Educational Choice, Rural-urban Migration and Economic Development: The Role of *Zhaosheng* in China*

Pei-Ju Liao† Ping Wang‡ Yin-Chi Wang§ Chong Kee Yip¶

This draft: April 30, 2015

Abstract

Observing the rapid economic growth accompanied by fast urbanization in China, we examine the contribution of migration to China's development with a focus on education-based migration. In addition to the conventional work-based channel, we construct a dynamic equilibrium model of migration decision with educational choice to highlight the education-based channel of migration. We then calibrate our model to fit the data from China over 1980-2007. We find that the effects of education-based migration are almost comparable to that of work-based migration. There are rich interactions between education- and work-based migration on urban output and employment shares. We further conduct various policy experiments focusing on migration regulations to assess their quantitative significance.

JEL Classification: O15, O53, R23, R28.
Keywords: Educational choice, rural-urban migration, urbanization and development.

---

*Acknowledgment: We are grateful to the comments from the participants at the Macro reading group at Washington University in St. Louis, Asian Meeting of the Econometric Society, SAET Conference on Current Trends in Economics, the IEAS and the CCER joint workshop, the seminars at the IEAS and the National Graduate Institute for Policy Studies (GRIPS), the lunch meeting group at the Chineses University of Hong Kong, and the Conference on Chinese Economy at Tsinghua University. Travel supports from the Chinese University of Hong Kong, Tsinghua University and the Center for Dynamic Economics at the Washington University are greatly appreciated. We want to express special thanks to B. Ravikumar for his comments when this project was at the early stage, and Suqin Ge and Dennis Tao Yang for providing data on skill premium in China. All errors remain to be ours.

†Institute of Economics, Academia Sinica. E-mail: pjliao@econ.sinica.edu.tw
‡Washington University in St. Louis & NBER. E-mail: pingwang@wustl.edu
§Chinese University of Hong Kong. Email: ycwang@cuhk.edu.hk
¶Chinese University of Hong Kong. Email: chongkeeyip@cuhk.edu.hk

1 Introduction

Over the past three decades, we have observed rapid economic growth and urbanization in China. While its annual per capita real GDP growth over our sample period of 1980 to 2007 is about 6.0 percent, the comparable figure since Deng Xiao-Ping's Southern Trip in 1992 is 7.6 percent. Over the same sample period, 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 (see Figure 1).

China's fast development process is accompanied by large rural to urban flows and sharp increases in tertiary education shares and years of schooling. Over our sample period, the migration flows (proxied by changes in urban population) over rural population have almost quadruppled, from 0.5 to 1.9 percent, despite a barrier to migration due to the household registration system – *hukou*.[1] Even more drastically, the urban tertiary education employment shares have increased seven-fold, from 2.2 to 15.4 percent. Figure 2 plots these two series. Thus, while conventional rural to urban migration emphasizes on employment related migration, the above observations suggest another potentially important channel of rural to urban migration via higher education, which is referred as *zhaosheng*.

Zhaosheng is a specific channel that mitigates migration barriers. Particularly, students attended college by passing the National College Entrance Examination, *gaokao,* to migrate to cities. Before 1994, all college graduates were assigned jobs by the government before graduation, hereafter *government job assignment* (GJA), and could officially obtain an urban *hukou.*[2] Starting 1995, GJA was abolished and jobs are not guaranteed. Thus, the abolishment of GJA and migration regulations become two natural policy expertiments to evaluate their roles played in the development process of China.

In this paper, we explore not only the conventionally work-based migration but also the education-based migration that is unique in the Chinese economy. We construct a dynamic equilibrium model with intergenerational migration decision based on education that enables human capital enhancement and mobility across generations. We then perform quantitative analysis to investigate the effects of rural-urban migration on various key economic variables. We further decompose these

[1] Under the *hukou* registration system started from the 1950s, it was difficult to obtain an urban *hukou.* Different from other countries, China used the *hukou* registration for government to regulate its citizens' life, including the potential opportunities for jobs, children's education, as well as other benefits.

