NBER WORKING PAPER SERIES

EDUCATIONAL CHOICE, RURAL-URBAN MIGRATION AND ECONOMIC DEVELOPMENT

Pei-Ju Liao
Ping Wang
Yin-Chi Wang
Chong Kee Yip

Working Paper 23939
http://www.nber.org/papers/w23939

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2017

We are grateful for comments from Kaiji Chen, Chang-Tai Hsieh, B. Ravikumar, and Michael Song, as well as participants at the AREUEA-ASSA Annual Meetings, the Asian Meetings of the Econometric Society, the Conference on Chinese Economy at Tsinghua University, the Asian Bureau of Finance and Economic Research Conference, the Midwest Macro Meetings, the Public Economic Theory conference, the Society for Advanced Economic Theory Conference, the Symposium on Growth and Development at the Chinese University of Hong Kong, the Academia Sinica-Center for Chinese Economic Research joint workshop, and the annual meeting of the Society for Economic Dynamics, and seminar participants at Academia Sinica, Chinese University of Hong Kong, National Taiwan University and National Sun Yat-sen University, and Washington University in St. Louis. Travel support from the Chinese University of Hong Kong, Tsinghua University, the Center for Dynamic Economics of Washington University, and the Ministry of Science and Technology of Taiwan are greatly appreciated. We wish to express special thanks to Suqin Ge and Dennis Tao Yang for providing the data on the skill premium in China. Needless to say, the usual disclaimer applies. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w23939.ack

Educational Choice, Rural-urban Migration and Economic Development
Pei-Ju Liao, Ping Wang, Yin-Chi Wang, and Chong Kee Yip
NBER Working Paper No. 23939
October 2017
JEL No. O15,O53,R23,R28

ABSTRACT

Observing rapid structural transformation accompanied by a continual process of rural to urban migration in many developing countries, we construct a micro founded dynamic framework to explore how important education-based migration is, as opposed to work-based migration, for economic development, urbanization and city workforce composition. We then calibrate our model to fit the data from China over the period from 1980 to 2007, a developing economy featuring not only large migration flows but major institutional reforms that may affect work and education based migration differently. We find that, although education-based migration only amounts to one-fifth of that of work-based migration, its contribution to the enhancement of per capita output is larger than that of work-based migration. Moreover, the abolishment of the government job assignment for college graduates and the relaxation of the work-based migration have limited effects on economic development and urbanization. Furthermore, the increase in college admission selectivity for rural students plays a crucial but negative role in China's development, lowering per capita output and worsening the high-skilled employment share in urban areas.

Pei-Ju Liao
Department of Economics
National Taiwan Unversity
Taipei, Taiwan
pjliao@ntu.edu.tw

Ping Wang
Department of Economics
Washington University in St. Louis
Campus Box 1208
One Brookings Drive
St. Louis, MO 63130-4899
and NBER
pingwang@wustl.edu

Yin-Chi Wang
Department of Economics
Chinese University of Hong Kong
Shatin, NT, Hong Kong
ycwang@cuhk.edu.hk

Chong Kee Yip
Department of Economics
Chinese University of Hong Kong
Shatin, NT, Hong Kong
chongkeeyip@cuhk.edu.hk

1 Introduction

During the post-WWII period, many developing countries have experienced rapid structural transformation from traditional agricultural societies to modern economies. Accompanied by industrialization is a continual process of rural to urban migration, with labor force shifting toward more productive sectors in cities. Its importance has led to a renewed interest in studying structural change induced rural-urban migration, almost four decades after the celebrated contribution by Todaro (1969) and Harris and Todaro (1970). This newer literature has restricted the attention to work-based migration, with one noticeable exception by Lucas (2004) who emphasizes that the reason for migration is to accumulate human capital when working in a city (i.e., work-based migration with educational purposes). In the real world, we have seen parents sending children to urban areas for better quality of education or taking city jobs for their children to benefit from the urban provision of education. That is, the education-based migration may take place prior to the participation in the job market. Moreover, as stressed by Heckman (1976) and Rosen (1976), schooling not only leads to higher initial human capital at the entry to the job market but also improves the efficacy of on-the-job learning. That is, those sent by parents to take higher education in cities are expected to accumulate human capital at higher rates than low-skilled migrant workers under the learning mechanism elaborated by Lucas (2004).

Yet, this "education-based migration" channel has been completely overlooked in the literature. Just how important is the role played by education-based migration in economic development, urbanization and city workforce composition and how significant is this channel compared to work-based migration? To address these questions, we develop a micro-founded dynamic framework in which rural to urban migration is an integral part of economic development and structural transformation and in which the two distinct channels of migration, work versus education based, are both incorporated.

While the framework we plan to construct is general, to quantify the causes and consequences of the two distinct channels of migration requires disciplining the model to an economy of particular interest. We choose to explore the issues for the case of China. This is not only because of its large migration flows but also its specific institutions that may affect work and education based migration differently. Such institutions include, particularly, the *zhaosheng* system, the *hukou* system and the college admission reforms, which we will further elaborate below.[1] A natural question arises

[1] We will briefly discuss the institutional backround and reforms in Section 3 when we calibrate our model to fit

immediately: What are the effects of various institutions on the two channels of migration and the resulting changes in economic development, urbanization and workforce composition? Moreover, despite its rapid structural transformation and economic growth, China has experienced over the past three decades much faster migration of low-skilled workers compared to the education-based high-skilled. So what are the causes to the relative slowdown in the education-based channel and the imbalanced migration between the skilled and the unskilled, and what are its consequences for the process of development and urbanization?

Before turning to our modeling and calibration strategies, we would like to begin by summarizing some key observations using data from China from its open-door reform to the arrival of the worldwide Great Recession, 1980-2007. During this period, China has experienced rapid economic growth and urbanization. Real per capita GDP has grown at an annual rate of approximately 6.0 percent, whereas the comparable figure since Deng Xiao-Ping's Southern Trip in 1992 is 7.6 percent. Meanwhile, as shown in Figure 1(a,b), urbanization rates (urban population shares) and urban output shares have increased from 19.4 to 44.9 percent and from 66.7 to 87.3 percent, respectively, and the migration flows (proxied by changes in urban population) over rural population have nearly quadrupled, increasing from 0.5 to 1.9 percent.[2] Concurrently, more rural students were attending colleges because of the college expansion in the late 1990s, while empirical studies have pointed out the phenomenon that fewer rural students were admitted to top universities. The above observations motivate us to develop a theory to better understand the roles of skill development and educational choices in the decision of rural to urban migration.

Since most of the high quality universities are located in large metropolises in China, we consider cities as places for higher education to take place.[3] This compliments Lucas (2004) who views cities as places for immigrants to accumulate human capital when working. In so doing we explore a potentially important, education-based channel of rural to urban migration, which is called *zhaosheng* in China. Under the strict internal migration control in China, z*haosheng* has become a unique channel that mitigates migration barriers: Students attend colleges by passing the National College Entrance Examination, *gaokao*, to migrate to cities. This institutional migration channel enables us to examine the role of the education-based migration in the development of China and to compare

the Chinese economy. Detailed discussion of the institutional background in China is relegated to Appendix A.

[2]For urban output share, urbanization rates and migration outflows, the correlations range from 0.71 to 0.96.

[3]See Appendix A for the detailed information on the rural-urban disparities in college admission rates and the inequality in the distribution of educational resources in China.

the importance of this channel to that of work-based migration. Moreover, due to the college expansion and the facts that most universities are located in cities and urban high-skilled jobs are much better paid, one would expect that there shall be more youngsters migrating to cities for higher education. However, as shown in Table 1, the number and the annual growth rate of work-based (job transfer, job assignment and work or business) migration far outweighed those of education-based (studying or training). Therefore, it is worth to examine the factors that shaped rural youngsters' migration patterns and the causes leading to the growing difficulty for rural students to attend top colleges in urban areas.

To answer the aforementioned questions, we construct a dynamic equilibrium model that includes intergenerational migration decisions based on education. It allows human capital accumulation and intergenerational mobility to explore both work- and education-based rural to urban migration. Since college education is a large investment calling for full support by parents, it is natural to lay the framework in a two-period overlapping generations model with altruistic parents making crucial education-migration decisions for their children. Our theoretical analysis shows that as long as the expected net payoff of college education dominates the outside option inclusive of work-based migration, parents will send their children to cities to attend college. Parents will be more willing to send their children to attend colleges whenever their children are more talented, the college admission is less selective, the college education becomes cheaper, the job opportunities in urban areas improve, and the work-based migration cost increases. However, the effect of an increase in the education-based migration cost on parent's willingness to invest in college education is ambiguous.

We then calibrate the model to Chinese data during the period from 1980 to 2007. To properly capture some key policy changes, we separate our sample period into two regimes, 1980-1994 and 1995-2007. Such changes include the abolishment of the government job assignment (GJA) in 1994, the relaxation of *hukou*-induced migration barriers since the mid-1990s, as well as the rise in college tuition and the expansion of college admission in the 1990s.

Our quantitative analysis investigates the influences of both types of migration on development and urbanization, and captures such institutional changes and reforms during the sample period. We then further decompose the effects of these various factors on China's development and urbanization. Policy experiments on these institutional factors are also conducted. Our quantitative analysis suggests that the contribution of education-based migration is larger than that of work-based migration, accounting for 6.3 and 4.5 percent, respectively, of changes in per capita output during the sample period. Interestingly, even in the sub-sample period of 1995–2007, we obtain a similar pattern for the comparable contributions of education- and work-based migration (8.0 and 5.9 percent, respectively), although on average education-based migration only amounts to one-fifth of that of work-based migration as shown in Table 1. This finding suggests that without examining the education channel, the picture of rural to urban migration in China could be severely misleading.

Another important result is that total factor productivity (TFP) growth and the improvement in human capital together account for about two-third of changes in per capita output, while the impact of the termination of GJA and the relaxation of the work-based migration is limited. Last but not least, the increase in admission selectivity for rural students plays a crucial but negative role in China's development. Although the college expansion policy introduced in 1996 has provided broader access to students from rural areas, increasing admission selectivity for rural children lowers per capita output and seriously worsens the high-skilled employment share in urban areas. Therefore, rising admission selectivity not only makes the education-based migration slower but also hurts the skill-enhanced development in China.

Related Literature

The older literature on migration is mostly empirical adopting reduced-from approach or theoretical under static or partial equilibrium setting. One exception is Glomm (1992), which develops a dynamic general equilibrium model with persistent urbanization along the equilibrium path; another is Lucas (2004) which rests the analysis in a continuous time framework. Our paper adopts a dynamic macroeconomic model to study a new, namely, education-based, channel of rural-urban migration in China. It can therefore be compared with the recent, dynamic model based studies on job-related internal migration.

Bond, Riezman and Wang (2015) examined the effects of reductions in trade and migration barriers on China's growth and urbanization, focusing on China's accession to the World Trade Organization in 2002. Both Bond, Riezman and Wang (2015) and Restuccia, Yang and Zhu (2008) highlight migration barriers as a main driver for the surplus labor and low productivity in rural areas. Laing et. al (2005) construct a dynamic search equilibrium model to study the macroeconomic consequences of illegal migrants in China (the so-called *mangliu* or pleasant flood) due to the presence of surplus labor. Garriga, Tang and Wang (2015) studied the structural transformation and the consequent reallocation of labor from rural to urban areas in China. They found that two-thirds of the increases in housing prices can be attributed to this urbanization and development processes.

Tombe and Zhu (2015) quantified the effects of opening the domestic market to international trade on migration and regional income differences in China and found that effects of migration barrier reduction are much stronger than international trade liberalization. As rural to urban migration depends crucially on the trade off between rural land productivity versus urban wage premium and fringe benefits, Ngai, Pissarides and Wang (2016) showed that land policy is a major barrier on industrialization in China.[4]

The remainder of this paper is organized as follows. Section 2 presents the model and conducts the theoretical analysis. The calibration strategy, simulation, decomposition analysis on education- and work-based migration as well as the policy experiments are given in Section 3. Finally, Section 4 concludes the paper.

2 The Model

We develop a dynamic spatial equilibrium model with education and location choices to study the continual process of rural-urban migration and to evaluate the policies that govern it. We use (i, j, k) to denote three consecutive generations. There are two geographical regions, rural (R) and urban (U). Our optimization problem focuses only on the decision of rural parents (generation i) to send their children to urban areas to have their education. Assume that there is an initial mass of workers in urban areas given by (N_H, N_L), where N_H (N_L) denotes the total number of workers with high (low) skills. To focus on rural-urban migration, we only allow for exogenous reverse migration from urban to rural areas. The residences of urban households are assumed to pass from one generation to another. To simplify the analysis, we assume zero population growth so that the total population in rural and urban areas is constant over time.

2.1 Production

In urban areas, output is produced from the following non-homothetic CES production function:

$$Y_U = AF\left\{(N_H + \psi)\, h, N_L\right\}, \tag{1}$$

[4] There are other studies adopting computable general equilibrium frameworks or numerical simulation methods to investigate the relationship between migration barriers and rural to urban income inequality in China. These studies usually consider static or partial equilibrium settings with different research approaches.

where $A > 0$ is the technology scaling factor in urban areas (also called the urban TFP; hereafter, we will use these two terms interchangeably), ψ is a constant endowment that resides in urban areas, and h is the level of human capital possessed by high-skilled workers. The outcome of education is the acquisition of h, which is assumed to be a constant level.[5] The presence of ψ allows firms to produce positive output even if there are no high-skilled workers.

As it is well documented, the skilled labor wage of planning economies is commonly depressed.[6] To capture the distortion, we introduce a distortionary wedge $\tau \in (-1, \infty)$ faced by urban firms when hiring high-skilled workers. Denoting w_H as the effective high-skilled wage received by high-skilled workers and w_L as the low-skilled wage, we obtain the urban wage rates as follows:

$$(1+\tau)\, w_H = \frac{\partial Y_U}{\partial \left[(N_H + \psi)\, h\right]} = AF_H, \tag{2}$$

$$w_L = \frac{\partial Y_U}{\partial N_L} = AF_L, \tag{3}$$

where $F_H = \partial F / \partial \left[(N_H + \psi)\, h\right]$ and $F_L = \partial F / \partial N_L$. Then, the skilled-unskilled wage ratio is:

$$\frac{w_H h}{w_L} = \frac{h}{1+\tau} \frac{F_H}{F_L}. \tag{4}$$

Rural production uses only raw (or unskilled) labor, and constant returns requires a linear production technology:

$$Y_R = BN_R, \tag{5}$$

where N_R is the number of raw laborers in the rural area and B is the technology scaling factor in the rural area. A competitive labor market implies that the rural wage rate is:

$$w_R = B. \tag{6}$$

2.2 Rural households

Rural households (generation i) are altruistic and derive utility from both their own consumption (c^i) and their children's consumption (c^j). Each adult agent gives birth to a child. Assuming the utility function $u(\cdot)$ is strictly increasing and strictly concave, the representative household's objective is:

$$\Omega^i \left(\mathbf{I}^j | \mathbf{I}^i = 0, \mathbf{I}^k, x^j \right) = \max_{\mathbf{I}^j} u\left(c^i\right) + \beta \mathbb{E}_{\mathrm{X}} u\left(c^j\right), \tag{7}$$

[5] We can think of h as an index on labor quality or human capital that results from the total number of years in higher education.

