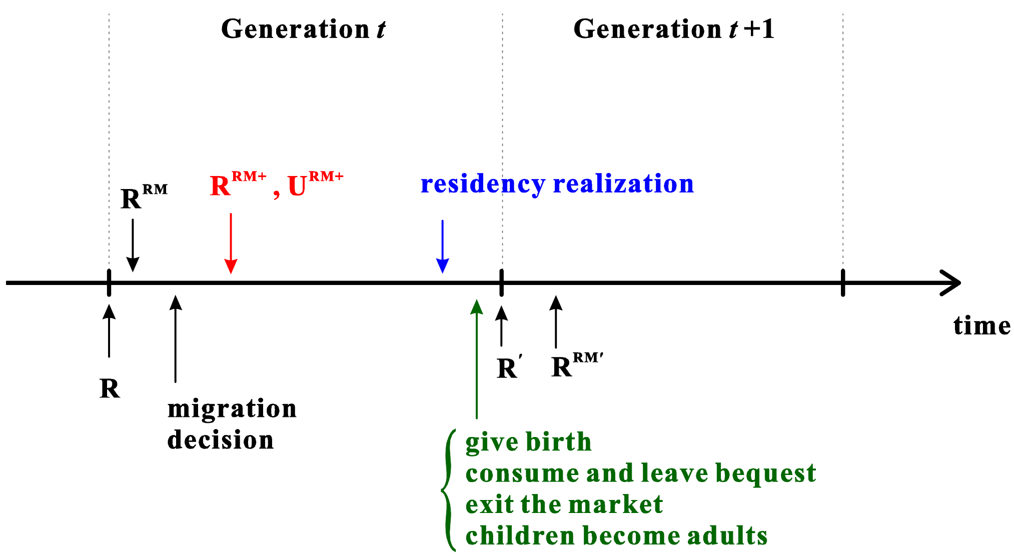
So the indifference boundaries can be decomposed into two terms: an expected locational quality gain versus an expected fertility loss.

For the expected locational quality gain, it is straightforward to show that it is positive. Given that urban amenities B are non-rival and non-excludable, whereas τ is the individual embarked tax, we get $B > \tau$. Also, since migration cost is a fraction of wage incomes, so the overall locational gain must be positive. For the expected fertility loss, Proposition 3 implies that it is negative. As a result, we obtain a locational quantity-quality trade-off for the migration decision.■

B. Mathematical Proofs for the Generalized Model

B.1 Reverse Migration Consider the urban residency shock that leads to relocate a constant fraction of newborns of generation-t parents from urban to rural occurring at the beginning of $t+1$. Figure IIB.1 depicts the timeline with reverse migration.

Figure IIB.1: Timeline for Reverse Migration



Denote the fraction of reverse migration as Λ. The beginning-of-the-period population identity equations in urban and rural areas are $U^{RM} = (1-\Lambda)\,U = (1-\Lambda)\left(S + P^H + P^L\right)$ and $R^{RM} = R + \Lambda U = H + L + \Lambda\left(S + P^H + P^L\right)$ as given in the main text. The actual-working populations in urban areas under the equilibrium we examine are:

$$
\begin{aligned}
S^{RM+} &= (1-\Lambda)\,S + \pi\left[\zeta + (1-\zeta)\,\Gamma_H\right]\left[H + \Lambda\left(S + P^H\right)\right] \\
&= S^+ - \Lambda\left\{S - \pi\left[\zeta + (1-\zeta)\,\Gamma_H\right]\left(S + P^H\right)\right\}
\end{aligned}
\tag{48}
$$

$$
\begin{aligned}
P^{RM,H+} &= (1-\Lambda) P^H + (1-\pi)\left[\zeta + (1-\zeta)\Gamma_H\right]\left[H + \Lambda\left(S + P^H\right)\right] \\
&= P^{H+} - \Lambda\left\{P^H - (1-\pi)\left[\zeta + (1-\zeta)\Gamma_H\right]\left(S + P^H\right)\right\}
\end{aligned} \tag{49}
$$

$$
\begin{aligned}
P^{RM,L+} &= (1-\Lambda) P^L + \zeta\Gamma_L\left(L + \Lambda P^L\right) \\
&= P^{L+} - \Lambda P^L\left(1 - \zeta\Gamma_L\right).
\end{aligned} \tag{50}
$$

Combining (48) and (49), we have

$$
\begin{aligned}
S^{RM+} + P^{RM,H+} &= S^+ - \Lambda\left\{S - \pi\left[\zeta + (1-\zeta)\Gamma_H\right]\left(S + P^H\right)\right\} \\
&\quad + P^{H+} - \Lambda\left\{P^H - (1-\pi)\left[\zeta + (1-\zeta)\Gamma_H\right]\left(S + P^H\right)\right\} \\
&= S^+ + P^{H+} - \Lambda\left\{1 - \left[\zeta + (1-\zeta)\Gamma_H\right]\right\}\left(S + P^H\right)
\end{aligned}
$$

so that reverse migration reduces the actual-working populations of both skill types (H and L) in urban areas. Total number of agents working in urban areas after migration inflows is thus equal to:

$$
\begin{aligned}
U^{RM+} &= U^{RM} + \left[\zeta + (1-\zeta)\Gamma_H\right]\left[H + \Lambda\left(S + P^H\right)\right] + \zeta\Gamma_L\left(L + \Lambda P^L\right) \\
&= U^+ - \Lambda\left[(1-\zeta)(1-\Gamma_H)\left(S + P^H\right) + (1-\zeta\Gamma_L) P^L\right]
\end{aligned} \tag{51}
$$