[2] Wu and Treiman (2004) show that education was one of the main determinants of the upward mobility in China.

effects into *zhaosheng* and work-based migration. Finally, we conduct experiments concerning the abolishment of GJA and migration regulations.

The main findings of our paper are as follows. The contributions of *zhaosheng* and work-based migration are almost comparable, accounting for 3.4 and. 4.9 percent, respectively, of changes in per capita output over our sample period. Interestingly, there are rich interactions between *zhaosheng* and work-based migration under our dynamic general equilibrium setting. Such interactions in terms of total output per capita, urban output share and urban employment share are negative as a result of reallocation of migration flows from one channel to another. On the contrary, the interaction is positive on skilled labor employment, which is due to the rising skilled premium resulting from increased low-skilled migration. On policy experiments, we find that the abolishment of the GJA is not too harmful, whereas the reduction in migration regulations can encourage disproportionately high-skilled migration, that boosting up urban output per capita.

Related Literature

The research on rural-urban migration is pioneered by the classic work Todaro (1969) and Harris and Todaro (1970). Since their work, from the macro perspectives, researchers have attempted to understand the contributions of rural-urban migration to the development process of a country, the urban unemployment problem, and income inequality across urban and rural regions. On the micro level, researchers have managed to understand the motivations and the determinants of individuals' migration decision and the welfare of migrant workers and their descendants. Stark (1991) pointed out that lots of migratory phenomena would not have occurred if the financial markets and labor markets are perfect and complete. He further hypothesized that migrants functioning as financial intermediaries to overcome the credit constraints for rural households. Lucas (2004) proposed that, not only the better job opportunities in cities, but also the accumulation and spillovers of human capital and the associated returns attract rural workers to migrate to cities.

The *hukou* system and the strictly controlled population movement have made the urbanization process in China different from other developing countries. It is believed that the *hukou* system helped to keep labor force in the rural area and prevent shanty towns to emerge, resulting in a relative low labor cost compared to other countries. As China expands its participation in the world trade, the massive cheap labor from rural area then serves as the main driving force of growth of China (see Li, Li, Wu and Xiong, 2012 for a discussion on cheap Chinese labor).

Most of the studies on the issues of rural-urban migration in China are empirical studies. As for the welfare of migrant workers, Knight and Gunatilaka (2010) studied why rural migrant households settled in urban China are less happier than their rural counterpart. They found that the unhappiness arises from certain features of migrant conditions and their high aspirations for achievement influenced by their new reference groups. Regarding the literature on education and migration, previous studies typically focus on the issues whether higher education empowers rural workers a greater mobility to migrate to cities, and the educational opportunities faced by migrant children. Xia (2004) found that education drives rural workers away from land, while Hare (1999) did not reach the same conclusion. Zhao (1997) pointed out that the expected return to senior high school education in rural China is high as a senior high school degree will improve the access of rural workers to urban jobs. Wu and Treiman (2004) found that educational attainment and Chinese Communist Party membership are the most important determinants for rural-urban mobility, greatly affecting the chance of obtaining urban *hukou*. Liu (2008) had an interesting finding that human capital externalities in rural China lead to better local off-farm job opportunities, thus resulting in less migration. However, as migration decision and education are endogenous and closely related to the income level and the development process of a county, it calls for a deeper empirical analysis with proper instrument variables to examine this issue. Tan (2010) and Arnadottir (2012) studied the inequity and challenges in education faced by migrant children. They argued that migrant children should be granted the same access to education as children with permanent urban *hukou*. Zhao, Yu, Wang and Glauben (2014) examined the impacts of parental migration on left-behind children's performance in school. They found that parental migration does jeopardize a child's performance.