[6] For instance, see Maurer-Fazio (1999) for a discussion of the case of China.

where β is the altruistic factor on children, and $\mathbf{I}^j$ is an indicator function of migration such that[7]

$$\mathbf{I}^j = \begin{cases} 0 \;\; if \text{ the household does not send generation } j \text{ (children) to college in an urban area} \\ 1 \;\; if \text{ the household sends generation } j \text{ (children) to college in an urban area.} \end{cases}$$

There are two types of costs in raising children. First, there is a basic requirement for resources, which we assume to be a constant child-rearing cost, denoted by ϕ^i. Second, there are costs to improve the child's quality, which we can summarize as education costs. Since talent matters for education because people who are more talented study more efficiently, we assume that part of the education costs depend on the talent of the child. Specifically, we denote the education cost as x^j, which is a random variable that consists of two parts; one part is inversely related to the talents of the child z^j and the college admission selectivity a, whereas the other part is a constant for basic education expenses b:

$$x^j \equiv \frac{1}{az^j} + b. \tag{8}$$

We note that z^j is drawn from a distribution with cumulative distribution function denoted by $G\left(z^j\right)$. Finally, as the education of children takes place only in urban areas, there is a constant migration cost for education denoted by σ_e which captures the basic moving expenses.[8] Thus, the budget constraint for a rural parent is:

$$c^i + \mathbf{I}^j \cdot \left(x^j + \sigma_e\right) + \phi^i = w_R. \tag{9}$$

Children who are sent to urban areas become skilled workers after receiving their education. They can obtain jobs that need high (low) skills in urban areas, earning a wage w_H (w_L) with probability γ_H (γ_L); otherwise, they are forced to move back to rural areas, earning a rural wage w_R. Children that remain in the rural area do not incur any cost in education or migration for their parents. When these children turn adults, they either get recruited as low-skilled workers in urban areas and earn w_L (with probability π) or work in a rural area and earn w_R. Here, π should be

[7] Because most tertiary schools in China are located in cities, we ignore the possibility of parents sending children to rural colleges. According to the data from the Chinese Ministry of Education, up to 2015, there are 2553 junior colleges, colleges and universities in China, and only 12 of them (0.47%) are located in county-level cities. The rest (99.53%) are located in prefectural-level cities or municipalities.

[8] The migration costs can be interpreted as the costs of obtaining the legal right to stay in cities, transportation costs between hometowns and cities and urban living costs. As college students usually enjoy cheaper housing provided by the universities, we differentiate migration costs into education-based and work-based migration costs.

regarded as "net" rural-urban migration for low skilled workers (i.e., migration inflows to cities net of outflows).[9] Their income (generation j) in the adulthood is given by:

$$W^j = \mathbf{I}^j \left[\gamma_H w_H h + \gamma_L w_L + (1-\gamma_H-\gamma_L) w_R\right] + \left(1-\mathbf{I}^j\right)\left[(1-\pi) w_R + \pi (w_L - \sigma_w)\right], \tag{10}$$

where σ_w is the constant work-based migration cost for the low-skilled workers. Then, the children's budget constraint is given by:

$$c^j + \mathbf{I}^k \cdot \left[\mathbf{I}^j (1-\gamma_H-\gamma_L) + \left(1-\mathbf{I}^j\right)(1-\pi)\right]\left(x^k + \sigma_e\right) + \phi^j = W^j, \tag{11}$$

where

$$\mathbf{I}^k = \begin{cases} 0 & if \text{ children do not send generation } k \text{ (grandchildren) to college in an urban area} \\ 1 & if \text{ children send generation } k \text{ (grandchildren) to college in an urban area} \end{cases}$$

and x^k is the education cost of grandchildren going to college in cities. When households of generation i decide $\mathbf{I}^j$, x^k is unknown. We use X to denote the random variable of education cost in their value function Ω^i. For illustration purposes, we plot the timeline of the model in Figure 2.

An agent's discrete choice problem is to decide whether to send his or her child to an urban area to attend college ($\mathbf{I}^j = 1$ versus $\mathbf{I}^j = 0$). That is, the agent compares $\Omega^i\left(1|0, \mathbf{I}^k, x^j\right)$ to $\Omega^i\left(0|0, \mathbf{I}^k, x^j\right)$ and chooses the highest value between the two. By substituting $c^i = w_R - \mathbf{I}^j \cdot \left(x^j + \sigma_e\right) - \phi^i$ and $c^j = W^j - \mathbf{I}^k \cdot \left[\mathbf{I}^j (1-\gamma_H-\gamma_L) + \left(1-\mathbf{I}^j\right)(1-\pi)\right]\left(x^k + \sigma_e\right) - \phi^j$ into the value functions, where W^j is given by (10), we have:

$$\begin{aligned} \Omega^i\left(1|0, \mathbf{I}^k, x^j\right) &= u\left(w_R - x^j - \sigma_e - \phi^i\right) \\ &\quad + \beta \mathbb{E}_{\mathrm{X}} u \begin{pmatrix} \gamma_H w_H h + \gamma_L w_L + (1-\gamma_H-\gamma_L) w_R \\ -\mathbf{I}^k(X)(1-\gamma_H-\gamma_L)(X+\sigma_e) - \phi^j \end{pmatrix} \end{aligned}$$

and

$$\Omega^i\left(0|0, \mathbf{I}^k, x^j\right) = u\left(w_R - \phi^i\right) + \beta \mathbb{E}_{\mathrm{X}} u\left[(1-\pi) w_R + \pi (w_L - \sigma_w) - \mathbf{I}^k(X)(1-\pi)(X+\sigma_e) - \phi^j\right].$$

Defining the net gain in value for sending children to urban areas to continue their education as $\Delta^i\left(\mathbf{I}^k, x^j\right) \equiv \Omega^i\left(1|0, \mathbf{I}^k, x^j\right) - \Omega^i\left(0|0, \mathbf{I}^k, x^j\right)$, we obtain:

$$\begin{aligned} \Delta^i\left(\mathbf{I}^k, x^j\right) &= u\left(w_R - x^j - \sigma_e - \phi^i\right) - u\left(w_R - \phi^i\right) \\ &\quad + \beta \mathbb{E}_{\mathrm{X}} \left\{ \begin{array}{c} u\left(\gamma_H w_H h + \gamma_L w_L + (1\text{-}\gamma_H\text{-}\gamma_L) w_R \text{-} \mathbf{I}^k(X)(1\text{-}\gamma_H\text{-}\gamma_L)(X+\sigma_e)\text{-}\phi^j\right) \\ \text{-}u\left((1\text{-}\pi) w_R + \pi (w_L\text{-}\sigma_w)\text{-}\mathbf{I}^k(X)(1-\pi)(X+\sigma_e)\text{-}\phi^j\right) \end{array} \right\}. \end{aligned} \tag{12}$$

[9] This is true from the macro perspective. Net flows from rural to urban areas have always been positive.

Further, we define $n \equiv (N_H + \psi)h/N_L$ to be the high-skilled to low-skilled labor ratio. Then, the high-skilled and low-skilled effective wage in (2) and (3) can be rewritten as:

$$(1+\tau)\, w_H = Af'(n), \quad w_L = A\left[f(n) - nf'(n)\right],$$

where $Af(n) = AF[n,1] = Y_U/N_L$. With w_H (w_L) is decreasing (increasing) in n, the skilled-unskilled wage ratio is decreasing in n. Defining n_s such that $w_H h/w_L = 1$ when $n = n_s$, we impose the following condition:

Condition S $w_H(n_s)\, h = w_L(n_s) > B + \sigma_w$.

If Condition S holds, then any urban job pays (net of the work-based migration cost) better than the rural job. To better understand Condition S, we plot the high- and low-skilled wages against n in Figure 3. Condition S requires that urban wages, net of work-based migration costs, are higher than rural wages. It is a condition to guarantee that, as long as children can find a job in cities, rural parents will send them to urban areas to attend college. Our next concern is the likelihood of finding a job in the urban area. We impose an assumption on the probabilities of acquiring an urban job: the probability of finding an urban job via education must be higher than that of finding one through other channels.

Assumption 1 $\gamma_H + \gamma_L > \pi$.

Assumption 1 states that the probability of securing an urban job after receiving a college education cannot be lower than the probability that a rural worker finds an urban job through non-educational channels. Thus, Condition S and Assumption 1 together imply that the expected urban wage income is higher than the rural wage income. Since urbanization and development depend on the composition and relative size of the urban workforce, Condition S and Assumption 1 simply highlights the fact that urban jobs are more attractive than rural jobs to the household. As a result, rural parents will consider sending their children to cities to attend college.

Although our framework is general, we can easily connect our model to the institutions in China. First of all, the relaxation of internal migration restrictions that has raised migrants' chance to get urban jobs is summarized by the probability parameters γ_H, γ_L and π. The changes in these three parameters represent the effects of policy reforms including the GJA and the loosened control on work-based migration. Next, changes in the education policy that alter the value of the education-based migration are given by the admission selectivity parameter a and the basic expenditure parameter b in the education cost variable x^j. These education parameters provide a short cut to study the effects of the *gaokao* system as well as the reforms like the college education expansion and increases in college tuition since 1996. Finally, the resulting reduction in the moving costs of rural-urban migration is captured by σ_e and σ_w.

2.3 Comparative statics

To have a better understanding of the comparative statics, we separate the effect of migration of (12) into two parts:

$$\begin{aligned} \Delta^i\left(\mathbf{I}^k, x^j\right) &= \underbrace{u\left(w_R - x^j - \sigma_e - \phi^i\right) - u\left(w_R - \phi^i\right)}_{\text{direct consumption effect}} \\ &\quad \underbrace{+\beta \mathbb{E}_{\mathrm{X}}\left\{u\left(c_U^j\right) - u\left(c_R^j\right)\right\}}_{\text{intergenerational effect}}. \end{aligned}$$

The direct consumption effect is always negative because parents' consumption is lower due to the education and migration costs, whereas the intergenerational effect is ambiguous. Condition S and Assumption 1 together assure that the intergenerational effect is positive which is necessary for parents to send their children to cities to attend college:[10]

Proposition 1 *Under Assumption 1 and Condition S, the intergenerational effect of migration is positive.*

The intuition of the above proposition is straightforward. If the probability of finding an urban job via education is reasonably high (Assumption 1) and rewarding (Condition S), then the higher expected urban wage provides an incentive for parents to pay the educational and migration costs of their children's education via altruism. Otherwise, this choice would not be a good "investment" from the parents' perspective. Thus, we conclude

Corollary *Under Assumption 1 and Condition S, if the positive intergenerational effect dominates the negative direct consumption effect, then parents will send their children to cities to attend college.*

We then examine how the net gain in education $\Delta^i\left(\mathbf{I}^k, x^j\right)$ responds to changes in the parameterization, i.e., we examine whether the "marginal" parent (a parent who is indifferent between sending her child to attend college in an urban area or keeping the child in the rural area so that

[10] All the proofs are relegated to Appendix B.

$\Delta^i\left(\mathbf{I}^k, x^j\right) = 0$) will send her child to receive an education. Based on the straightforward computation of comparative statics, we obtain the following proposition:

Proposition 2 *Under Condition S and Assumption 1, more parents will be willing to send their children to urban areas to attend college*

1. *when their children become more talented ($z_j \uparrow$), or when education becomes cheaper ($b \downarrow$).*
2. *when the chances that their children can obtain an urban job are higher ($\gamma_H \uparrow$, $\gamma_L \uparrow$).*
3. *when the chance of being low-skilled workers decreases ($\pi \downarrow$).*

Finally, to examine how changes in migration costs (σ_e and σ_w) affect parents' decisions, we first compute the effects of σ_e:

$$\frac{d\Delta^i\left(\mathbf{I}^k, x^j\right)}{d\sigma_e} = -u^i_{c_U} - \beta\mathbb{E}_{\mathrm{X}}\left\{\left[u^j_{c_U}\left(1-\gamma_H-\gamma_L\right) - u^j_{c_R}\left(1-\pi\right)\right]\mathbf{I}^k\left(X\right)\right\}.$$

The first term on the RHS highlights the standard negative direct effect of an increase in the education-based migration cost on parents' consumption. However, once the child is sent to the urban area, the higher future education-based migration cost of the grandchild can be saved. This indirect or intergenerational positive effect on the education-based migration cost is captured by the second term on the RHS (which is shown to be positive below). However, for the work-based migration cost σ_w, its effect only works through the intergenerational channel:

$$\frac{d\Delta^i\left(\mathbf{I}^k, x^j\right)}{d\sigma_w} = \pi\beta u^j_{c_R} > 0.$$

This is intuitive since a rise in the work-based migration cost makes education investment relatively more attractive to the parents. Thus, we can conclude the following:

Proposition 3 *Under Condition S and Assumption 1, if the intergenerational effect dominates the direct consumption effect, then more parents will be willing to send their children to attend college in urban areas when the education-based migration cost increases. For the work-based migration cost, its increase always makes parents willing to send their children to attend college.*

We have studied the rural-urban migration decision as an outcome of two opposing effects: a negative direct consumption effect on the parents and a positive intergenerational effect on the offsprings. If the latter dominates the former, then education-based migration takes place. Our comparative statics fit some salient features of the Chinese rural-urban migration experience. Specifically, we analyze the abolishment of the GJA policy as a fall in γ_H; the changes in college admission selectivity or the easiness for rural students to attend colleges as changes in a; the rising college tuition since 1992 as an increase in b; the loosening control over rural unskilled migrant workers in the 1990s as both an increase in π and a fall in the work-based migration cost σ_w. We will quantify our model to study the impacts of the changes in these parameters on the threshold talents – the education-based migration – as well as the important economic variables of China.