In a similar manner, we write down the actual-working populations in rural areas after migration outflows as:

$$
\begin{aligned}
R^{RM+} &= R^{RM} - \left[\zeta + (1-\zeta)\Gamma_H\right]\left[H + \Lambda\left(S + P^H\right)\right] - \zeta\Gamma_L\left(L + \Lambda P^L\right) \\
&= R^+ + \Lambda\left[(1-\zeta)(1-\Gamma_H)\left(S + P^H\right) + (1-\zeta\Gamma_L) P^L\right]
\end{aligned} \tag{52}
$$

As children inherit the skill levels (i.e., jobs) and residency status directly from parents and all agents live for one period, the evolutions of workers in the SOE and the private sectors are:

$$
S^{RM\prime} = \left[\zeta n_F^*|_{\underline{\beta}} + (1\text{-}\zeta)\, n_F^*|_{\bar{\beta}}\right](1\text{-}\Lambda)\, S + \left[\zeta n_F^*|_{\underline{\beta}} + (1\text{-}\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right]\pi\left[H + \Lambda\left(S + P^H\right)\right], \tag{53}
$$

$$
\begin{aligned}
P^{RM,H\prime} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1\text{-}\zeta)\, n_F^*|_{\bar{\beta}}\right](1\text{-}\Lambda)\, P^H \\
&\quad + \left[\zeta n_F^*|_{\underline{\beta}} + (1\text{-}\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right](1\text{-}\pi)\,\rho\left[H + \Lambda\left(S + P^H\right)\right],
\end{aligned} \tag{54}
$$

$$
P^{RM,L\prime} = \left[\zeta n_F^*|_{\underline{\beta}} + (1\text{-}\zeta)\, n_F^*|_{\bar{\beta}}\right](1\text{-}\Lambda)\, P^L + \zeta\Gamma_L n_F^*|_{\underline{\beta}}\,\rho\left(L + \Lambda P^L\right), \tag{55}
$$

where Z' denotes the next period value of Z. The evolution equation for U^{RM} can be written accordingly follows:

$$
\begin{aligned}
U^{RM\prime} &= S^{RM\prime} + P^{RM,H\prime} + P^{RM,L\prime} \\
&= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\, n_F^*|_{\bar{\beta}}\right](1-\Lambda)\left(S + P^H + P^L\right) \\
&\quad + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\,\Gamma_H n_F^*|_{\bar{\beta}}\right]\left[\pi + (1-\pi)\,\rho\right]\left[H + \Lambda\left(S + P^H\right)\right] + \zeta\Gamma_L n_F^*|_{\underline{\beta}}\,\rho\left(L + \Lambda P^L\right)
\end{aligned} \tag{56}
$$

The evolution equations for rural high- and low-skilled workers and total rural workers with rural residency can be written accordingly:

$$
H^{RM\prime} = \left[H + \Lambda\left(S + P^H\right)\right]\left\{(1\text{-}\rho)(1\text{-}\pi)\left[\zeta n_I^*|_{\underline{\beta}} + (1\text{-}\zeta)\,\Gamma_H n_I^*|_{\bar{\beta}}\right] + (1\text{-}\zeta)(1\text{-}\Gamma_H)\, n_R^*|_{\bar{\beta}}\right\}, \tag{57}
$$

$$
L^{RM\prime} = \left(L + \Lambda P^L\right)\left\{(1\text{-}\rho)\,\zeta\Gamma_L n_I^*|_{\underline{\beta}} + \zeta\,(1\text{-}\Gamma_L)\, n_R^*|_{\underline{\beta}} + (1\text{-}\zeta)\, n_R^*|_{\bar{\beta}}\right\}, \tag{58}
$$

$$
R^{RM\prime} = H^{RM\prime} + L^{RM\prime}. \tag{59}
$$

Along a BGP, we get the ratios reported in the text using the fact that

$$\frac{H+\Lambda\left(S+P^{H}\right)}{S}=\frac{H}{S}+\Lambda\left(1+\frac{P^{H}}{S}\right)=\Lambda+\left(1+\Lambda\frac{P^{H}}{H}\right)\frac{H}{S}$$

B.2 Left-Behind Children Consider the case of left-behind children by migrant workers. The timeline is the same as that in the baseline economy.

Value Function of Staying in Rural Areas Recall the property associated with Leontief preferences that the optimal number of children n_R^{LBC*} is independent of agents' income and wealth, so both Propositions 1 and 2 for rural agents remains unchanged. As a result, we reproduce the value function $V^{LBC,R}(\beta)$ for convenience:

$$V^{LBC,R}(\beta)=(1-\theta)\,\theta\left\{x-n_R^{LBC*}|_\beta\left[\phi_R^0+I^\beta\left(1-\frac{\bar{n}_R}{n_R^{LBC*}|_\beta}\right)\bar{\phi}_R\right]\right\}+\beta\left(n_R^{LBC*}|_\beta\right)^\varepsilon \tag{60}$$

where

$$n_R^{LBC*}|_\beta=\left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\phi_R^0+I^\beta\bar{\phi}_R\right]}\right]^{\frac{1}{1-\varepsilon}}=n_R^*|_\beta.$$

So we have

$$V^{LBC,R}(\beta)=V^R(\beta).$$

That is, rural stayers' decisions remain unaffected.