We now turn to the determinants for work-based migration. The higher wages and better career perspectives, education opportunities and social benefits are the "pull" factors in cities that attract rural migrants, while social network, land arrangement, and having elderly members in the households are plausible "retain"factors that keep rural workers away from cities. Rozelle, Taylor, and deBrauw (1999) examined the effects of labor leaving from the villiages on the agricultural productivity and studied whether the remittances from migrants exacerbate or compensate for the labor leaving villages in Hebei and Liaoning provinces in China. They found that less family members on farmwork due to migration depresses farm yields, but remittances from migrant family members can partly offset the negative effects of the lost labor. Zhao (2003) pointed out that networks indeed play an important role for rural-urban migration. Families with friends or relatives working in cities have a higher chance to migrate. Given household land arrangement, Yang (1997) developed a partial equilibrium static model on time allocation, studying how rural households optimally allocate time between farming, non-farm activities and migration decision. Using household survey data of Sichuan province for 1994 and 1995, Zhao (1999) found that the shortage of farmland and the abundance in household labor were the most important determinants of labor migration decision. He also found that, unwilling to give up farmland and unequal education opportunities faced by their children, most migrants are on transitory basis. Using the China General Social Survey data, Hu, Xu and Chen (2011) examined the household characteristics on the probability of being permanent migrants. They found that, more educated and experienced rural workers have a higher probability to be permanent migrants because of their higher pay and greater affordability of an urban *hukou*. Also, the migration destination for permanent migrants is usually within the same province of their hometown, while temporary migrants are more prone to move to remote cities. They also found that permanent migrant workers are more likely to obtain urban *hukou* of small or medium-sized cities. Land tenure uncertainty and the possibility of reallocation of land based on the on-farm family workers can impede households' incentives in investing in land and retard their incentives to migrate to cities. This could lead to underinvestment and inefficient use of land as well as misallocation of labor across sectors. Zhao (1999), Brandt, Huang, Li and Rozelle (2002), Mullan, Grosjean and Kontoleon (2011), Wang, Tong, Su, Wei and Tao (2011) discussed such problems and the dilemma faced by rural households.

In regard to the literature on rural-urban income gaps, Liu (2005), Lu and Song (2006), and Hertel and Zhai (2006) indicated that migration restrictions (due to the *hukou* system) are the major causes for rural-ruban income gap. Hu (2002) adopted a special general equilibrium model and found that the improving trade conditions and the increasing urban-rural mobility are possible to explain the widened income gap between the coastal and the hinterland area in China. Whalley and Zhang (2007) adopted numerical simulation method to study the role of the hukou restrictions in supporting the rural-urban inequalities in China. They found that eliminating hukou system not only reduces the degree of inequality, but also increases GDP per capita, output per worker and total output. Zhu (2002) found that income gaps between rural and urban areas significantly influence migration decisions with survey data of Hubei province.

The rest of the paper is organized as follows. Section 2 summaries the basis of China's *hukou* system, its reforms, and the migration channels in China. Section 3 presents the model and the theoretical analysis. The calibration strategy, the simulation, the decomposition analysis on education-and work-based migration different as well as policy experiments on migration regulations are provided in Section 4. Finally, Section 5 concludes this paper.

2 Institutional Background

In this section, we will provide institutional background for the *hukou* system and the relevant migration policies that relaxes the hukou regulations.

2.1 The Hukou System

China implemented the *hukou* system in 1958 in order to solve the serious problem of the "blind flows" (illegal rural labor to cities) in the early 1950s. The system was still the basis of China's registration system in the last decade.

The power of China's *hukou* system did not come from the regulation of migration itself. Instead, it was from its integration with other social and economic controls. In the rural areas, with the commune system, all rural residents in communes had to participate in agricultural production to receive food rations for their households. In the urban areas, most recruitment and job transfers were controlled by the state government. There were few jobs outside the state enterprises. *Danwei* was the basic administrative unit for most urban adults. It assigned most social services for its employees, such as the access to good jobs, grain rations, education for children, health benefits, and purchasing house. Without a work unit, it was difficult to survive in a city. Therefore, in the pre-reform period, it was hard for people to survive outside their *hukou* registration place without permissions.

A citizen's *hukou* was classified by two part: *hukou suozaidi* (the place of *hukou* registration) and *hukou leibie* (the type or status of *hukou* registration).[3]

- *Hukou suozaidi* was a person's presumed regular residence. Everyone was required to register in one and only one place of his residence. The common categories of place of *hukou* registration were cities, towns, villages, and state farms. It determined where a person would receive his benefits and social welfare.