2.4 Evolution of workers

In this section, we study the population dynamics of rural-urban migration. Recall that adult agents supply labor to the market and that each agent gives birth to only one child, so the entire adult population participates in the labor market. Let $\left(N_H^t, N_L^t\right)$ be the skilled and unskilled workers in the urban area and N_R^t be the rural labor force, all at time t. Denote $J, K = \{H, L\}$ as the type of jobs for generation-j and generation-k urban workers. Let δ_{JK} be the transitional probability for an urban generation-k worker born to a generation-j urban worker with job J, working as a type K worker in an urban area. Thus, δ_{JK} captures job mobility across generations in the urban areas. In general, we expect that $\delta_{JJ} > \delta_{JK}$ for $J \neq K$, implying that the child is more likely to work as a type-J worker when the parent is a type-J worker. Under the assumption that the residences of urban households are passed from one generation to another, we have:

$$\sum_K \delta_{JK} = 1. \tag{13}$$

Then, the populations of skilled, unskilled and rural laborers evolve according to the following law of motion equations:

$$N_H^{t+1} = \delta_{HH} N_H^t + \delta_{LH} N_L^t + N_R^t \int \mathbf{I}^j\left(z^j, \mathbf{I}^k\right) \gamma_H dG(z^j), \tag{14}$$

$$N_L^{t+1} = \delta_{HL} N_H^t + \delta_{LL} N_L^t + N_R^t \left\{ \int \mathbf{I}^j\left(z^j, \mathbf{I}^k\right) \gamma_L dG(z^j) + \int \left[1 - \mathbf{I}^j\left(z^j, \mathbf{I}^k\right)\right] \pi dG(z^j) \right\}, \tag{15}$$

$$\begin{aligned} N_R^{t+1} = {} & \left(1 - \delta_{HH} - \delta_{HL}\right) N_H^t + \left(1 - \delta_{LH} - \delta_{LL}\right) N_L^t \\ & + N_R^t \left\{ \int \mathbf{I}^j\left(z^j, \mathbf{I}^k\right)\left(1 - \gamma_H - \gamma_L\right) dG(z^j) + \int \left[1 - \mathbf{I}^j\left(z^j, \mathbf{I}^k\right)\right]\left(1 - \pi\right) dG(z^j) \right\}, \end{aligned} \tag{16}$$

where the initial urban and rural labor forces are denoted by N_H^0, N_L^0 and N_R^0, respectively. Using (13), we can simplify (14)–(16) as follows:

$$N_H^{t+1} = \delta_{HH} N_H^t + (1-\delta_{LL}) N_L^t + N_R^t \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) \gamma_H dG(z^j), \tag{17}$$

$$N_L^{t+1} = (1-\delta_{HH}) N_H^t + \delta_{LL} N_L^t + N_R^t \left\{\pi + \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) (\gamma_L - \pi)\, dG(z^j)\right\}, \tag{18}$$

$$N_R^{t+1} = N_R^t \left\{(1-\pi) - \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) (\gamma_H + \gamma_L - \pi)\, dG(z^j)\right\}. \tag{19}$$

Finally, combining (17) and (18), we can see that the *residences* of urban households are passed from one generation to another:

$$N_U^{t+1} = N_U^t + N_R^t \left\{\pi + \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) (\gamma_H + \gamma_L - \pi)\, dG(z^j)\right\},$$

where $N_U^t \equiv N_H^t + N_L^t$ denotes the total urban workforce at time t.

Figure 4 plots the flows of workers in the model economy. For the quantitative analysis in the next section, we will focus on the case where all high-skilled parents produce high-skilled children, i.e., $\delta_{HH} = 1$.

2.5 Equilibrium

In equilibrium, all labor markets clear under the factor prices $\{w_H, w_L, w_R\}$ given by (2), (3) and (6):

$$N_J^{dt} = N_J^t, \quad J = H, L, R, \tag{20}$$

where N_J^{dt} denotes labor demand of type J. In addition, there is the overall population restriction for each period:

$$N_H^t + N_L^t + N_R^t = N, \tag{21}$$

where N is the constant population size in each period.

To conclude this section, we define the competitive equilibrium for our model.

Definition. *A* ***dynamic competitive spatial equilibrium*** *(DCSE) of the model consists of migration choice* $\left\{\mathbf{I}^j\right\}$ *and wage rates* $\{w_H, w_L, w_R\}$*, such that*

(i) *(Optimization)* given *wage rates* $\{w_H, w_L, w_R\}$, $\left\{\mathbf{I}^j\right\}$ *solves (7) subject to (9), (10) and (11);*

(ii) *(Market clearing) wage rates* $\{w_H, w_L, w_R\}$ *satisfy (2), (3) and (6), and labor markets clear according to (20); and*

(iii) *(Population identity) given the initial population* $\left\{N_H^0, N_L^0, N_R^0\right\}$ *and the distribution of talent* $G(z^j)$*, the population evolves according to (17)–(19) and is restricted by the identity (21).*

3 Quantitative Analysis

We are interested in studying the contribution of the education-based migration to the Chinese economy within the post-reform regime but before the financial tsunami, namely, the 1980–2007 period. Because the GJA policy was eliminated in 1994, it is natural to break the entire period into two sub-periods: Regime 1, spanning from 1980 to 1994, and Regime 2, ranging from 1995 to 2007.[11] We first conduct the two-regime calibration by fitting the model to the Chinese data. Based on the two-regime calibration, annual urban TFPs and distortions τ are drawn out. We then simulate the model annually, from 1980 to 2007, based on the calibrated parameters. Finally, the simulated economy is taken as our benchmark model and we proceed to perform analyses on decomposition and policy experiments to examine the roles of education-based migration in China's development.

3.1 Calibration and simulation

Because reverse migration of education-based is negligible in practice, we rule it out in our calibration. Below we briefly describe how we calibrate the model, while relegating calibration details and data sources to Appendix C.

3.1.1 Two-regime calibration

A model period is 25 years. Total population is normalized to one in every period. Urban (rural) population in the model is equal to the share of urban (rural) to total population and is computed using the data on populations by rural and urban residence. We term workers with educational attainment of college and above (below) as high (low)-skilled. Then, using the data on urban employment by educational attainment and the share of urban population, we compute the stocks of high- and low-skilled workers.

The utility function is assumed to take the standard CRRA form:

$$u(c) = \frac{c^{1-\varepsilon} - 1}{1 - \varepsilon}, \quad \varepsilon > 1,$$

where ε is the inverse of the elasticity of intertemporal substitution (EIS). In the literature, the Pareto distribution is commonly associated with wealth and income, which are believed to be closely

[11] Analogous to our theoretical model, we consider the whole Chinese economy to be two geographical regions, rural and urban, and dismiss the differentiation of within- and cross-provicial migration.

related to one's talent. Therefore, we assume that children's talents z^j follow a Pareto distribution, with the CDF given by:

$$G\left(z^j\right) = 1 - \left(\frac{z_{\min}}{z^j}\right)^{\theta}, \quad z^j \geq z_{\min},$$

where $z_{\min}$ and θ are the location and shape parameters of the Pareto distribution, respectively. Below, we first describe the preset common parameters and then the preset regime-specific parameters. Then we elaborate on the methods of identifying the remaining parameters.

China is well known for its high saving rates and low annual time preference rates. We thus set the annual time preference at 1 percent, which is close to Song et al. (2011). The parental altruistic factor for children β is hence equal to 0.7798. The inverse of the EIS parameter ε is set at 1.5, which is common in the literature. There is no nationwide survey of child-rearing costs for rural China. We follow the estimate in the literature to set ϕ such that the percentage of the child-rearing cost to rural household income $\tilde{\phi}$ is 17.4 percent in both regimes. For the Pareto distribution parameters, we set $z_{\min}$ to one according to the literature on firm size distribution, productivity and international trade.[12] Since talents are unobservable but are found to be correlated with income levels, we set θ to 2.5 using rural household net income data from the Chinese Household Income Project (CHIP). Our value is close to the estimate for the United States. The last preset common parameter is the elasticity of substitution between high- and low-skilled labor $1/(1-\rho)$. As pointed out by previous studies, the estimated values for Asian economies are usually larger, mostly falling between 2 and 7, than those for developed countries, ranging from 1 to 3. We thus choose the elasticity of substitution between high- and low-skilled labor to be 3 so that ρ equals 0.6667.

Denote $\tilde{\sigma}_e$ ($\tilde{\sigma}_w$) as the education-based (work-based) migration cost as a percentage of rural household income. Considering work-based migration cost as urban living costs and the required costs for moving to urban areas, we compute $\tilde{\sigma}_w$ from CHIP 2002 and obtain a value of 55.54% and 30.79% of rural household income for regimes 1 and 2, respectively. Education-based migration cost includes the costs of food and dormitory for a college student. Assuming that a student stays in college for four years and adjusting for model periods, we obtain the education-based migration cost $\tilde{\sigma}_e$ to be 0.1021 for regime 2. The data on education-based migration cost prior to 1996 is not available, so we compute $\tilde{\sigma}_e$ for regime 1 by assuming that $\tilde{\sigma}_e$ and $\tilde{\sigma}_w$ grow at the same rate across the two regimes and obtain $\tilde{\sigma}_e$=0.1841 for regime 1.

The main spirits of China's education reforms are captured by the endogenous threshold in

[12] See, for example, Ghironi and Melitz (2005), Bernard et al. (2003), and Eaton and Kortum (2002).

talents and parameters controlling admission selectivity and cost of college education, a and b. We will address the pinning down of the threshold talent and a using model equations later. College education was almost free of charge before 1990. Thus, b in regime 1 only includes stationary, materials and textbooks while b in regime 2 further includes tuition costs. Using Urban Household Survey (UHS) 2007 and 2008, b equals 0.48% and 5.28% of rural household income in regimes 1 and 2. The sharp change in b reflects the increase in college tuition in the late 1990s of China.

Under the linear rural production technology, the scaling factor B is equal to the rural wage rate. Being interested in the relative economic positions of rural and urban China and understanding how regional technological disparities shape individuals' migration decisions, we normalize rural per capita income in 2007 to 1. Then we compute the rural per capita income over 1980–2007. The average of B is 0.3685 and 0.7177 for regimes 1 and 2, respectively. It is notable that such normalization of rural per capita income together with $z_{\min} = 1$ imply that only parents with relatively talented children can afford to send their children to college. This is because rural parents have to maintain their own consumption and pay the child-rearing cost in addition to costs of college education and education-based migration.

We now turn to the rates at which college graduates find jobs and the migration probability for rural workers. All the job finding probabilities are the ones facing by each cohort. During the years of the GJA policy (1951–1994), a college graduate was assigned a stable job (either in the government or in state-owned enterprises), usually in an urban work unit. In contrast, after the termination of the GJA policy, jobs for college graduates were no longer guaranteed. In line with the GJA policy, we set $\gamma_H = 1$ and $\gamma_L = 0$ in regime 1, meaning that college graduates from rural China are fully employed ($\gamma_H + \gamma_L = 1$). For regime 2, the data on the employment rate of college graduates from rural China is not available. We thus use urban employment rates from CHIP in 1995, 2002 and 2007 to proxy for the employment rate of college graduates from rural areas. The average value, 0.9209, is set to be the employment rate in city districts for college graduates in regime 2. Note that γ_L is the job mismatching rate for college graduates, which we do not have information for. We set γ_L to 0.05, and γ_H is thus solved as 0.8709 in regime 2.[13] For the probability capturing the rate of work-based net migration flows π, as there is no nationwide survey on rural-urban migration in China during the periods under examination, we use changes in urban population as a proxy for rural-urban migration flows and compute π based on migration

[13] The calibration results are not sensitive to our choice of γ_L.

flows due to employment to rural population ratio.[14] As reported in Table 2, the average migration probabilities for rural workers π in regime 1 and regime 2 are 0.0036 and 0.0083, respectively.[15]

The next one is the human capital possessed by high-skilled workers relative to low-skilled workers, h. We first compute the average years of schooling for high- and low-skilled workers and take Mincerian coefficients from the literature for the two regimes. Following the Mincerian method, we then compute the regime-specific h and obtain 1.3529 and 1.5928 for regimes 1 and 2, respectively. The last preset regime-specific parameters are intergenerational mobility. Assuming that the residences of urban households are passed from one generation to another and allowing upward mobility, we have $\delta_{HH} = 1$, $\delta_{HL} = 0$ and $\delta_{LH} + \delta_{LL} = 1$ in both regimes.[16] The probabilities of remaining low-skilled workers across generations (δ_{LL}) in the two regimes are calibrated to match the N_H/N_L ratios using (17)-(19) and the *zhaosheng* flows data (computed in the same way as that for migration flows due to employment as reported in Table 2). δ_{LL} thus equals 0.9996 and 0.9883 in regimes 1 and 2, respectively.[17] The decreasing δ_{LL} shows that intergenerational mobility in China has improved over 1995–2007.

The regime-specific distortions τ faced by urban firms when hiring high-skilled workers, the urban TFPs A in the two regimes, the CES production high-skilled labor share α and the non-homothetic term ψ are calibrated to match the regime average skill premiums ($w_H h/w_L$), urban premiums (w_L/w_R) and urban output shares (Y_U/Y). Note that $\psi > 0$ in the CES production function implies decreasing returns to scale technologies, and ψ can be perceived as an urban infrastructure or producer rent that is used to facilitate production in cities. The targets of urban output shares thus contain additional information in addition to employment and wage measures and can serve to calibrate both α and ψ. The calibrated α and ψ are equal to 0.8461 and 0.0618, the regime-specific distortions τ are 7.1103 and 5.4763, and the urban TFPs in the two regimes are equal to 5.3877 and 11.0573, respectively. Our results show that urban TFP is growing faster relative to rural TFP: the

[14] Although Longitudinal Survey on Rural Urban Migration in China provides migration information, it only starts in 2008 which is not the period that we examine in this paper.

[15] We notice that migrants with different *hukou* status would have different urban benefits. However, our focus is the overall contribution of work-based migration compared to that of education-based. Our calibration is thus employment-based, rather than *hukou*-based. Considering workers' *hukou* status will not change our main results.

[16] The average years of schooling in China for people aged 15 and over have increased from 4.86 years in 1980 to 7.51 years in 2010, showing an overall pattern of upward mobility in education.

[17] We have matched the N_H/N_L data series and considered *xiagang* when computing urban employment rate. Thus, the reform of *xiagang* is being taken care of.

implied annual urban TFP growth rate is 5.47 percent. In addition, the distortion τ faced by urban firms in regime 2 is reduced by more than 22 percent compared to that in regime 1, indicating that the market distortions due to the planned economy have been greatly alleviated.

Denote $\hat{z}$ the threshold in children's talent such that when a child is equipped with the talent $\hat{z}$, his parents are indifferent between sending him to college or keeping him at home. When a child is talented such that $z^j \equiv 1/\left[a\left(x^j - b\right)\right] \geq \hat{z}$, his parents will definitely send him to colleges ($\Delta^i(\mathbf{I}^k, x^j) \geq 0$). The endogenous threshold $\hat{z}$ therefore dichotomizes the "destiny" of rural children. Specifically, define $\tilde{N}_E^t$ as the *zhaosheng* flow at time t. $\tilde{N}_E^t$ can be written as:

$$\tilde{N}_E^t = N_R^t \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) dG(z^j) = N_R^t \left(\frac{z_{\min}}{\hat{z}}\right)^\theta . \tag{22}$$

Therefore, $\hat{z} = z_{\min}(N_R^t/\tilde{N}_E^t)^{1/\theta}$ and $\hat{z}$ can be obtained using the *zhaosheng* flows data. The computed average $\hat{z}$ for regimes 1 and 2 are equal to 17.7632 and 13.1391, respectively. The decrease in $\hat{z}$ captures the college expansion in China: more rural students are going to colleges. With $\hat{z}$, we can solve the last parameter a by the indifference boundary equation (12). The calibrated a are 1.1489 and 0.4701 for regimes 1 and 2. The decrease in a reflects the fact of the draining in rural talents so that college admission is becoming more selective for rural students. This is consistent with the data that it becomes more difficult for rural students to attend top universities in China. Table 3 reports the calibration results. Based on the above parameters, our next step is to calibrate the annual urban TFP and distortions for 1981-2007 and to perform a simulation to serve as our benchmark model.