Value Functions of Urban Workers Because SOE workers are granted with urban hukou immediately, their value functions remain the same. The main effects of left-behind children fall on the value functions of urban workers in the private sector.

1. High-skilled private sector workers

For high-skilled rural migrants working in the private sector, after staying in urban areas for $(1-\mu)$ of their lifetime, they obtain urban residency ($\rho^{LBC,H}=1$). The value function of a high-skilled migrant worker in the private sector is:

$$V^{LBC,P,H}(\beta)=\left\{\begin{array}{l}\max\limits_{c,b,n}u^U(c,b,n;\beta)\,|_{I^F=0,I^T=1}\\ s.t.\ c+nb+n\phi_U^0+\max\{n-\bar{n}_U,0\}\,\bar{\phi}_U=w_P^H-\mu\tau\end{array}\right\}$$

and we can solve:

$$n_F^{LBC*}|_\beta=\left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\phi_U^0+I^\beta\bar{\phi}_U\right]}\right]^{\frac{1}{1-\varepsilon}}=n_F^*|_\beta$$

Recall $n_R^*|_\beta$ from (19) and that $\phi_U^0>\phi_R^0$ and $\bar{\phi}_U\geq\bar{\phi}_R$. It is straightforward to show that the urban fertility choices are always below the rural counterpart, with formal urban workers' fertility lower than informal urban worker's, as stated in Proposition 5a.

Substituting the solutions of the maximization problem into $V^{LBC,P,H}(\beta)$ and noting that $n_F^{LBC*}|_\beta=n_F^*|_\beta$, we have

$$V^{LBC,P,H}(\beta)=(1-\theta)\,\theta\left\{w_P^H-\mu\tau-n_F^*|_\beta\left[\phi_U^0+I^\beta\left(1-\frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\}+\beta\left(n_F^*|_\beta\right)^\varepsilon+\mu B,$$

which can be compared with $V^{P,H}(\beta)$ to yield the expression of $V^{LBC,P,H}(\beta) - V^{P,H}(\beta)$.

2. Low-skilled workers

By applying the same procedure for deriving $V^{LBC,P,H}(\beta)$, together with $\rho^{LBC,L} = 0$ and $\Xi > 1$, the value function for low-skilled migrant workers in urban private sector becomes:

$$V^{LBC,P,L}(\beta) = (1-\theta)\theta\left\{w_P^L - n_I^{LBC*}|_\beta\left[\Xi\phi_R^0 + I^\beta\left(1 - \frac{\bar{n}_R}{n_I^{LBC*}|_\beta}\right)\bar{\phi}_R\right]\right\} + \beta\left(n_I^{LBC*}|_\beta\right)^\varepsilon$$

Because all urban low-skilled workers are without urban residency (i.e., they are informal private-sector workers) and their children are all left behind, their fertility choices are

$$n_I^{LBC*}|_\beta = \left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\Xi\phi_R^0 + I^\beta\bar{\phi}_R\right]}\right]^{\frac{1}{1-\varepsilon}}$$

Proposition 5b follows immediately.

Recalling $V^{P,L}(\beta)$ in the baseline model, we conclude:

$$\begin{aligned}
& V^{LBC,P,L}(\beta) - V^{P,L}(\beta) \\
= \; & (1-\theta)\theta\left\{w_P^L - n_I^{LBC*}|_\beta\left[\Xi\phi_R^0 + I^\beta\left(1 - \frac{\bar{n}_R}{n_I^{LBC*}|_\beta}\right)\bar{\phi}_R\right]\right\} + \beta\left(n_I^{LBC*}|_\beta\right)^\varepsilon \\
& -\rho\left\{(1-\theta)\theta\left\{w_P^L - \mu\tau - n_F^*|_\beta\left[\phi_U^0 + I^\beta\left(1 - \frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\} + \beta\left(n_F^*|_\beta\right)^\varepsilon + \mu B\right\} \\
& -(1-\rho)\left\{(1-\theta)\theta\left\{w_P^L - n_I^*|_\beta\left[\phi_U^0 + I^\beta\left(1 - \frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right\} + \beta\left(n_I^*|_\beta\right)^\varepsilon\right\},
\end{aligned}$$

which leads to the expression in the main text.

Migration Decisions From the above value functions under left-behind children, we obtain Proposition 5c. From Proposition 5c, the indifference boundaries and migration conditions are as derived in the main text, which lead to Proposition 5d.