[3]Before 1997, *hukou* registration place and status were inherited from a person's mother. Since 1997, they can be inherited from a person's mother or father.

- *Hukou leibie* referred to "agricultural" and "non-agricultural" *hukou*. It was 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 of land rental and the right of inheriting the land that his parents rented.

The above two classifications were different. Urban areas contained both agricultural and non-agricultural *hukou* population. Non-agricultural *hukou* population may exist in both urban and rural areas. *Hukou* registration place and status mattered because it determined, for example, access to good jobs, grain rations, education for children, health benefits, and purchasing house.[4] A formal rural-urban migration involved both a change in *hukou* residential place and in entitlement status. In order to complete the process, a person had to satisfy both the migration requirements and obtained a space under the quota control.

2.2 Migration Policies and Hukou Reform

In the late 1970s, because of the development of the market-oriented economy, more and more people stayed (i.e., actually worked and lived) outside their *hukou* registration place. They were so-called the *nongmingong*, or peasant workers. The increase in the mobility resulted in a series of *hukou* reforms since 1980. Due to the economic reform, people have become easier to survive outside their *hukou* registration place. The economic reform relaxed the administrative control, such as the abolition of the commune system, the introduction of the household responsibility system, and the erosion of the rigid *danwei*-based rationing system. Furthermore, the growing market-oriented economy demanded more cheap labor. Both the push and pull factors increased migrations from rural to urban areas. It was noted that in 1995 there was a "floating population" of some 80-100 million, who stayed outside their own *hukou* registration place.[5] The increase in mobility had forced the government to change its *hukou* policies. In 2005, the deputy minister of public security stated that eleven provinces had begun or would soon begin to implement unified urban-rural household registration system, removing the distinctions between agricultural and non-agricultural *hukou* status.[6] The ultimate aim is to abolish the regulations of migration in the *hukou* system.

[4] See Whalley and Zhang (2007) for the detail.

[5] See Chan and Zhang (1999) for the detail.

[6] However, Beijing was tentatively excluded. See Chan and Buckingham (2008) for the detail. An updated statement in 2007 repeated the same points and included a list of twelve provincial-level units.

Changing from agriculture to non-agriculture was commonly known as *nongzhuanfei*. The annual quota was controlled by the central government at about 0.15-0.2 percent of the non-agricultural *hukou* population in each area. The regular channels of *nongzhuanfei* included recruitment by (i) a state-owned enterprise (*zhaogong*), (ii) enrollment in an institution of higher education (*zhaosheng*), and (iii) promotion to a senior administrative job (*zhaogan*). Therefore, *zhaosheng* was one of the formal channels for rural students to obtain an urban *hukou* in China.

While the *zhaogan* channel is related to the official permit, the so-called *red stamp* (which need not be due to economic factors and hence ignored in our paper), the interesting part of *hukou* reforms were to relax the migration regulations for the general public as captured by the two other channels. For example, state governments implemented a new type of urban *hukou* with "self-supplied food grain" in 1984. People actually living in urban areas other than their place of *hukou* registration were required to apply for a temporary residence certificate. In addition, due to the demand of economic development, some local governments introduced the *blue-stamp* urban *hukou* in the early 1990s in order to attract professional workers and investors.[7] The blue-stamps *hukou* system allowed them to obtain a temporary urban *hukou*. It could be upgraded to a formal urban *hukou* under some conditions and after some years. However, applying the blue-stamp *hukou* required an urban infrastructure construction fee for any new comer, ranging from a few thousand to some fifty thousand yuans. Blue-stamps *hukou* was different from the urban *hukou* obtained through *nongzhuanfei* that it only had limited rights and obligations compared to regular urban residents.

We now turn to the zhaosheng channel. The system of *gaokao* was established at the beginning of the 1950s. It was abolished for several years during the periods of Cultural Revolution and then was restored in 1977. Because of the scarcity of education resources, the acceptance rates were very low, especially in the 1980s. Students who passed *gaokao* could move their *hukou* to cities, enjoyed lots of social benefits, and were expected to become a pillar of society in the future.