3.1.2 Calibration of the annual urban TFP and distortions

To calibrate the annual urban TFP and τ, we first need the annual N_R, N_H and N_L based on the model. Following the same method in the two-regime calibration, we compute the annual *zhaosheng* flows. Together with the data on N_R, N_H and N_L in 1980 and the calibrated parameters (including γ_H, γ_L, π, δ_{HH}, δ_{HL}, δ_{LH}, δ_{LL} and θ), we solve the threshold $\hat{z}$ of 1980 based on (22). The 1980–1981 work-based migration flows are also solved according to the equation: Work-based migration flow$_t = \pi N_R^{t-1}\left[1 - (z_{\min}/\hat{z})^\theta\right]$. Furthermore, from the evolution of workers equations (14)–(16), we compute the model implied N_R, N_H and N_L for 1981. We then repeat this procedure to obtain annual series for $\hat{z}$, N_R, N_H, and N_L. Assuming that the annual growth rate of human capital is constant over 1980–2007, we compute the annual series of h so that the average human capital in regimes 1 and 2 are exactly equal to those in the two-regime calibration. Finally, with the time series data on rural per capita income, the annual urban TFP A and distortions τ are solved to match the urban premium (w_L/w_R) and skill premium ($w_H h/w_L$).

Figure 5 plots the calibrated urban TFP and rural TFP for 1981–2007. It can be observed that the urban TFP grows relatively faster than the rural TFP after 1985, corresponding to China's economic reform, the privatization of state-owned enterprises and the deregulation of price controls. As reported in Table 3, the relative urban-rural TFP growth rate over our sample period is approximately 0.39 percent per year. Figure 6 provides a comparison between the model and the data on urban per capita output and total output per capita.[18] We define the urbanization rates in the model as the shares of urban workers. Figure 7 compares the model to the data on urbanization rates and the stocks of urban high- and low-skilled workers. Our model shows a lower urbanization rate and a smaller stock of low-skilled workers than the data do, with the discrepancies between the model and the data widening over time. The gaps can be explained by the migration flows inputted when we calibrate the model. Because our model only considers two channels of migration, the data on migrants who migrated for non-educational and non-employment reasons (accounting for approximately 50 percent of total migration) are thus excluded in the calibration. However, these migrants could migrate due to other reasons but became low-skilled workers later. As a result, our model underestimates the stock of low-skilled workers and the urbanization rate. As the model generates fewer workers in urban areas, especially fewer low-skilled workers, the urban per capita output in the model is slightly higher than that observed in the data. Additionally, as there are more rural workers in the model and rural technology is less productive, total output per capita in the model is slightly lower than that observed in the data.

This calibrated economy serves as our benchmark model for the decomposition analysis and for the policy experiments. Table 4 summarizes the annual averages of important macroeconomic variables in the benchmark model for regimes 1 and 2 as well as for the entire sample period. As expected, total output per capita in regime 2 is more than double that in regime 1, urban output shares increase, and urban employment shares increase. These imply that urban production becomes more important in regime 2. Furthermore, our model shows that the high-skilled employment shares in urban areas in regime 2 are more than quadruple the shares in regime 1, while the skill premium

[18] Based on equations (1) and (5), we use the N_H, N_L and N_R data to calculate urban per capita output and total output. See Appendix C for details.

still increases. These trends are all consistent with the experience of China's development.

3.2 Decomposition: The contributions of *zhaosheng* and work-based migration

To identify the contribution of each migration channel and to study the total effects of migration on China's development process, we eliminate the migration channels sequentially. The effect of the channel under study is thus the difference in results between the model with the channel being excluded and the benchmark model.

Figure 8 plots urban per capita output, total per capita output, and urbanization rates under the decomposition. The benchmark model and the three scenarios are plotted for comparison: (1) work-based migration is eliminated; (2) *zhaosheng* is eliminated; and (3) both migration channels are eliminated. In the first scenario, when the work-based migration is eliminated, the only "new" source of low-skilled workers coming from countryside is unlucky college graduates. Consequently, there are much fewer productive low-skilled workers in cities, resulting in a larger high-to-low skilled labor ratio and a higher urban per capita output. Furthermore, as the migration volume via the work-based migration is large, the urbanization rate in this scenario is much lower than the benchmark case. In the second scenario in which *zhaosheng* is eliminated, once again, as the volume of migration through the *zhaosheng* channel is small, the urbanization rate in this case is very close to that in the benchmark model. This shows that the majority of rural-urban migration is work-based. With fewer productive high-skilled workers in the cities, urban per capita output is now slightly lower than that in the benchmark case.

To identify the magnitude of the contribution of migration types to major macroeconomic variables, Table 5 reports the percentage change relative to the benchmark model for the above three scenarios. Given the large volume of work-based migration, the conventional wisdom is that the effects of work-based migration on output levels should far outweigh the effects of education-based migration. However, our results in Table 5 show that the contribution of education-based migration cannot be overlooked: *zhaosheng* and work-based migration explain 6.3 percent and 4.5 percent of total output per capita in the benchmark model over the entire sample period, respectively.[19] We also find that *zhaosheng* contributes to roughly one-third of the high-skilled employment share in

[19] Our framework does not consider spillover effects in cities, as mentioned in Lucas (2004). If low-skilled workers can benefit from high-skilled workers in urban areas, the role of education-based migration in our model will be even more important.

the benchmark model and thereby lowers the skill premium, while work-based migration reduces the high-skilled employment share and boosts the skill premium. Furthermore, the result suggests that *zhaosheng* is more important in regime 2 than in regime 1: *zhaosheng* in regime 2 explains 8.0 percent of total output per capita in the benchmark model, while it only explains 2.0 percent of total output per capita in regime 1. There are several conflicting forces influencing the effects of *zhaosheng*: a higher skill premium, a higher human capital level and a lower education-based migration cost in regime 2 attract more migration through the *zhaosheng* channel, whereas the higher tuition cost and the termination of the GJA policy depress education-based migration. Our quantitative results show that the three positive effects dominate the two negative effects. The effects of *zhaosheng* on total output per capita and urban employment share in regime 2 are larger than those in regime 1.

As shown in Table 5, our results also show rich interactions between *zhaosheng* and work-based migration on the skill premium, high-skilled employment share, total output per capita, and urban output and employment shares. It is intuitive that the interactive effect is strongest on the high-skilled employment share (accounting for 11 percent of its change over the entire sample period) because work-based migration leads to a higher skill premium that attracts more education-based migration. For the other variables, several conflicting forces are involved in the resulting interactive effect. First, if work-based migration is not allowed, rural residents can still move to urban areas via the *zhaosheng* channel. Second, high-skilled workers (mainly from *zhaosheng*) and low-skilled workers (mainly from work-based migration) are substitutes in production. Third, when there is a larger stock of low-skilled workers in the cities, the skill premium is boosted up. The higher skill premium thus encourages more parents to send their children to cities to attend college. Fourth, there exists upward intergenerational mobility. The last two forces are positive, while the first two are negative. The results show that the skill premium is the dominant effect; thereby, a minor but positive interaction between *zhaosheng* and work-based migration is observed.

3.3 Factor decomposition

China experienced huge changes in the past decades, such as institutional reforms on education, market intervention and migration regulation. Here we provide an eleven-factor decomposition in regime 2 to study the contribution of each factor of our concern to the development of China. The eleven factors refer to the abolishment of the GJA policy (lower γ_H in regime 2), better work-based job opportunities (higher π in regime 2), an increase in the education-based migration cost (higher σ_e), an increase in the work-based migration cost (higher σ_w), increases in both urban and rural TFP, an improvement in human capital h, an increase in child-rearing cost (higher ϕ), less market distortion (lower τ in regime 2), better intergenerational mobility (lower δ_{LL} in regime 2), rising admission selectivity (lower a in regime 2), and an increase in college tuition (higher b in regime 2). Each experiment is conducted by setting the corresponding parameter back to the level of regime 1, while others remain unchanged. The percentage change relative to the regime 2 benchmark model (as reported in Table 4) is then computed.

The results of the factor decomposition are provided in Table 6 and are summarized below. First of all, the TFP growth, the improvement in human capital and the better intergenerational mobility contribute the most to the increases in the total output per capita, accounting for 52.9%, 10.8% and 12.3% respectively, whereas the rising admission selectivity greatly damp the total output per capita (depressed by 24.8%). Secondly, the better intergenerational mobility, the improvement in human capital and the TFP growth also matter for the increase in urban output share, accounting for 3.2%, 3.0% and 1.8% of the increase, respectively. However, the effect is offset by the rising admission selectivity (-4.9%). Third, urban employment share rises due to better work-based job opportunities, accounting for 8.2% of the increase, but is depressed by the rising admission selectivity (-12.4%). Fourth, intergenerational mobility and TFP growth are both important in increasing the skilled employment share (accounting for 49.3% and 5.5% respectively), whereas the high-skilled employment share is decreased by the rising admission selectivity, the higher college tuition and better work-based job opportunities (-64.2%, -6.1% and -7.3%, respectively). Finally, among all the factors, lower labor market distortion is found to be the most important factor that leads to the increase in skill premium. Other factors contributing to the increase in skill premium include the improvement in the quality of human capital and the rising admission selectivity, whereas the improvement in intergenerational mobility drags the skill premium down.

Compared with other factors, we find that the rising college admission selectivity plays a crucial but negative role in China's development during 1994-2007. Admissions are becoming more selective for rural students. This could be due to the fact that skilled parents tend to move to cities, resulting in the brain drain phenomenon from rural to urban areas. Since it is more difficult for rural students to attend top universities, rural parents have lower incentives to send their children for higher education in urban areas (fewer education-based migration). This provides a possible explanation to the imbalanced migrations between the skilled and the unskilled.

3.4 Policy experiments

Based on the benchmark model, we are now ready to provide policy experiments. Two experiments are conducted in this section. The first one is related to the education-based migration, discussing the scenario without the GJA policy throughout the history in China. As the GJA had been in force in China from the 1950s to the mid 1990s, we wonder how would the economy perform if China had not implemented the GJA policy throughout its history. The second experiment explores the effect of a more relaxed regulation on the work-based migration since 1980. Because of the *hukou* reforms, the regulations on work-based migration have been gradually relaxed. We are curious what China would look like if the government had maintained looser regulations for migrant workers. Table 7 summarizes the percentage changes relative to the benchmark model for the two experiments. The details are given below.

In the first experiment, the value of γ_H in regime 1 is set to that of regime 2. That is, jobs are not guaranteed for college graduates in regime 1 anymore. As shown in Table 7, there are two opposite effects of this policy. Without guaranteed high-skilled jobs, college education becomes less rewarding, resulting in fewer education-based migration. However, the skill premium increases because of the decreasing supply of high-skilled workers, which makes college education more rewarding. Our quantitative result suggests that the former effect is larger. Therefore, without the GJA policy throughout the history, urban employment would decrease by 0.5 percentage, the share of high-skilled employment would decrease by approximately 7 percentage, the skill premium would increase by 0.7 percentage and the total output per capita would decline by 1.2 percentage. We thus conclude that the impact of no GJA on China's development is relatively small. The only more negative effect of not having the GJA is that the high-skilled employment share is smaller. The minor effect of no GJA could be due to the fact that the rate at which college graduates find jobs in the benchmark model is very close to one.

In the second experiment, the value of π in regime 1 is increased to that of regime 2. The result in Table 7 suggests that, with a relaxed regulation on work-based migration, there would be more work-based migrants, resulting in a larger share of urban employment and an increase in both urban output share and total output per capita. However, the relaxation leads to a lower share of high-skilled employment; thereby a higher skill premium. Compared with the GJA policy, the regulation on work-based migration has a larger impact on China's urbanization and development.

4 Conclusions

Economic development is usually associated with a process of structural transformation and urbanization. Rural to urban migration triggers the process. In this paper we have constructed a dynamic spatial equilibrium model with a focus on a non-conventional migration channel: education-based migration. We have then conducted quantitative analysis, taking China as an example of special interest to examine the causes and consequences of education and work-based rural-urban migration in its development process. We have performed various decomposition analysis and policy experiments.

The main takeaway of our quantitative analysis is that migration indeed has an important contribution to the development of China: rural-urban migration accounted for nearly 11 percent of per capita output throughout the 1981-2007 period. Particularly, we find that the effect of education-based migration is larger than that of work-based migration, explaining 6.3 and 4.5 percent of output per capita, respectively. Because of the considerable impact of education-based migration, ignoring the education channel would severely under-estimate the effects of migration, particularly the skill-enhanced process of migration. This strong skill enhancing effect of education is consistent with the celebrated contribution by Heckman (1976) and Rosen (1976). Moreover, our results suggest that TFP growth and the improvement of human capital together account for about two-thirds of changes in per capita output. The abolishment of GJA and the relaxation of work-based migration have limited impacts on the development. Furthermore, the more selective college admission for rural students plays a significantly negative role in the development of China, especially lowering the high-skilled employment share seriously. It offsets the skill-enhanced development process in China.

Along these lines, it would be interesting to extend our framework to study various migration issues in developing countries. A possible extension is to allow urban unskilled workers to accumulate human capital in cities, as in Lucas (2004). This will further enhance the importance of the education-based migration channel. Another extension might examine different underlying channels of the work-based migration, in particular, the early sample stage of the *zhaogong* channel into state-owned enterprises and the later stage into both state-owned enterprises and private sector jobs. Moreover, the investment-oriented channel via the blue-stamp scheme for setting up private businesses as well as investments in properties and factories is worth exploring. We will leave these topics for future research.

References

[1] Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, 2003, "Plants and productivity in international trade," *American Economic Review*, 93(4), pp. 1268-1290.

[2] Bond, Eric, Raymond Riezman, and Ping Wang, 2015, "Urbanization and economic development: A tale of two barriers," Working Paper.

[3] Eaton, Jonathan and Samuel Kortum, 2002, "Technology, geography and trade," *Econometrica*, 70, pp. 1741–1779.

[4] Garriga, Carlos, Yang Tang, and Ping Wang, 2015, "Rural-urban migration, structural transformation, and housing markets in China," Federal Reserve Bank of St. Louis, Working Paper 2014-028A.

[5] Ghironi, Fabio and Marc J. Melitz, 2005, "International trade and macroeconomic dynamics with heterogeneous firms," *The Quarterly Journal of Economics,* 120, pp. 865-915.

[6] Glomm, Gerhard, 1992, "A model of growth and migration," *The Canadian Journal of Economics*, 25(4), pp. 901-922.

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

[8] Heckman, James J., 1976, "A life-cycle model of earnings, learning, and consumption," *Journal of Political Economy*, Vol. 84, No. 4, Part 2: Essays in labor economies in honor of H. Gregg Lewis, pp. S11-S44.

[9] Laing, Derek, Chuhwan Park, and Ping Wang, 2005, "A modified Harris-Todaro model of rural-urban migration for China," In: F. Kwan and E. Yu (Eds.), *Critical Issues in China's Growth and Development*. London: Ashgate, pp. 245-264.

[10] Lucas, Robert E., 2004, "Life earnings and rural-urban migration," *Journal of Political Economy*, Vol. 112, No. S1, pp. S29-S59.

[11] Maurer-Fazio, Margaret, 1999, "Earnings and education in China's transition to a market economy: Survey evidence from 1989 and 1992," *China Economic Review*, 10, pp. 17-40.