Evolution of Workers The beginning-of-the-period population identity equations in urban and rural areas are the same as in the benchmark case:

$$\begin{aligned}
U^{LBC} &= U = S + P^H + P^L, && (61) \\
R^{LBC} &= R = H + L. && (62)
\end{aligned}$$

The actual-working populations in urban areas under the equilibrium we examine are:

$$S^{LBC+} = S + S_F = S + \left[\zeta + (1-\zeta)\Gamma_H\right]\pi H \quad (63)$$

$$P^{LBC,H+} = P^H + P_F^H = P^H + \left[\zeta + (1-\zeta)\Gamma_H\right](1-\pi)H \quad (64)$$

$$P^{LBC,L+} = P^L + P_I^L = P^L + \zeta\Gamma_L L \quad (65)$$

Total number of agents working in urban areas after migration inflows is thus equal to:

$$\begin{aligned}
U^{LBC+} &= S^{LBC+} + P^{LBC,H+} + P^{LBC,L+} && (66) \\
&= S + P^H + P^L + \left[\zeta + (1-\zeta)\Gamma_H\right]H + \zeta\Gamma_L L \\
&= U + \left[\zeta + (1-\zeta)\Gamma_H\right]H + \zeta\Gamma_L L.
\end{aligned}$$

In a similar manner, we write down the actual-working populations in rural areas after migration outflows as:

$$R^{LBC+} = (1-\zeta)(1-\Gamma_H)H + (1-\zeta\Gamma_L)L. \tag{67}$$

As children inherit the skill levels (i.e., jobs) and residency status directly from parents and all agents live for one period, the evolutions of workers in the SOE and the private sectors are:

$$\begin{aligned} S^{LBC\prime} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta) n_F^*|_{\bar{\beta}}\right] S + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\Gamma_H n_F^*|_{\bar{\beta}}\right] \pi H, & (68) \\ P^{LBC,H'} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta) n_F^*|_{\bar{\beta}}\right] P^H + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\Gamma_H n_F^*|_{\bar{\beta}}\right] (1-\pi) H, & (69) \\ P^{LBC,L'} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta) n_F^*|_{\bar{\beta}}\right] P^L. & (70) \end{aligned}$$

The evolution equation for U can be written accordingly follows:

$$\begin{aligned} U^{LBC\prime} &= S^{LBC\prime} + P^{LBC,H'} + P^{LBC,L'} & (71) \\ &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta) n_F^*|_{\bar{\beta}}\right] \left(S + P^H\right) \\ &\quad + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\Gamma_H n_F^*|_{\bar{\beta}}\right] H + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta) n_F^*|_{\bar{\beta}}\right] P^L. \end{aligned}$$

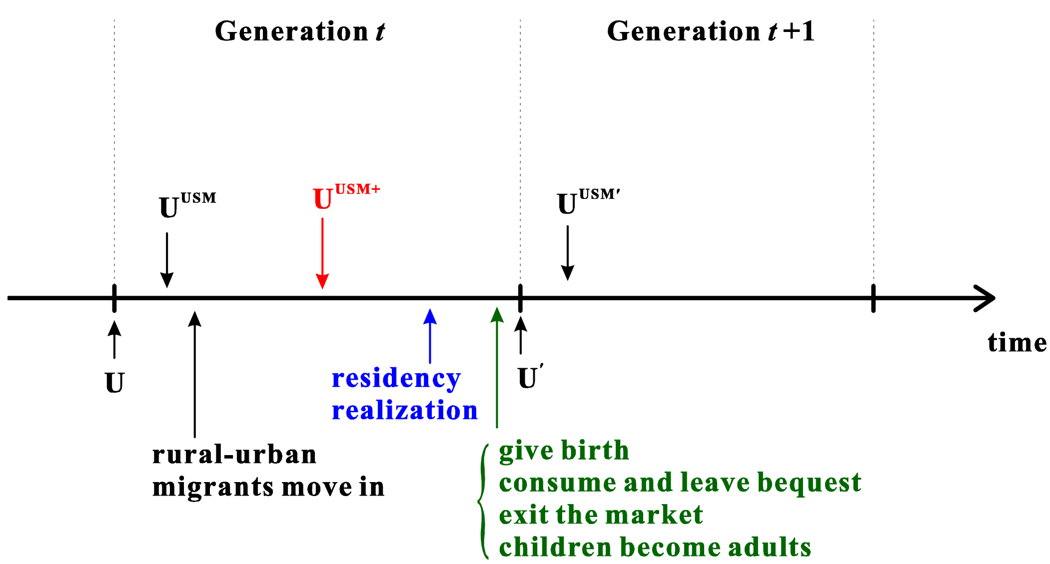
The evolution equations for rural high- and low-skilled workers and total rural workers with rural residency can be written accordingly:

$$\begin{aligned} H^{LBC\prime} &= H(1-\zeta)(1-\Gamma_H) n_R^*|_{\bar{\beta}}, & (72) \\ L^{LBC\prime} &= L\left\{\zeta\Gamma_L n_I^{LBC*}|_{\underline{\beta}} + \left[\zeta(1-\Gamma_L) + (1-\zeta)\right] n_R^*|_{\bar{\beta}}\right\}, & (73) \\ R^{LBC\prime} &= H^{LBC\prime} + L^{LBC\prime}. & (74) \end{aligned}$$

These enables us to derive the a BGP population ratios.