At the beginning of entering an university or a college, a freshman could be voluntary to change his own *hukou* registration place to his school (a collective joint household). This would also change his *hukou* status to become non-agricultural. During the periods with the implementation of GJA (1951-1994), a graduate would be distributed to a stable government job. Thus, a graduate simply moved his *hukou* to the working unit after graduation. He still kept the non-agricultural status.[8]

[7] This was distinguished by a blue stamp, different from the formal urban *hukou* book with a red stamp.

[8] The reform of the GJA started in 1989; but was officially abolished in 1996. Tibet, which abolished the system

In the periods that the GJA was abolished, a graduate could move his *hukou* to the working unit, temporarily put to the collective joint household of personal exchange center, or moved back to his own hometown.[9] Thus, under China's *hukou* system, entering college through *gaokao* implied a rural to urban migration. It was an upward mobility in the society.[10]

3 The model

In this section, we build a model to study the transition process of rural-to-urban migration under the *hukou* system of China and to evaluate the policies on rural-urban migration. We use (i, j, k) to denote three consecutive generations and let (H, L) denote high skill and low skill. There are two geographical regions, rural and urban area (R and U), respectively. Our optimization problem focuses only on the decision of rural parents (generation i) in sending their children to urban area to get education. Therefore, our specification of the urban households is rather simple. Assume that there is an initial mass of workers in urban area given by (N_H, N_L), where N_H (N_L) denotes the total number of high (low) skilled workers. To focus on rural-to-urban migration, we shut down the channel of endogenous decision of "moving back" from urban to rural and assume that the *hukou* of urban households are passed from one generation to another. Also, to simplify our analysis, we assume that there is no population growth. That is, the total population in rural and urban area is constant over time.

3.1 Production

In urban area, production is taken place with the following nonhomothetic CES production function:

$$Y_U = A\left[\alpha\left[\left(N_H + \psi\right)h\right]^{\rho} + (1-\alpha)\,N_L^{\rho}\right]^{1/\rho},\ \alpha \in (0,1) \tag{1}$$

where $A > 0$ is the technology scaling factor in urban (or urban TFP, hereafter we will use the two terms interchangeably), ψ is a constant that resides in urban area, $\rho < 1$ and $1/(1-\rho)$ is the elasticity of substitution in production for high and low skilled labor, and h is the level of human

in 2007, was the last place to terminate the distribution system of graduation.

[9] Even though a graduate moved his *hukou* back to his hometown, he still kept the non-agricultural status. This implied that he could not rent land and inherit the land rented by his parents.

[10] Since 1996, China has introduced a series of education reforms. In particular, since 2003, universities in Hong Kong are allowed to enroll students in Mainland China. Since the reform, attending universities in Hong Kong is getting popular.

capital possessed by the high-skilled worker. With this specification, firms can still operate and have positive output even if there is no high-skilled workers.

As documented by Maurer-Fazio (1999), skilled labor wage in China was depressed due to the planned economy system. To capture this phenomenon, we introduce a wedge $\tau \in (-1, \infty)$ facing by urban firms when hiring high skilled workers. Denote w_H as the effective high skilled wage received by high skilled workers, and w_L as the low skilled wage. The urban wage rates are determined by:

$$(1+\tau)\, w_H = \frac{\partial Y_U}{\partial (N_H + \psi)\, h} = \alpha A \left[\alpha \left[(N_H+\psi)\, h\right]^{\rho} + (1-\alpha)\, N_L^{\rho}\right]^{1/\rho - 1} \left[(N_H+\psi)\, h\right]^{\rho-1}, \quad (2)$$

$$w_L = \frac{\partial Y_U}{\partial N_L} = (1-\alpha)\, A \left[\alpha \left[(N_H+\psi)\, h\right]^{\rho} + (1-\alpha)\, N_L^{\rho}\right]^{1/\rho - 1} N_L^{\rho-1}, \quad (3)$$

and the skilled-unskilled wage ratio is thus given by:

$$\frac{w_H h}{w_L} = \frac{1}{1+\tau} \frac{\alpha}{1-\alpha} \left(\frac{N_L}{N_H+\psi}\right)^{1-\rho} h^{\rho}$$