[12] Ngai, L. Rachel, Christopher A. Pissarides, and Jin Wang, 2016, "China's mobility barriers and employment allocations," Working paper.

[13] Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu, 2008, "Agriculture and aggregate productivity: A quantitative cross-country analysis," *Journal of Monetary Economics*, 55(2), pp. 234–250.

[14] Rosen, Sherwin, 1976, "A theory of life earnings," *Journal of Political Economy*, Vol. 84, No. 4, Part 2: Essays in labor economies in honor of H. Gregg Lewis, pp. S45-S67.

[15] Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, 2011, "Growing like China," *American Economic Review*, 101(1), pp. 196-233.

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

[17] Tombe, Trevor and Xiaodong Zhu, 2015, "Trade, migration and regional income differences: evidence from China," Working paper.

Table 1: MIGRATION BY REASONS

Year	Population outflow	Job Transfer	Job Assignment	Work or Business	Study or Training	Other reasons
Percentage						
1985	100.00	29.57	8.04	3.08	11.26	48.05
2000	100.00	5.32	3.76	33.55	6.84	50.53
Average	100.00	17.44	5.90	18.32	9.05	49.29
Population (thousand persons)						
1985	10770.00	3184.23	866.43	331.75	1212.81	5174.78
2000	21580.00	1148.73	810.36	7240.20	1475.87	10904.83
Annual Growth	4.74%	-6.57%	-0.45%	22.82%	1.32%	5.09%

Data source: Migration by reasons (percentage) is obtained from *The 10 Percent Sampling Tabulation on the 1990 Population Census of the People's Republic of China* and *The Tabulation on the 2000 Population Census of the People's Republic of China*. Migration reasons include migration due to job transfer, job assignment, work or business, study and training, to relative and friend, retired or resigned (1985 data only), moved with family, marriage, pull down and move (2000 data only) and other reasons. We categorize migration due to job transfer, job assignment and work or business as work-based migration, and migration due to study or training as migration via *zhaosheng*. Note: There is no available national-wide survey on population outflow (rural-urban migration) in China. Thus, we use changes in urban population as a proxy for population outflow. In the table, migrant population by reasons is computed based on the proxy for population outflow.

Table 2: *ZHAOSHENG* FLOW AND THE PROBABILITY OF WORK-BASED MIGRATION

	Zhaosheng flow	Prob. of work-based migration
Regime 1	0.00058946	0.003554486
Regime 2	0.00114381	0.008281515

Source: Authors' calculation using the average of 1985 and 2000 migration reasons in Table 1.

Table 3: CALIBRATION

	Parameter values		Target	Target values		Explanation/Source
	Regime 1	Regime 2		Regime 1	Regime 2	
Preset						
γ_L	0	0.05				preset
δ_{HH}	1	1				preset
Calibrated: Regime-common						
β	0.7798					annual discount factor=1%
ε	1.5					elasticity of intertemporal substitution=2/3
z_{min}	1					standard setup in the literature
$\tilde{\phi}$	0.174	0.174				percentage of rural income.
θ	2.5					computed using rural household net income data in CHIPS 1995 and 2002
ρ	0.6667					elasticity of substitution between high/low skilled=3
α	0.8461		Y_U/Y	0.6922		match urban income shares
ψ	0.0618		Y_U/Y		0.8294	match urban income shares
Calibrated: Regime-specific						
γ	-	0.9209				urban employment rate, using data in CHIPS 1995, 2002 and 2008
γ_H	1	0.8709				$\gamma_H = \gamma - \gamma_L$
π	0.0036	0.0083				probability of work-based migration in Table 2
δ_{LL}	0.9996	0.9883	N_H/N_L	0.0424	0.1466	match the ratio of high-low skilled worker
B	0.3685	0.7177				$w_R = 1$ in 2007
h	1.3529	1.5928				Mincerian rate of return
A	5.3877	11.0573	w_L/w_R	1.7781	2.0076	match urban premium
τ	7.1103	5.4763	$w_H h/w_L$	1.2296	1.6576	match skill premium
$\hat{z}$	17.7632	13.1391		0.059%	0.114%	match regime average of *zhaosheng* flow in Table 2
$\tilde{\sigma}_e$	0.1841	0.1021				percentage of rural household income
$\tilde{\sigma}_w$	0.5554	0.3079				percentage of rural household income
a	1.1489	0.4701				solved by the indifference boundary equations
b	0.0048	0.0528				percentage of rural household income, using data in UHS 2007 and 2008
Model implications						
$1-G(\hat{z})$	0.075%	0.160%				*zhaosheng* proportion
A/B	14.6188	15.4071				urban-rural TFP ratio
ψ_{cost}	0.6459	0.4380				unit cost reduced by ψ
A_g		5.47%				average annual growth rate of A in 1981-2007
$(A/B)_g$		0.39%				average annual growth rate of A/B in 1981-2007

Note: Calibrated results are not sensitive to the value of γ_L. α, ψ, A and τ are actually solved together to match the corresponding targets listed in the table.

Table 4: BENCHMARK MODEL

Period	Total output per capita Y/N	Urban output Y_U/Y	Urban employment $(N_H+N_L)/N$	High-skilled employment share $N_H/(N_H+N_L)$	Skill premium $(w_H h/w_L)$
Whole: 1981-2007	1.6206	0.7148	0.2516	0.0784	1.4571
Regime 1: 1981-1994	0.8811	0.6585	0.2174	0.0327	1.2575
Regime 2: 1995-2007	2.4169	0.7754	0.2883	0.1277	1.6720

Table 5: DECOMPOSITION - MIGRATION CHANNELS

Unit: Percentage change

Period	Total output per capita Y/N	Urban output Y_U/Y	Urban employment $(N_H+N_L)/N$	High-skilled employment share $N_H/(N_H+N_L)$	Skill premium $(w_H h/w_L)$
Zhaosheng					
Whole: 1981-2007	6.3%	1.9%	2.8%	30.8%	-3.1%
Regime 1: 1981-1994	2.0%	1.0%	1.1%	30.6%	-1.2%
Regime 2: 1995-2007	8.0%	2.8%	4.2%	30.8%	-4.7%
Work-based migration					
Whole: 1981-2007	4.5%	3.3%	19.9%	-21.7%	7.2%
Regime 1: 1981-1994	0.8%	1.7%	9.7%	-11.5%	3.5%
Regime 2: 1995-2007	5.9%	4.8%	28.1%	-24.5%	10.2%
Interactive migration					
Whole: 1981-2007	0.1%	0.4%	0.2%	11.0%	0.1%
Regime 1: 1981-1994	0.0%	0.0%	0.0%	4.4%	0.1%
Regime 2: 1995-2007	0.2%	0.7%	0.4%	12.8%	0.2%
Non-migration factors					
Whole: 1981-2007	89.1%	94.4%	77.1%	79.9%	95.8%
Regime 1: 1981-1994	97.3%	97.3%	89.2%	76.5%	97.6%
Regime 2: 1995-2007	85.8%	91.8%	67.2%	80.8%	94.3%

Note: Numbers reported in the table are the percentage changes relative to the benchmark model. For example, total output per capita is 1.6206 for the whole period in the benchmark model and 1.5178 in the scenario with the channel of work-based migration only. Therefore, the channel of *zhaosheng* explains 6.3% of total output per capita in the benchmark model.

Table 6: FACTOR DECOMPOSITION

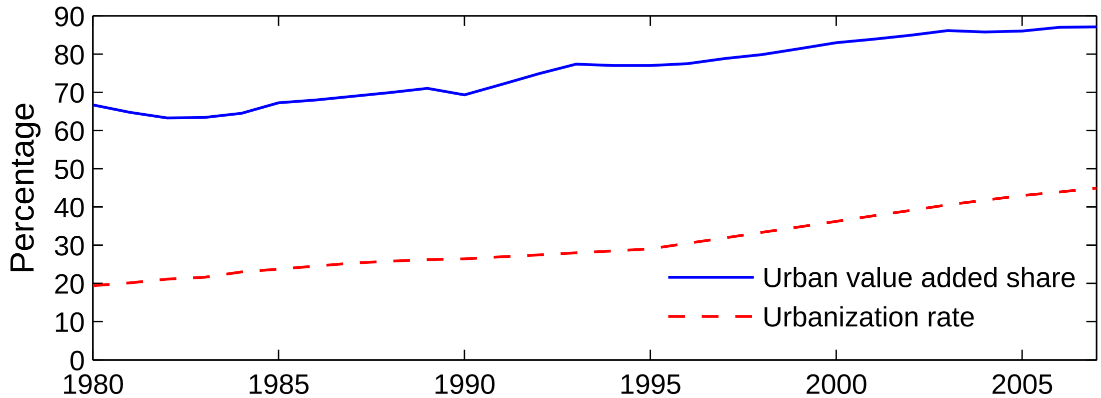
Unit: Percentage change

Factors	Total output per capita Y/N	Urban output Y_U/Y	Urban employment $(N_H+N_L)/N$	High-skilled employment share $N_H/(N_H+N_L)$	Skill premium $(w_H h/w_L)$
Abolishment of the GJA (lower γ_H)	-0.9%	-0.2%	-0.4%	-2.9%	0.4%
Better work-based job opportunities (higher π)	1.5%	1.2%	8.2%	-7.3%	2.9%
An increase in the education-based migration cost (higher σ_e)	-0.3%	-0.1%	-0.1%	-0.8%	0.1%
An increase in the work-based migration cost (higher σ_w)	0.0%	0.0%	0.0%	0.0%	-0.0%
Increases in urban and rural TFP	52.9%	1.8%	1.0%	5.5%	-0.8%
An improvement in human capital (higher h)	10.8%	3.0%	0.3%	2.1%	9.8%
An increase in the child-rearing cost (higher ϕ)	-1.1%	-0.3%	-0.5%	-3.2%	0.5%
Lower market distortion (lower τ)	1.2%	0.3%	0.6%	3.5%	21.4%
Better intergenerational mobility (lower δ_{LL})	12.3%	3.2%	-0.0%	49.3%	-9.9%
Rising admission selectivity (lower a)	-24.8%	-4.9%	-12.4%	-64.2%	8.5%
An increase in college tuition (higher b)	-2.0%	-0.5%	-1.0%	-6.1%	0.9%

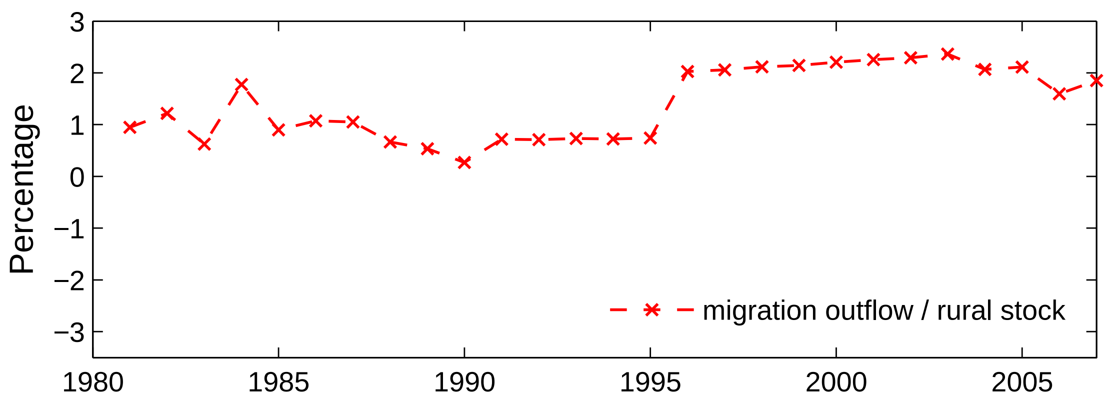
Table 7: POLICY EXPERIMENTS

Unit: Percentage change

Period	Total output per capita Y/N	Urban output Y_U/Y	Urban employment $(N_H+N_L)/N$	High-skilled employment share $N_H/(N_H+N_L)$	Skill premium $(w_H h/w_L)$
No GJA in regime 1					
Whole: 1981-2007	-1.2%	-0.4%	-0.5%	-7.0%	0.7%
Regime 1: 1981-1994	-0.7%	-0.3%	-0.3%	-10.0%	0.4%
Regime 2: 1995-2007	-1.4%	-0.5%	-0.7%	-6.1%	0.9%
Better job oppertunities in regime 1: $\pi_1 = \pi_2$					
Whole: 1981-2007	2.8%	2.5%	14.4%	-6.6%	4.2%
Regime 1: 1981-1994	0.9%	2.2%	12.5%	-11.1%	4.1%
Regime 2: 1995-2007	3.6%	2.7%	16.0%	-5.4%	4.2%



(A) URBANIZATION RATES AND URBAN OUTPUT SHARES



(B) MIGRATION OUTFLOWS

Figure 1: URBANIZATION PROCESS IN CHINA OVER 1980-2007

Note: Urbanization rate is defined as urban population shares out of total population. Agricultural sectors are excluded when urban output shares are computed. Because there is no good data on migration, we use changes in urbanization as a proxy for migration outflow.