B.3 Upward Skill Mobility We now consider upward skill mobility in the sense that children of high-skilled parents brought up to urban are high-skilled but those of urban low-skilled parents have a probability ϑ to become high-skilled. Figure IIB.2 depicts the timeline with upward skill mobility.

Figure IIB.2: Timeline for Upward Skill Mobility



Evolution of Workers The actual-working populations in urban areas under the equilibrium we examine are:

$$S^{USM+} = S + S_F = S + \left[\zeta + (1-\zeta)\Gamma_H\right]\pi H \tag{75}$$

$$P^{USM,H+} = \underbrace{P^H + \vartheta P^L}_{P^{USM,H}} + \underbrace{\left[\zeta + (1-\zeta)\Gamma_H\right](1-\pi)H}_{P_F^H + P_I^H} \tag{76}$$

$$P^{USM,L+} = \underbrace{(1-\vartheta)P^L}_{P^{USM,L}} + \underbrace{\zeta \Gamma_L L}_{P_F^L + P_I^L} \tag{77}$$

Total number of agents working in urban areas after migration inflows is thus equal to:

$$\begin{aligned} U^{USM+} &= S + P^H + P^L + \left[\zeta + (1-\zeta)\Gamma_H\right]H + \zeta\Gamma_L L \\ &= U + \left[\zeta + (1-\zeta)\Gamma_H\right]H + \zeta\Gamma_L L. \end{aligned} \tag{78}$$

The actual-working populations in rural areas after migration outflows are

$$R^{USM+} = H^+ + L^+ = (1-\zeta)(1-\Gamma_H)H + (1-\zeta\Gamma_L)L. \tag{79}$$

As children inherit the skill levels (i.e., jobs) and residency status directly from parents and all agents live for one period, the evolutions of workers in the SOE and the private sectors are:

$$S^{USM\prime} = \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)n_F^*|_{\bar{\beta}}\right]S + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\Gamma_H n_F^*|_{\bar{\beta}}\right]\pi H \tag{80}$$

$$\begin{aligned} P^{USM,H\prime} &= \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)n_F^*|_{\bar{\beta}}\right]P^H + \left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)\Gamma_H n_F^*|_{\bar{\beta}}\right](1-\pi)\rho H \\ &\quad + \vartheta\left\{\left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)n_F^*|_{\bar{\beta}}\right]P^L + \zeta\Gamma_L n_F^*|_{\underline{\beta}}\rho L\right\} \end{aligned} \tag{81}$$

$$P^{USM,L\prime} = (1-\vartheta)\left\{\left[\zeta n_F^*|_{\underline{\beta}} + (1-\zeta)n_F^*|_{\bar{\beta}}\right]P^L + \zeta\Gamma_L n_F^*|_{\underline{\beta}}\rho L\right\} \tag{82}$$

The evolution equations for rural high- and low-skilled workers and total rural workers with rural residency can be written accordingly:

$$H^{USM\prime} = H\left\{(1-\rho)(1-\pi)\left[\zeta n_I^*|_{\underline{\beta}} + (1-\zeta)\Gamma_H n_I^*|_{\bar{\beta}}\right] + (1-\zeta)(1-\Gamma_H)n_R^*|_{\bar{\beta}}\right\}, \tag{83}$$

$$L^{USM\prime} = L\left\{(1-\rho)\zeta\Gamma_L n_I^*|_{\underline{\beta}} + \zeta(1-\Gamma_L)n_R^*|_{\underline{\beta}} + (1-\zeta)n_R^*|_{\bar{\beta}}\right\}, \tag{84}$$

$$R^{USM\prime} = H^{USM\prime} + L^{USM\prime}. \tag{85}$$

These imply the BGP population ratios in the main text.

The General-Equilibrium Wage Effect Although the common growth property remains valid under upward skill mobility, there is a general equilibrium wage effect so that the indifference boundaries of migration are affected.

Since $P^{USM,H+} = P^{H+} + \vartheta P^L$ and $P^{USM,L+} = P^{L+} - \vartheta P^L$, it is straightforward to see there is a negative general equilibrium labor endowment effect on high-skilled wage and a positive effect on low-skilled wage:

$$\begin{aligned} w_P^{USM,H} &= \alpha\eta^\sigma A_P\left[\alpha\eta^\sigma + (1-\alpha)\left(\frac{P^{L+} - \vartheta P^L}{P^{H+} + \vartheta P^L}\right)^\sigma\right]^{\frac{1-\sigma}{\sigma}} < w_P^H, \\ w_P^{USM,L} &= (1-\alpha)A_P\left[\alpha\left(\eta\frac{P^{H+} + \vartheta P^L}{P^{L+} - \vartheta P^L}\right)^\sigma + (1-\alpha)\right]^{\frac{1-\sigma}{\sigma}} > w_P^L \end{aligned}$$

Because of changes in wages, there is a dynamic migration effect encouraging more low-skilled but less high-skilled to migrate. Given the Leontief preferences, the fertility choices are not affected under upward skill mobility. However, the value functions of the urban private-sector workers are altered due to the changes in wages. Specifically, we have

$$\begin{aligned} V^{USM,P,H}(\beta) - V^{P,H}(\beta) &= (1-\theta)\theta\left(w_P^{USM,H} - w_P^H\right) < 0 \\ V^{USM,P,L}(\beta) - V^{P,L}(\beta) &= (1-\theta)\theta\left(w_P^{USM,L} - w_P^L\right) > 0 \end{aligned}$$

which leads to Proposition 6a. As a result, the indifference boundaries of the urban private-sector workers are affected as given in the main text, which lead to Proposition 6b.