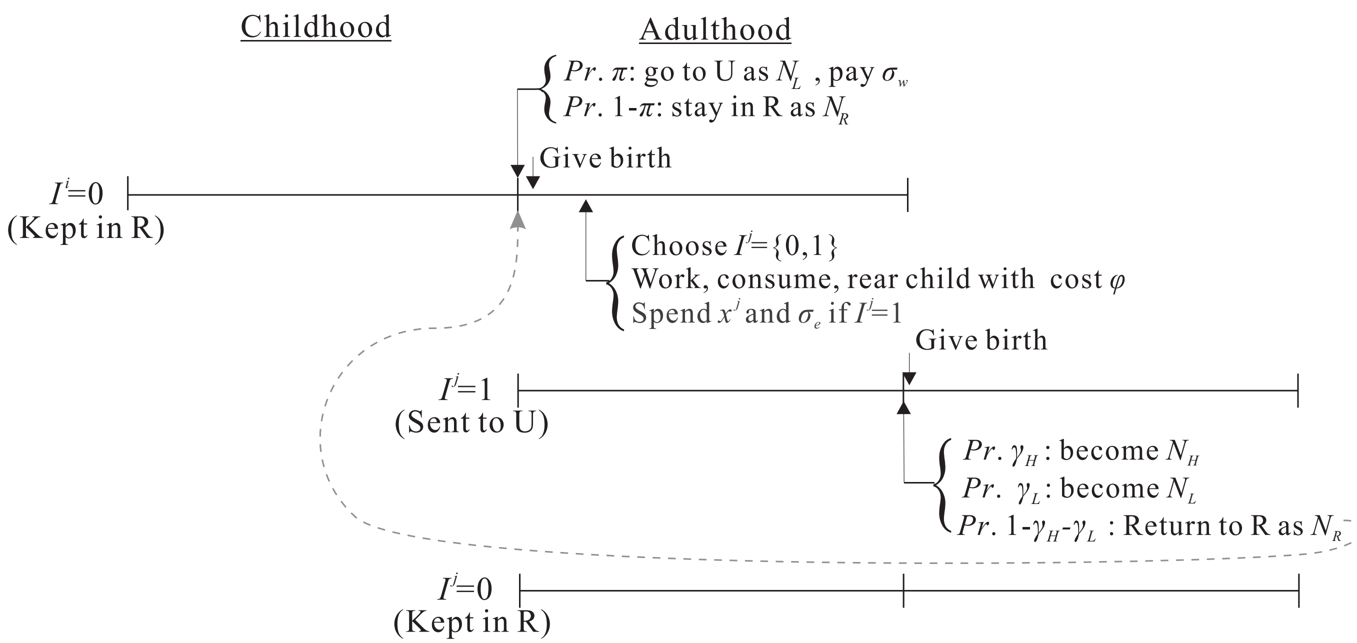


Figure 2: TIMELINE OF THE MODEL

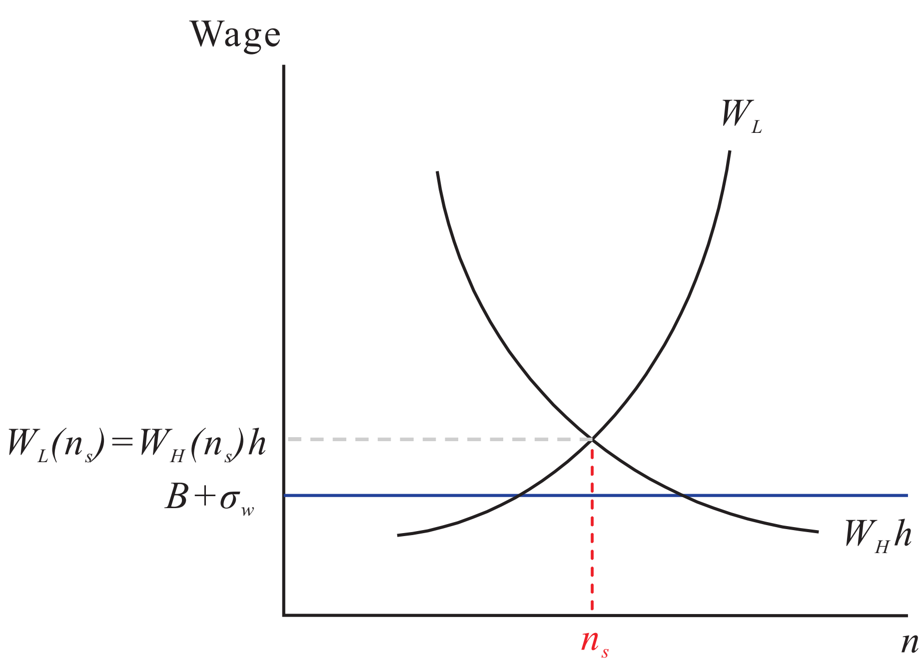


Figure 3: HIGH- AND LOW-SKILLED WAGES AND RURAL WAGE RATE

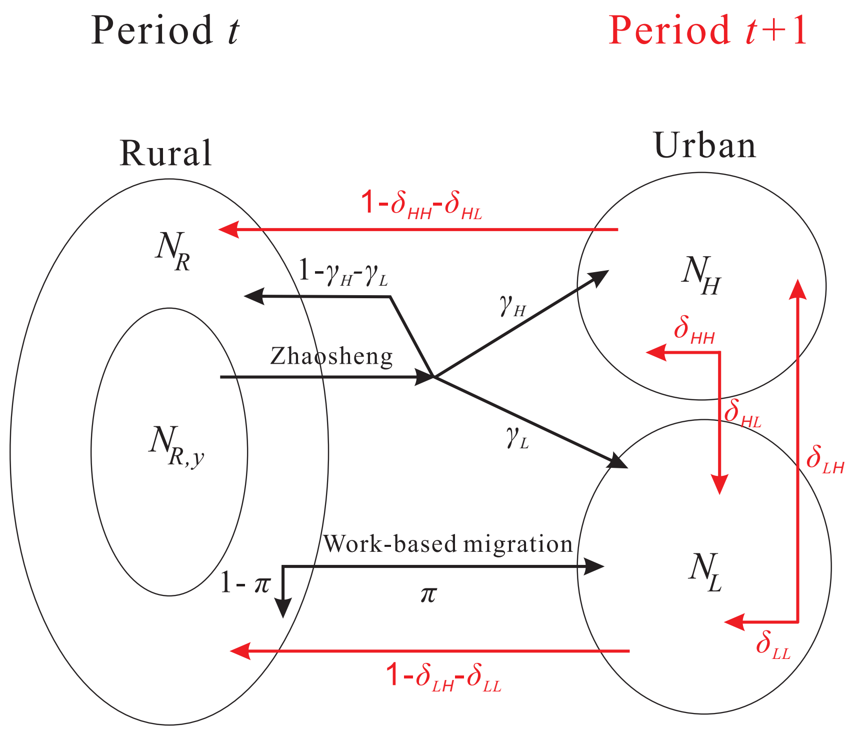


Figure 4: FLOWS OF WORKERS

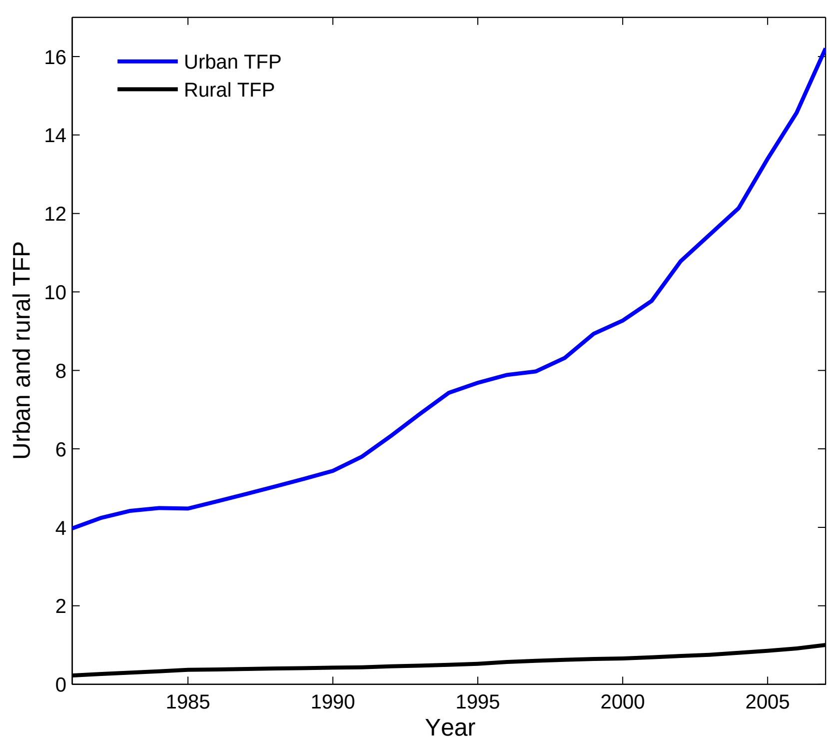


Figure 5: CALIBRATED URBAN AND RURAL TFP DURING 1981-2007

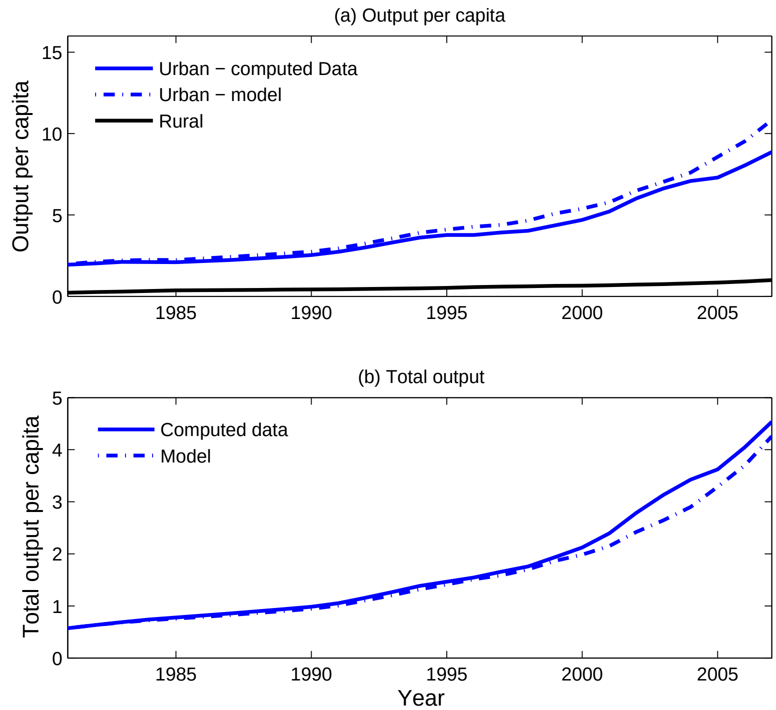


Figure 6: BENCHMARK MODEL - OUTPUT

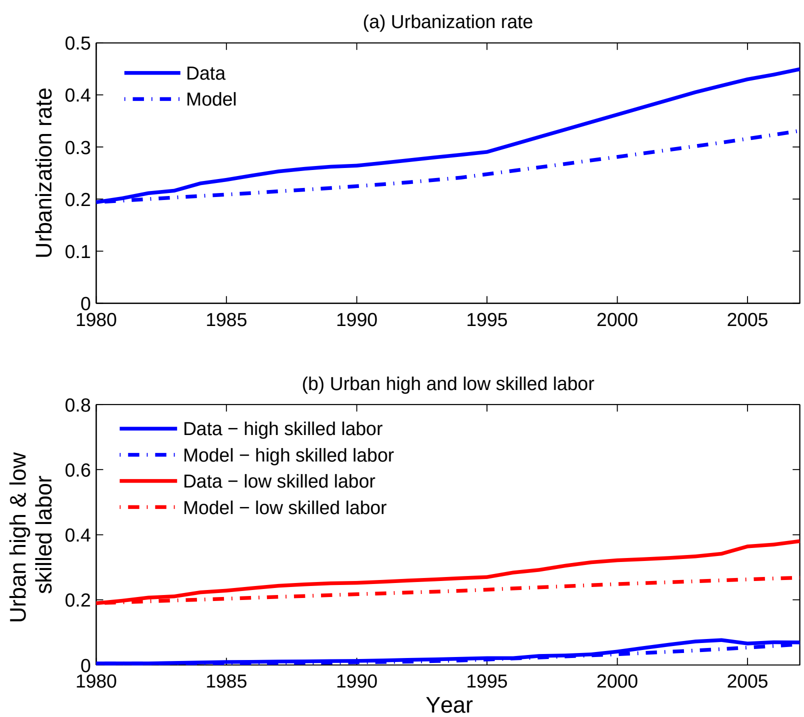


Figure 7: BENCHMARK MODEL - URBANIZATION RATE AND LABOR SHARE

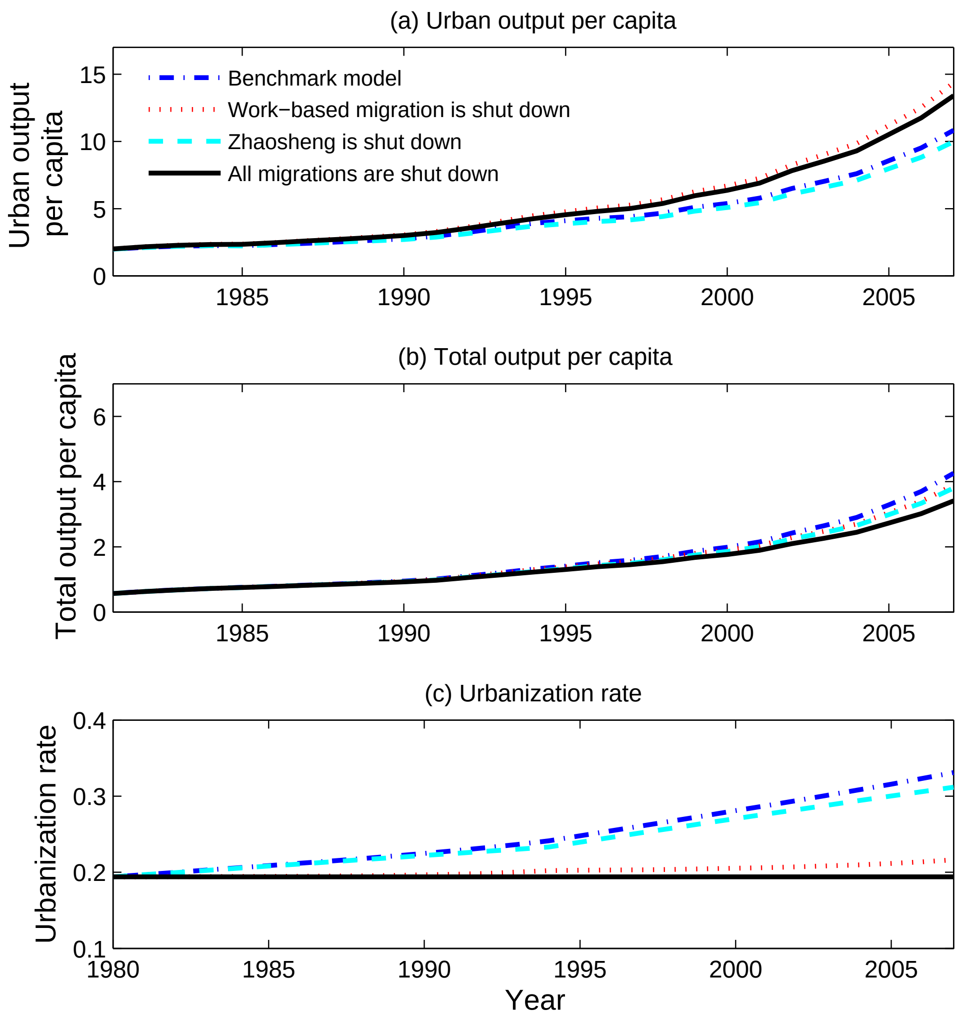


Figure 8: DECOMPOSITION FOR MIGRATION CHANNELS

Appendix (Not Intended for Publication)

This appendix is divided into three parts: Appendix A summarizes institutional background, Appendix B presents mathematical proofs, and Appendix C provides data sources and calibration details.

Appendix A: Institutional background

China implemented the *hukou* system to solve a serious problem of the so-called "blind flows" (of rural workers into cities) in the early 1950s. The important role of *zhaosheng* in China's development is indeed due to the *hukou* regulation. Therefore, here we briefly review the institutional background of the *hukou* system, its reforms and *zhaosheng*.

A.1 The *hukou* system and its reforms

China introduced the *hukou* regulation system in 1958. A citizen's *hukou* contained two parts: *Hukou suozaidi* (the place of *hukou* registration) and *Hukou leibie* (the type of *hukou* registration: "agricultural" and "non-agricultural"). *Hukou suozaidi* was a person's presumed regular residence, such as cities, towns, villages or state farms. Everyone was required to register in one and only one place of residence. This determined the place where the person received benefits and social welfare. *Hukou leibie* was mainly used to determine a person's entitlements to state-subsidized food grain (commodity grain). A citizen with "non-agricultural" *hukou* status would lose the right to rent land and the right to inherit the land that his parents rented. The above two classifications were different. Urban areas contained both agricultural and non-agricultural *hukou* populations. People with non-agricultural *hukou* may live in both urban and rural areas. Therefore, a "formal urban *hukou* holder" refers to an urban and non-agricultural *hukou* holder. Before 1997, *hukou* registration place and type were inherited from a person's mother. Since 1997, they can be inherited from a person's mother or father.

Under the *hukou* system, *nongzhuanfei*, changing from agriculture to non-agriculture, was the only method to obtain an official urban *hukou*. The regular channels of *nongzhuanfei* included (i) recruitment by a state-owned enterprise (*zhaogong*), (ii) promotion to a senior administrative job (*zhaogan*) and (iii) enrolment in an institution of higher education (*zhaosheng*). Official rural-urban migration involved both changes in *hukou* registration place and in registration type. To complete the *nongzhuanfei* process, a person had to satisfy both the migration requirements and obtain a quota, which was controlled by the central government at approximately 0.15–0.2 percent of the non-agricultural *hukou* population in each area.

The *hukou* system not only regulated internal population movement but also governed the social and economic aspects of a citizen's life. In rural areas, which were organized through the commune system, all rural residents had to participate in agricultural production to receive food rations for their households. In urban areas, under the pre-reform periods, state governments essentially controlled job assignments, grain rations, education for children, health benefits and housing purchase rights. There were few jobs outside the state-owned enterprises. Without an urban *hukou*, people were not able to survive. Therefore, people in China lost their freedom of migration before the economic reform.

A series of economic reforms began in the late 1970s. Since then, the increasingly market-oriented economy, the rural-urban income gap and the demand for cheap labor from rural areas have greatly increased informal rural-urban migrants, which has led to the continual relaxation of the *hukou* system.

The interesting part of the *hukou* reforms was the relaxation of migration regulations for the general public. For example, state governments implemented a new type of urban *hukou* with "self-supplied food grain" in 1984. In addition, due to the demands of economic development, several state governments introduced the blue-stamp urban *hukou* in the early 1990s to attract professional workers and investors. The blue-stamp *hukou* required an urban infrastructure construction fee for any newcomer, ranging from a few thousand to some fifty thousand yuan. It allowed people to obtain a temporary urban *hukou*. However, the blue-stamp *hukou* was different from the official urban *hukou* obtained through *nongzhuanfei* in that it provided limited rights and obligations and was only valid in that city. The blue-stamp *hukou* could be upgraded to an official urban *hukou* under certain conditions and after some years.

In 2005, the deputy minister of public security stated that eleven provinces had begun or would soon begin to implement a unified urban-rural household registration system, removing the distinctions between agricultural and non-agricultural *hukou* types. An updated statement in 2007 repeated the same points and included a list of twelve provincial-level units. In the statement of 2014, the government further adjusted migration policies according to the size of a city. The ultimate aim of the *hukou* reforms is to establish a unified *hukou* registration system, abolish the regulations of migration and provide social benefits to all residents.

A.2 *Zhaosheng*

As mentioned in the last section, *zhaosheng* was one of the official and important channels for rural people, especially for rural youths, to move to cities. Here we briefly review the procedure of rural students to obtain an urban *hukou* through the channel of *zhaosheng*. Educational reforms in the late 1990s of China are also discussed.

To obtain formal urban *hukou* through *zhaosheng*, rural students in China must pass *gaokao* to be admitted to universities. The *gaokao* system was established at the beginning of the 1950s, abolished during the Cultural Revolution, and restored in 1977. Because of the scarcity of education resources, acceptance rates were very low, especially in the 1980s. As most universities and colleges in China were located in urban areas, they were considered as urban collective units. Once a rural student was admitted to a university or a college, upon starting his freshman year, the student could voluntarily move his *hukou* to the school and obtain an urban *hukou*. However, such urban *hukou* was temporary. The youth's *hukou* would be removed from the school after graduation and moved to his work unit if he successfully found a job; otherwise, he was required to move his *hukou* back to his hometown. During the years of the GJA policy (1951–1994), a college graduate was assigned a stable government job, usually in an urban work unit. His *hukou* was thus transferred to the urban work unit immediately after graduation, allowing him to keep an urban *hukou* henceforth. However, after the termination of the GJA policy, governmental jobs for college graduates were no longer guaranteed. More specifically, the reform of the GJA policy started in 1989, but it was officially ended in 1996. Tibet, which abolished the system in 2007, was the last place to terminate the distribution system of graduation. With the abolishment of the GJA policy, if a college graduate failed to find an urban job upon graduation, he could temporarily assign his *hukou* to the collective joint household of a personal exchange center if he was still searching for an urban job or moved his *hukou* back to his hometown. Therefore, under China's *hukou* system, entering college through the *gaokao* provided a formal channel for rural-urban migration, and it provided rural youths with greater upward mobility in society.

A.3 Education reforms

Since 1996, China has introduced a series of educational reforms, notably the college education expansion and increases in college tuition. The expansion policy has provided broader access to students from rural areas. For example, Gou (2006) shows that, from 1996 to 2005, the number of rural students admitted to colleges have increased from 507,500 to 3,038,100 people, while the number of urban students have increased from 520,300 to 2,692,700 people. The admission rate for rural students also increased from 18.7% in 1989 to 62.9% in 2005. Meanwhile, the rise in college tuition has placed a heavier burden on rural parents for children's college education. Researchers have noticed the phenomenon that fewer and fewer rural students were admitted to top universities, and the rural-urban disparity in access to top universities has been discussed in studies such as Li (2007) and Qiao (2010). College expansion, increases in college tuition and rural-urban inequality in access to top schools all affect parents' education-based migration decisions. Therefore, our model is designed to capture the main spirits of these educational reforms in China.

Besides, the regional inequality in the distribution of educational resource is also observed in China. Wu and Luo (2012) point out that about two-thirds of higher education institutions either affiliated directly under the Ministry of Education or supported by the 211 Project are located in province capital cities, and few higher education institutions are located in cities that are smaller than prefecture level. Hu and Vargas (2015) found that college location is significantly associated with salary levels after controlling for job locations in China. Based on these facts, we thus assume that urban areas are the only places for higher education, i.e. college education is not available in rural areas.

Appendix B: Mathematical appendix

Proof of Proposition 1.

Denote c_U^j as the consumption of children if they are sent to an urban area and c_R^j as the consumption of children if they are kept in a rural area. From (10) and (11) we have:

$$\begin{aligned} c_U^j &= \gamma_H w_H h + \gamma_L w_L + (1-\gamma_H-\gamma_L)\, w_R - \mathbf{I}^k(X)(1-\gamma_H-\gamma_L)(X+\sigma_e) - \phi^j, && (23) \\ c_R^j &= (1-\pi)\, w_R + \pi (w_L - \sigma_w) - \mathbf{I}^k(X)(X+\sigma_e) - \phi^j. && (24) \end{aligned}$$

By subtracting (24) from (23) and rearranging terms, under Condition S, we have:

$$\begin{aligned} c_U^j - c_R^j &= \gamma_H w_H h + \gamma_L w_L + (\pi-\gamma_H-\gamma_L)\, w_R + \mathbf{I}^k(X)(\gamma_H+\gamma_L)(X+\sigma_e) - \pi(w_L-\sigma_w) \\ &= \gamma_H w_H h + \gamma_L w_L - \pi w_L + (\pi-\gamma_H-\gamma_L)\, w_R + \mathbf{I}^k(X)(\gamma_H+\gamma_L)(X+\sigma_e) + \pi\sigma_w \\ &> (\gamma_H+\gamma_L-\pi)\left[w_L(n_s) - w_R\right] + \mathbf{I}^k(X)(\gamma_H+\gamma_L)(X+\sigma_e) + \pi\sigma_w \\ &> 0. \end{aligned}$$

Because $u(\cdot)$ is strictly increasing and strictly concave, we have:

$$u\left(c_U^j\right) > u\left(c_R^j\right).$$

Thus, Assumption 1 and Condition S together guarantee that $\mathbb{E}_{\mathrm{X}}\left(u\left(c_U^j\right) - u\left(c_R^j\right)\right) > 0$ for all $x^k \in (0, x_{\max}]$. ■

Proof of Proposition 2.

Denote $u_{c_U}^j = u_c\left(c_U^j\right)$ as the marginal utility. We compute:

$$\begin{aligned} \frac{d\Delta^i\left(\mathbf{I}^k, x^j\right)}{dx^j} &= -u_{c_U}^j < 0 \\ \frac{d\Delta^i\left(\mathbf{I}^k, x^j\right)}{d\gamma_H} &= \beta\mathbb{E}_{\mathrm{X}}\left\{u_{c_U}^j\left[(w_H - w_R) + \mathbf{I}^k(X)(X+\sigma_e)\right]\right\} > 0 \\ \frac{d\Delta^i\left(\mathbf{I}^k, x^j\right)}{d\gamma_L} &= \beta\mathbb{E}_{\mathrm{X}}\left\{u_{c_U}^j\left[(w_L - w_R) + \mathbf{I}^k(X)(X+\sigma_e)\right]\right\} > 0 \\ \frac{d\Delta^i\left(\mathbf{I}^k, x^j\right)}{d\pi} &= \beta\mathbb{E}_{\mathrm{X}} u_{c_R}^j\left[w_R - (w_L - \sigma_w) - \mathbf{I}^k(X)(X+\sigma_e)\right] < 0. \end{aligned}$$

Since x^j is decreasing in a and z^j, but increasing in b, the results follow.■

Proof of Proposition 3.

Under Condition S, $c_U^j > c_R^j$ and $u_{c_U}^j < u_{c_R}^j$. Define $\Lambda \equiv \left[u_{c_U}^j (1-\gamma_H-\gamma_L) - u_{c_R}^j (1-\pi)\right] \mathbf{I}^k(X)$, then we have:

$$\begin{aligned}
\Lambda &\equiv \left[u_{c_U}^j (1-\gamma_H-\gamma_L) - u_{c_R}^j (1-\pi)\right] \mathbf{I}^k(X) \\
&< \left[u_{c_U}^j (1-\gamma_H-\gamma_L) - u_{c_U}^j (1-\pi)\right] \mathbf{I}^k(X) \\
&= u_{c_U}^j \underbrace{(\pi-\gamma_H-\gamma_L)}_{(-)} \mathbf{I}^k(X) \\
&< 0.
\end{aligned}$$

Therefore, if $-\beta \mathbb{E}_X \Lambda > u_{c_U}^i$, $\frac{d\Delta^i(\mathbf{I}^k, x^j)}{d\sigma_e} > 0$. ■

Appendix C: Data and Calibration

1. Population

(1) Rural and urban population

Table 1-4 of the *China Population and Employment Statistical Yearbook* 2010 reported the fraction of rural (urban) population as a percentage of total population in China during the 1952–2009 period. We directly borrow the time series data from 1980 to 2007 for our rural and urban population (N_R and N_U) data. The data in the calibration for regime 1 are the simple average of 1980–1994; for regime 2, the simple average of 1995–2007.

(2) High-skilled and low-skilled workers

The *China labor Statistical Yearbook* reported the educational attainment composition of urban employment (as a percentage of total urban employment). Thus, workers whose educational attainment is categorized as college and above are defined as high-skilled workers. However, urban data are only available for 2002–2007. Thus, we first use 2002–2007 data to compute an urban to national ratio (a ratio of educational attainment composition of urban employment to that of the entire country). The ratio is approximately 2.457. Second, for the years 1982, 1990, 1995–1999, and 2001, the fraction of high-skilled workers as a percentage of total urban employment is computed using nationwide data and is adjusted by the urban to nationwide ratio. The national data for 1996–1999 and 2001–2007 are also from the *China labor Statistical Yearbook.* The data for 1982 are from *1 Percent Sampling Tabulation on the 1982 Population Census of the People's Republic of China.* The data for 1990 are from the *China Population Statistical Yearbook* 1994. The data for 1995 are available in *1995 China 1% Population Sampling Survey Data.* For the years 1980, 1985, and 2000, the educational attainment for total population in Barro and Lee (2001) is adjusted by the urban to nationwide ratio to obtain N_H/N_U. Third, we interpolate data for the years for which no data are available. Finally, the fraction of high-skilled workers as a percentage of total urban employment is multiplied by N_U to obtain N_H. Then, N_L is the difference between N_U and N_H.

The data in the calibration for regime 1 reflect the simple average of 1980–1994, and for regime 2, the average is for 1995–2007.

(3) Rural to urban migration flows

There is no available nationwide survey on rural to urban migration for the periods of China that we examine. Here, we use changes in urban population as a proxy for total rural to urban migrants. We are aware that changes in urban population is equal to the amount of rural-urban migrants only if births and deaths in urban areas are net out exactly. However, as shown in Table C.1, we find that the net birth rates (net of death) in urban and rural areas are quite stable during the periods that we examine. Since there is no nationwide available data on rural-urban migration, we use changes in urban population as a proxy. In addition, we believe that the actual rural-urban migration could be larger than our proxy because the birth rates and mortality rates are both higher in rural areas than those in urban areas.

The total number of rural to urban migrants is then divided by the stock of rural population to obtain the flow of migrants (as a percentage of the rural population). In the calibration, we take the simple average on the flow of migrants for 1981–1994 to be the flow of migrants in the first regime. The second regime is the average of the 1995–2007 flows. Finally, the average flows of migrants are multiplied by the working-related and studying or training reasons (the average of 1985 and 2000) to obtain the probabilities of working migration and *zhaosheng* flow, respectively.

(4) Migration reasons

The *10 Percent Sampling Tabulation on the 1990 Population Census of the People's Republic of China* reported the number of immigrants by type of usual residence and cause of migration for 1985. We choose "the number of immigrants from town and county of this province" and "the number of immigrants from town and county of other provinces" to be rural to urban migration in 1985. Then, the fraction of migrants due to each reason as a percentage of total rural to urban migration is computed. The *Tabulation on the 2000 Population Census of the People's Republic of China* only reported the number of emigrants and the reasons for emigration. We thus choose the number of emigrants from towns and counties to represent rural to urban migration in 2000. Then, the fraction of migrants for each reason as a percentage of total rural to urban migration is computed. Finally, we categorize migration due to job transfers, job assignments, and work or business as working-related reasons. The migration due to studying or training is categorized as migration via *zhaosheng*.

2. Human capital

Table C.2 summarizes average years of schooling for the group of college and above and the group of less than college. The urban employment by education in 1995 data are from the *China Statistical Yearbook* 1998. The 2002 and 2009 data are from the *China Labor Statistical Yearbook* 2002 and 2009, respectively. We further assume that the years of schooling for graduate school is equal to 18 years, 16 years for college, 14 years for junior college, 12 years for senior high, 9 years for junior high, 6 years for primary school, and 1 year for semi-illiterate or illiterate. Then, weighted average years of schooling for college and above and less than college are computed. Table C.3 provides the average years of schooling for 1981, 1988, 1995, and 2002. For years without data, they are computed by backward extrapolation based on 1995, 2002, and 2009 data. In the calibration, years of schooling in regime 1 (8.02 and 14.10) is the average of 1981 and 1988 and regime 2 (8.95 and 14.52) is the average of 1995 and 2002.