B.4 Directed Urban Benefits To the end, we consider directed urban benefits. The timeline is unchanged and the same as Figure 3. An urban worker's optimization is modified as follows. The utility function is:

$$\begin{aligned} u^{DB,U}(c,b,n;\beta)|_{I^F,I^T} = &\min\left\{\theta c, (1-\theta)nb\left[1+\left(I^F+\left(1-I^F\right)I^T\mu\right)B_E\right]\right\} \\ &+\beta n^{\varepsilon} + \left[I^F+\left(1-I^F\right)I^T\mu\right]B_A. \end{aligned} \tag{86}$$

The budget constraint is:

$$\begin{aligned} &c+nb+n\left[\phi_U^0 - \left(I^F+\left(1-I^F\right)I^T\mu\right)B_C\right] + \left[I^F+\left(1-I^F\right)I^T\right]\max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U \\ &+\left\{1-\left[I^F+\left(1-I^F\right)I^T\right]\right\}\max\left\{n-\bar{n}_R,0\right\}\bar{\phi}_R \\ = &\ w+I^H\left[I^F+\left(1-I^F\right)I^T\mu\right]B_W - \left[I^F+\left(1-I^F\right)I^T\mu\right]\tau, \end{aligned} \tag{87}$$

where $I^H = 1$ for high-skilled workers (0 otherwise) and $w \in \{w_S, w_P^H, w_P^L\}$. All urban workers face the same childrearing costs regardless of their residency status, albeit the benefits and obligations are associated with one's residency.

So an urban worker with urban residency ($\left(I^F=1, I^T=0\right)$ or $\left(I^F=0, I^T=1\right)$) faces budget constraints that can be rewritten as follows:

$$\begin{aligned} S &: \ c+nb+n\left(\phi_U^0 - B_C\right) + \max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U = w_S + B_W - \tau \\ P^H &: \ c+nb+n\left(\phi_U^0 - \mu B_C\right) + \max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U = w_P^H + \mu B_W - \mu\tau \\ P^L &: \ c+nb+n\left(\phi_U^0 - \mu B_C\right) + \max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U = w_P^L - \mu\tau \end{aligned}$$

On the other hand, the budget constraint of an urban worker with rural residency ($\left(I^F, I^T\right) = (0,0)$) is:

$$c+nb+n\phi_U^0 + \max\left\{n-\bar{n}_R,0\right\}\bar{\phi}_R = w_P^i, i=H,L.$$

SOE Workers For SOE workers $\left(I^F=1, I^T=0\right)$, the value function is:

$$V^{DB,S}(\beta) = \max_{c,b,n}\left\{\min\left[\theta c, (1-\theta)nb(1+B_E)\right] + \beta n^{\varepsilon} + B_A\right\}$$

$$s.t.\ \ c+nb+n\left(\phi_U^0 - B_C\right) + \max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U = w_S + B_W - \tau.$$

which is equivalent to

$$\begin{aligned} V^{DB,S}(\beta) = &\max_{b,n}\ (1-\theta)nb(1+B_E) + \beta n^{\varepsilon} + B_A \\ s.t.\ \ &\frac{1}{\theta}\left[(1-\theta)nb(1+B_E)\right] + nb + n\left(\phi_U^0 - B_C\right) + \max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U = w_S + B_W - \tau. \end{aligned}$$

Denote λ as the Lagrangian multiplier associated with the budget constraint. The first-order conditions for n and b are:

$$\begin{aligned}(1-\theta)\, b\,(1+B_E)+\varepsilon\beta n^{\varepsilon-1} &= \lambda\left[\frac{1}{\theta}(1-\theta)\, b\,(1+B_E)+b+\phi_U^0-B_C+I^\beta\bar{\phi}_U\right],\\ (1-\theta)\, n\,(1+B_E) &= \lambda n\left[\frac{1}{\theta}(1-\theta)(1+B_E)+1\right]\end{aligned}$$

and together yield

$$\frac{\varepsilon\beta}{n^{1-\varepsilon}}=\frac{\theta(1-\theta)(1+B_E)}{\theta+(1-\theta)(1+B_E)}\left(\phi_U^0-B_C+I^\beta\bar{\phi}_U\right).$$

Rearranging to get

$$n_F^{DB,S*}|_\beta=\left[\frac{1+(1-\theta)\,B_E}{1+B_E}\frac{\varepsilon\beta}{\theta(1-\theta)\left(\phi_U^0-B_C+I^\beta\bar{\phi}_U\right)}\right]^{\frac{1}{1-\varepsilon}}.$$

As a result, we have the ranking of fertility choice stated in the main text. Finally, we have

$$\left[1+\frac{(1-\theta)(1+B_E)}{\theta}\right]nb=w_S+B_W-\tau-n\left[\left(\phi_U^0-B_C\right)+I^\beta(n-\bar{n}_U)\,\bar{\phi}_U\right]$$

By substituting in the number of children chosen and the investment in children quality, we obtain a SOE worker's value function:

$$\begin{aligned}V^{DB,S}(\beta) &= \theta(1-\theta)\frac{1+B_E}{1+(1-\theta)\,B_E}\left\{w_S+B_W-\tau-n_F^{DB,S*}|_\beta\left[\phi_U^0-B_C+I^\beta\left(1-\frac{\bar{n}_U}{n_F^{DB,S*}|_\beta}\right)\bar{\phi}_U\right]\right\}\\ &\quad+\beta\left(n_F^{DB,S*}|_\beta\right)^\varepsilon+B_A\end{aligned}$$

In addition, we obtain the expression of $V^{DB,S}(\beta)-V^S(\beta)$.

Private-Sector High-Skilled Workers For private-sector high-skilled workers with urban residency $\left(I^F=0, I^T=1\right)$, the value function is:

$$\begin{aligned}V_F^{DB,P,H}(\beta) &= \max_{c,b,n}\left\{\min\left[\theta c,(1-\theta)\, nb\,(1+\mu B_E)\right]+\beta n^\varepsilon+\mu B_A\right\}\\ &\quad s.t.\;\; c+nb+n\left(\phi_U^0-\mu B_C\right)+\max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U=w_P^H+\mu B_W-\mu\tau\end{aligned}$$

which is equivalent to

$$\begin{aligned}V_F^{DB,P,H}(\beta) &= \max_{b,n}\;\;(1-\theta)\, nb\,(1+\mu B_E)+\beta n^\varepsilon+\mu B_A\\ s.t.\;\; &\frac{1}{\theta}\left[(1-\theta)\, nb\,(1+\mu B_E)\right]+nb+n\left(\phi_U^0-\mu B_C\right)+\max\left\{n-\bar{n}_U,0\right\}\bar{\phi}_U=w_P^H+\mu B_W-\mu\tau.\end{aligned}$$

Following the same procedure as in the case of SOE workers, we get

$$n_F^{DB,P,H*}|_\beta=\left[\frac{1+(1-\theta)\,\mu B_E}{1+\mu B_E}\frac{\varepsilon\beta}{\theta(1-\theta)\left(\phi_U^0-\mu B_C+I^\beta\bar{\phi}_U\right)}\right]^{\frac{1}{1-\varepsilon}}.$$

For private-sector high-skilled workers with rural residency $\left(I^F = 0, I^T = 0\right)$, the value function is identical to the benchmark case:

$$V_I^{DB,P,H}(\beta) = (1-\theta)\,\theta\left\{w_P^H - n_I^*|_\beta\left[\phi_U^0 + I^\beta\left(1-\frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right\} + \beta\left(n_I^*|_\beta\right)^\varepsilon,$$

where

$$n_I^*|_\beta = \left[\frac{\varepsilon\beta}{\theta(1-\theta)\left[\phi_U^0 + I^\beta\bar{\phi}_R\right]}\right]^{\frac{1}{1-\varepsilon}}.$$

These give Proposition 7a.

Putting together all the private-sector high-skilled workers, we have

$$\begin{aligned}
V^{DB,P,H}(\beta) &= \rho V_F^{DB,P,H}(\beta) + (1-\rho)V_I^{DB,P,H}(\beta)\\
&= \rho\theta(1-\theta)\frac{1+\mu B_E}{1+(1-\theta)\mu B_E}\cdot\\
&\quad\left\{w_P^H + \mu B_W - \mu\tau - n_F^{DB,P,H*}|_\beta\left[\phi_U^0 - \mu B_C + I^\beta\left(1-\frac{\bar{n}_U}{n_F^{DB,P,H*}|_\beta}\right)\bar{\phi}_U\right]\right\}\\
&\quad+\rho\left[\beta\left(n_F^{DB,P,H*}|_\beta\right)^\varepsilon + \mu B_A\right]\\
&\quad+(1-\rho)\left\{(1-\theta)\,\theta\left\{w_P^H - n_I^*|_\beta\left[\phi_U^0 + I^\beta\left(1-\frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right\} + \beta\left(n_I^*|_\beta\right)^\varepsilon\right\}.
\end{aligned}$$

Recall $V^{P,H}(\beta)$ in the benchmark case:

$$\begin{aligned}
V^{P,H}(\beta) &= \rho\left\{(1-\theta)\,\theta\left\{w_P^H - \mu\tau - n_F^*|_\beta\left[\phi_U^0 + I^\beta\left(1-\frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\} + \beta\left(n_F^*|_\beta\right)^\varepsilon + \mu B_A\right\}\\
&\quad+(1-\rho)\left\{(1-\theta)\,\theta\left\{w_P^H - n_I^*|_\beta\left[\phi_U^0 + I^\beta\left(1-\frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right\} + \beta\left(n_I^*|_\beta\right)^\varepsilon\right\}.
\end{aligned}$$

Then we get the expression of $V^{DB,P,H}(\beta) - V^{P,H}(\beta)$.