To compute the human capital possessed by high-skilled workers relative to low-skilled workers, the Mincerian method is employed. The education returns coefficients in China reported by Zhang et al. (2005) are 0.0497 and 0.0836 for 1980–1994 and 1995–2007, respectively. Thus, the human capital in regime 1 is equal to $\frac{e^{0.0479*14.1}}{e^{0.0479*8.02}}$. The human capital in regime 2 is $\frac{e^{0.0835*14.52}}{e^{0.0835*8.95}}$.

3. Urban employment rate

In the model, $\gamma_H + \gamma_L$ refers to the employment rate of college graduates who migrated from rural areas. However, no data are available. Thus, we use the urban employment rate as a proxy. Urban employment rate is computed by using the number of urban working or employed workers divided by the sum of the number of urban working or employed workers and the number of workers who are waiting for a job or unemployed. The urban employment rate of 1995 is computed using 1995 CHIP urban individual data; the value of 2002 is computed using 2002 CHIP urban individual income, consumption, and employment data; and the value of 2007 is computed using 2007 CHIP (or RUMiC 2008). The average of them is the urban employment rate in the calibration.

China has introduced lots of reforms in the public sector in the late 1990s. Many workers were "off post" or *xiagang* during the reforms. These workers still had their *hukou* with their employers (and hence stay in cities) as only by doing this they could obtain compensations for the loss of their jobs. *Xiagang* workers are usually low-skilled workers, senior in age and difficult to find a job again. See Lee (2000) for more information on the characteristics of *xiagang* workers. In the calibration, we have matched the N_H/N_L data series and considered *xiagang* when computing urban employment rate. Therefore, the employment composition change due to the reforms is being taken care of.

4. Labor income shares

Bai and Qian (2010) reported the sectoral labor share of GDP for the 1978–2004 period and the sectoral composition of value-added at factor cost in China. We thus compute a time series of labor income share in urban areas by assuming that industry, construction, and service belong to the urban sector. The labor income share is weighted by the corresponding sectoral composition of value-added at factor cost. In the calibration, the labor income share is the average of 1980–1994 in regime 1 and of 1995–2004 in regime 2.

5. Rural per capita income

The *China Statistical Yearbook* 2011 reported rural real per capita income from 1978 to 2011.

However, during the period before 1990, only data for 1978, 1980, and 1985 are available. We thus use interpolation to compute rural real income per capita for 1981–1984 and 1986–1989. Then, the rural real income per capita of 2007 is normalized to one. The rural real income per capita for other years is adjusted accordingly. In the calibration, the rural income of regime 1 is the average of rural real income per capita during the 1980–1994 period. The rural income of regime 2 is the average of 1995–2007.

6. Skill premium

Zhang et al. (2005) estimate the skill premium for China during the 1988–2001 period, while Ge and Yang (2014) estimate it for 1992–2007. Using the ratio of the skill premium in Zhang et al. (2005) to that in Ge and Yang (2014), we construct a time series for the skill premium for 1988–2007 based on Zhang et al. (2005). Furthermore, Lee (1999) estimates the skill premium for China in 1980 and 1988. However, the estimate in Lee (1999) is higher than that reported by others because the estimate is based on a survey of SOEs. Therefore, we first compute the growth rate of the skill premium from 1980 to 1988 in Lee (1999). Then, using the estimate of the skill premium in 1988 in Zhang et al. (2005) and the growth rate computed from Lee (1999), the skill premium in 1980 is obtained. Finally, curve fitting with polynomial 3 is used to construct a series for the skill premium from 1980 to 2007. In the calibration, the skill premium of regime 1 is the average of 1980–1994. The skill premium of regime 2 is the average of 1995–2007.

7. Urban premium

The urban premium is defined as the ratio of the low-skilled wage to the rural wage. The *China Statistical Yearbook 2011* also reported urban real income per capita for 1978–2011. Thus, we are able to compute a ratio of urban to rural income per capita. Because urban income per capita is a weighted average of the high-skilled wage and the low-skilled wage, we are now able to compute the urban premium using the skill premium data, urban-rural income per capita ratios, and the ratios of high- to low-skilled worker stocks. However, during the period before 1990, data are only available for 1978, 1980, and 1985. We thus use interpolation to compute the urban premiums for 1981–1984 and 1986–1989.

8. Sharp parameter of the Pareto distribution

Chinese Household Income Project (CHIP) 1995 and 2002 reports rural household net income data. We first compute the mean of the rural household net income for each year. Then the rural household net income is divided by the average number of rural household members to obtain the average of rural individual income. Similarly, we compute the standard deviation and the variance of the rural individual income. Finally, using the formulas for the mean and variance of Pareto distribution, we are able to back out the value of θ, which is roughly equal to 2.5138 for 1991-2002. We thus set θ to 2.5. Our estimated value is close to the value (2.11) reported by Feenberg and Poterba (1993) for the United State during the period from 1950 to 1990. The average number of rural household members is roughly equal to 4. The data on the number of rural household members is also from CHIP 1995 and 2002.

9. Elasticity of substitution between high- and low-skilled labor

The estimated value of the elasticity of substitution between high- and low-skilled labor in the production function $1/(1-\rho)$ for developed countries is between 1 and 3. For example, Autor, Katz and Krueger (1998), Acemoglu (2003), and Ciccone and Peri (2005). However, the elasticity of substitution between high- and low-skilled labor in developing countries are usually higher. For example, Toh and Tat (2012) estimate that the value for Singapore is 4.249. Te Velde and Morrissey (2004) use data from Singapore, Hong Kong, Korea, the Philippines and Thailand and obtained a value of 2.78. The results in Gindling and Sun (2002) imply that the value in Taiwan is between 2.3 and 7.4. We choose the value to be 3, the maximum value in developed countries and within the estimated range for developing countries.

10. Child-rearing cost

Zhu and Zhang (1996) estimated that the average child-rearing cost in rural villages in Xianyang, which is located in the Shaanxi province of China, was approximately 17.4 percent of family income for a child aged 0–16 in 1995. Since national-wide survey on child-rearing costs is not available for rural China, we adopt the value in Zhu and Zhang (1996) to be our child-rearing cost.

11. Work-based migration cost

CHIP 2002 rural-urban migrant individual data provides information on the expenditures occurred in the first month when migrant workers arrived at the city. In the calculation of the work-based migration cost, food and housing costs are counted as regular costs, while city expansion fee, certification fee and others are considered as one-time cost. Our work-based migration cost is thus the sum of the above costs, adjusted for model periods and expressed as a percentage of rural household income. Rural household income is computed by multiplying rural real per capita income by the average number of rural household members. Rural real per capita income is obtained from the *China Statistical Yearbook* 2011 and the average number of rural household members is from the *China Rural Statistical Yearbook*.

Tombe and Zhu (2015) found a high moving cost for Chinese migrant workers, roughly equal to the annual income of a rural worker. For the United States, the estimated migration costs are between one-half and two-thirds of average annual household income, such as Bayer and Juessen (2012) and Lkhagvasuren (2014). Our work-based migration cost is consistent with the literature.

12. Education-based migration cost

He and Dong (2007) reports the annual cost of food and dormitory for a college student in 1996-2005. It is about 63.78% of annual rural household income. We use the estimate in He and Dong (2007) and assume that a student stays in college for four years to compute our education-based migration cost. It is adjusted by model periods.

13. Direct college cost

The direct college cost b is computed based on Urban Household Survey (UHS) 2007 and 2008. Because college education was almost free of charge before 1990, the value of b in regime 1 includes stationary, materials and textbooks only, while the value of b in regime 2 includes not only stationary, materials and textbooks but also college tuitions. College tuition as a percentage of rural household income ranges from 22.8% in UHS to 35.2% in CHIP. We therefore assume college tuition is 30% of rural household income in the computation of b in regime 2. Then, the value of b equals 0.48% and 5.28% of rural household income in regimes 1 and 2, respectively.

14. Data in figures

(1) Urban output

The computed data for urban output is calculated by the urban production function. Using the calibrated parameters, the calibrated time series of urban TFP, the time series data of high-skilled workers, and the time series data of low-skilled workers, we are able to obtain the computed data for urban output. The computed data for urban per capita output is the computed data for urban output divided by the time series data for high- and low-skilled workers.

(2) Rural output

The computed data for rural output is obtained from the rural production function. Because we have time series data of rural per capita income (2007 is normalized to one) and the stock of the rural population, we are able to obtain the computed data for rural output.

(3) Total output

The computed total output is the sum of the computed data for urban output and rural output.

Appendix References

Acemoglu, Daron, 2003, "Cross-country inequality trends," *The Economic Journal*, 113, F121-F149.

Autor, David H., Lawrence F. Katz, and Alan B. Krueger, 1998, "Computing inequality: have computers changed the labor market?" *The Quarterly Journal of Economics*, 113(4), pp. 1169-1213.

Bai, Chong-En and Zhenjie Qian, 2010, "The factor income distribution in China: 1978-2007," *China Economic Review*, 21, pp. 650-670.

Barro, Robert J. and Jong-Wha Lee, 2001, "International data on educational attainment updates and implications," *Oxford Economic Papers* 53(3), pp.541-563.

Bayer, Christian and Falko Juessen, 2012, "On the dynamics of interstate migration: Migration costs and self-selection," *Review of Economic Dynamics*, 15, pp. 377-401.

Ciccone, Antonio and Giovanni 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.

Feenberg, Daniel R. and James M. Poterba, 1993, “Income inequality and the incomes of very high-income taxpayers: Evidence from tax returns,” in *Tax policy and the economy*, Vol. 7, edited by James Poterba, NBER, MIT Press.

Ge, Suqin and Dennis Tao Yang, 2014, “Changes in China’s wage structure,” *Journal of the European Economic Association*, 12(2), pp. 300-336.

Gindling, T.H. and Way Sun, 2002, “Higher education planning and the wages of workers with higher education in Taiwan,” *Economics of Education Review*, 21(2), pp. 153-169.

Gou, Renmin, 2006, “Examining equality in higher education from the perspective of rural-urban access to higher education,” *Research in Educational Development*, 5, pp. 29-31. (in Chinese).

He, Zhiyun and Naihan Dong, 2007, “A study of family consumption of higher education in Zhejiang Province based on the construction of a harmonious society – A comparative analysis with other seven provinces and municipalities,” *China Hihger Education Research*, 5, pp. 23-26.

Hu, Anning and Nicholas Vargas, 2015, Economic consequences of horizontal stratification in postsecondary education: evidence from urban China,” *Higher Education*, 70(3), pp. 337–358.

Lee, Hong Yung, 2000, “Xiagang, the Chinese style of laying off workers,” *Asian Survey*, 40(6), pp. 914-937.

Lee, Young, 1999, “Wages and employment in China’s SOEs, 1980-1994: Corporatization, market development, and insider forces,” *Journal of Comparative Economics*, 27, pp. 702-729.

Li, Wenli, 2007, “Family background, financial constraint and higher education attendance in China," *Economics of Education Review*, 26, pp. 725-735.

Lkhagvasuren, Damba, 2014, “Education, mobility and the college wage premium,” *European Economic Review*, 67, pp. 159-173.

Qiao, Jinzhong, 2010, “On the rural-urban disparity in access to higher education opportunities in China,” *Chinese Education and Society*, 43(4), pp. 22-31.

Te Velde, Dirk and Oliver Morrissey, 2004, “Foreign direct investment, skills and wage inequality in East Asia,” *Journal of the Asia Pacific Economy*, 9(3), pp. 348-369.

Toh, Ruby and Hui Weng Tat, 2012, “Trade liberalization, labor demand shifts and earnings inequality in Singapore,” *Review of Urban and Regional Development Studies*, 24(3), pp. 65-82.

Tombe, Trevor and Xiaodong Zhu, 2015, “Trade, migration and regional income differences: evidence from China,” Working paper.

Wu, Hongjun and Yingzi Luo, 2012, “Vertical distribution structure of colleges and universities in China: The phenomenon of assembling in political centers,” *Modern Education Management* (Chinese), 5, pp. 11–16. (in Chinese).

Zhang, Junsen, Yaohui Zhao, Albert Park, and Xiaoqing Song, 2005, “Economic returns to schooling in urban China 1988 to 2001,” *Journal of Comparative Economics*, 33, pp. 730-752.

Zhu, Chuzhu and Yougan Zhang, 1996, “Cost and benefit of rural children in the Xianyang province of China,” *Population and Economics* 5, pp. 13-22 (in Chinese).

Table C.1: NET BIRTH RATES IN CHINA

Unit: ‰

Year	Urban areas			Rural areas		
	Birth rate	Mortality rate	Net birth rate	Birth rate	Mortality rate	Net birth rate
1980	14.17	5.48	8.69	18.82	6.47	12.35
1981	16.45	5.14	11.31	21.55	6.53	15.02
1982	18.24	5.28	12.96	21.97	7.00	14.97
1983	15.99	5.92	10.07	19.89	7.69	12.20
1984	15.00	5.86	9.14	17.90	6.73	11.17
1985	14.02	5.96	8.06	19.17	6.66	12.51
1986	17.39	5.75	11.64	21.94	6.74	15.20
1987	-	-	-	-	-	-
1988	-	-	-	-	-	-
1989	16.73	5.78	10.95	23.27	6.81	16.46
1990	16.14	5.71	10.43	22.80	7.01	15.79
1991	15.49	5.50	9.99	21.17	7.13	14.04
1992	15.47	5.77	9.70	19.09	6.91	12.18
1993	15.37	5.99	9.38	19.06	6.89	12.17
1994	15.13	5.53	9.60	18.84	6.80	12.04
1995	14.76	5.53	9.23	18.08	6.99	11.09
1996	14.47	5.65	8.82	18.02	6.94	11.08
1997	14.52	5.58	8.94	17.43	6.90	10.53
1998	13.67	5.31	8.36	17.05	7.01	10.04
1999	13.18	5.51	7.67	16.13	6.88	9.25

Source: *China Statistical Yearbook* 1990 and 2000. Data for 1987, 1988, and years after 1999 are not available.

Table C.2: URBAN EMPLOYMENT BY EDUCATION

Education Attainment	Years of schooling	1995	2002	2009
College or above		10.6%	15.9%	16.2%
Graduate	18		0.3%	0.5%
College	16		4.4%	5.8%
Junior college	14		11.2%	9.9%
Average years of schooling college or above			14.63	14.84
Below college		89.4%	84.1%	83.8%
Senior high	12	24.6%	26.6%	20.7%
Junior high	9	39.7%	41.0%	45.6%
Primary	6	20.4%	13.6%	15.4%
Semi-illiterate or illiterate	1	4.7%	2.9%	2.1%
Average years of schooling below college		8.72	9.19	8.99

Source: *China Statistical Yearbook* and *China Labour Statistical Yearbook*.

Table C.3: AVERAGE YEARS OF SCHOOLING

Year	*Below college*	*College or above*
1981	7.79*	14.00*
1988	8.25*	14.21*
1995	8.72	14.42*
2002	9.19	14.63
Average: 1981-2002	8.49	14.31
Average: 1981 and 1988	8.02	14.10
Average: 1995 and 2002	8.95	14.52

Note: * denotes those numbers are obtained from backward extrapolation using on 1995, 2002 and 2009 data.