Private-Sector Low-Skilled Workers For private-sector low-skilled workers with urban residency $\left(I^F = 0, I^T = 1\right)$, the value function is:

$$\begin{aligned}
V_F^{DB,P,L}(\beta) &= \max_{c,b,n}\left\{\min\left[\theta c, (1-\theta)\,nb\,(1+\mu B_E)\right] + \beta n^\varepsilon + \mu B_A\right\}\\
&\quad s.t.\;\; c + nb + n\left(\phi_U^0 - \mu B_C\right) + \max\left\{n - \bar{n}_U, 0\right\}\bar{\phi}_U = w_P^L - \mu\tau
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
V_F^{DB,P,L}(\beta) &= \max_{b,n}\;\; (1-\theta)\,nb\,(1+\mu B_E) + \beta n^\varepsilon + \mu B_A\\
s.t.&\quad \frac{1}{\theta}\left[(1-\theta)\,nb\,(1+\mu B_E)\right] + nb + n\left(\phi_U^0 - \mu B_C\right) + \max\left\{n - \bar{n}_U, 0\right\}\bar{\phi}_U = w_P^L - \mu\tau.
\end{aligned}$$

Following the same procedure, we get

$$n_F^{DB,P,L*}|_\beta = \left[\frac{1+(1-\theta)\,\mu B_E}{1+\mu B_E}\frac{\varepsilon\beta}{\theta(1-\theta)\left(\phi_U^0 - \mu B_C + I^\beta\bar{\phi}_U\right)}\right]^{\frac{1}{1-\varepsilon}} = n_F^{DB,P,H*}|_\beta.$$

For private-sector low-skilled workers with rural residency $\left(I^F=0, I^T=0\right)$, the value function is identical to the benchmark case:

$$V_I^{DB,P,L}(\beta)=(1-\theta)\,\theta\left\{w_P^L-n_I^*|_\beta\left[\phi_U^0+I^\beta\left(1-\frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right\}+\beta\left(n_I^*|_\beta\right)^\varepsilon.$$

These imply Proposition 7b.

Putting together all the private-sector low-skilled workers, we have

$$\begin{aligned}
V^{DB,P,L}(\beta) &= \rho V_F^{DB,P,L}(\beta)+(1-\rho)V_I^{DB,P,L}(\beta)\\
&= \rho\theta(1-\theta)\frac{1+\mu B_E}{1+(1-\theta)\mu B_E}\cdot\\
&\quad\left\{w_P^L+\mu B_W-\mu\tau-n_F^{DB,P,L*}|_\beta\left[\phi_U^0-\mu B_C+I^\beta\left(1-\frac{\bar{n}_U}{n_F^{DB,P,L*}|_\beta}\right)\bar{\phi}_U\right]\right\}\\
&\quad+\rho\left[\beta\left(n_F^{DB,P,L*}|_\beta\right)^\varepsilon+\mu B_A\right]\\
&\quad+(1-\rho)\left\{(1-\theta)\,\theta\left\{w_P^L-n_I^*|_\beta\left[\phi_U^0+I^\beta\left(1-\frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right\}+\beta\left(n_I^*|_\beta\right)^\varepsilon\right\}.
\end{aligned}$$

Recall $V^{P,L}(\beta)$ in the benchmark case:

$$\begin{aligned}
V^{P,L}(\beta)=\rho&\left\{(1-\theta)\,\theta\left\{w_P^L-\mu\tau-n_F^*|_\beta\left[\phi_U^0+I^\beta\left(1-\frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\}+\beta\left(n_F^*|_\beta\right)^\varepsilon+\mu B_A\right\}\\
&+(1-\rho)\left\{(1-\theta)\,\theta\left\{w_P^L-n_I^*|_\beta\left[\phi_U^0+I^\beta\left(1-\frac{\bar{n}_R}{n_I^*|_\beta}\right)\bar{\phi}_R\right]\right\}+\beta\left(n_I^*|_\beta\right)^\varepsilon\right\}.
\end{aligned}$$

Then we can derive

$$\begin{aligned}
V^{DB,P,L}(\beta)-V^{P,L}(\beta) &= \rho\theta(1-\theta)\frac{1+\mu B_E}{1+(1-\theta)\mu B_E}\cdot\\
&\quad\left\{w_P^L+\mu B_W-\mu\tau-n_F^{DB,P,L*}|_\beta\left[\phi_U^0-\mu B_C+I^\beta\left(1-\frac{\bar{n}_U}{n_F^{DB,P,L*}|_\beta}\right)\bar{\phi}_U\right]\right\}\\
&\quad+\rho\left[\beta\left(n_F^{DB,P,L*}|_\beta\right)^\varepsilon+\mu B_A\right]\\
&\quad-\rho\left\{(1-\theta)\,\theta\left\{w_P^L-\mu\tau-n_F^*|_\beta\left[\phi_U^0+I^\beta\left(1-\frac{\bar{n}_U}{n_F^*|_\beta}\right)\bar{\phi}_U\right]\right\}+\beta\left(n_F^*|_\beta\right)^\varepsilon+\mu B_A\right\}
\end{aligned}$$

which yields the expression in the main text.

Migration Decision As a result, under $n_F^{DB,P,H*}|_\beta=n_F^{DB,P,L*}|_\beta\approx n_F^*|_\beta$, we can derive the indifference boundaries of the urban private-sector workers as in the main text. We then conclude with Proposition 7c.