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

$$Y_R = B N_R, \quad (4)$$

where N_R is the number of raw-labor workers in rural area and B is the technology scaling factor in rural area. Competitive labor market implies that the rural wage rate is

$$w_R = B. \quad (5)$$

3.2 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). There is no fertility decision and we assume that each adult agent gives birth to a child. Assuming the utility function $u(\cdot)$ is strictly increasing and strictly concave, then the representative household objective is:

$$\Omega^i \left(\mathbf{I}^j | \mathbf{I}^i = 0, \mathbf{I}^k, x^j\right) = \max_{\mathbf{I}^j} u\left(c^i\right) + \beta \mathbb{E}_{\mathrm{X}} u\left(c^j\right) \quad (6)$$

where β is the altruistic factor on children and $\mathbf{I}^j$ is an indicator function such that

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

We assume constant child-rearing cost ϕ^i. Education of the children only takes place in urban area with cost x^j, which is a random variable that is inversely related to the talents of the child z^j, i.e., $z^j \equiv 1/x^j$. We note that z^j is drawn from a distribution with cumulative distribution function denoted by $G\left(z^j\right)$. Finally, the migration cost (settle-down cost in cities) of moving from rural to urban is given by a constant σ. Then the budget constraint for a rural parent is

$$c^i + \mathbf{I}^j \cdot \left(x^j + \sigma\right) + \phi^i = w_R. \tag{7}$$

Children become skilled workers after receiving education in urban area. They can either get high (low) skilled job in urban area, earning a skilled wage income w_H (w_L) with probability γ_H (γ_L), or otherwise are forced to move back to rural area, earning a rural wage income of w_R. Children that remain in the rural area do not incur any cost in education or migration. They either get recruited to serve as low skilled worker in urban area and earn w_L (with probability π), or otherwise work as unskilled labor in rural area and earn w_R. The income for the children (generation j) when they become adults is

$$W^j = \mathbf{I}^j \left[\gamma_H w_H h + \gamma_L w_L + \left(1 - \gamma_H - \gamma_L\right) w_R\right] + \left(1 - \mathbf{I}^j\right) \left[\left(1 - \pi\right) w_R + \pi \left(w_L - \sigma\right)\right]. \tag{8}$$

Then the children's budget constraint is given by

$$c^j + \mathbf{I}^k \cdot \left[\mathbf{I}^j \left(1 - \gamma_H - \gamma_L\right) + \left(1 - \mathbf{I}^j\right)\left(1 - \pi\right)\right] \left(x^k + \sigma\right) + \phi^j = W^j \tag{9}$$

where

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

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 to them. We use X to denote the random variable of education cost. For illustration purpose, we plot in Figure 3 the timeline of the model.

An agent's discrete-choice problem is to decide whether to send his or her child to 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 among the two. By substituting $c^i = w_R - \mathbf{I}^j \cdot \left(x^j + \sigma\right) - \phi^i$ and $c^j = W^j - \mathbf{I}^k \cdot \left[\mathbf{I}^j \left(1 - \gamma_H - \gamma_L\right) + 1 - \mathbf{I}^j\right] \left(x^k + \sigma\right) - \phi^j$ into the value functions, with W^j given by

(8), we have:

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

and

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

Defining $\Delta^i\left(\mathbf{I}^k,x^j\right)\equiv\Omega^i\left(1|0,\mathbf{I}^k,x^j\right)-\Omega^i\left(0|0,\mathbf{I}^k,x^j\right)$, we have:

$$\Delta^i\left(\mathbf{I}^k,x^j\right) = u\left(w_R-x^j-\sigma-\phi^i\right)-u\left(w_R-\phi^i\right) +\beta\mathbb{E}_{\mathrm{X}}\left\{\begin{array}{c}u\left(\gamma_H w_H h+\gamma_L w_L+(1\text{-}\gamma_H\text{-}\gamma_L)\,w_R\text{-}\mathbf{I}^k(X)\,(1\text{-}\gamma_H\text{-}\gamma_L)\,(X+\sigma)\text{-}\phi^j\right)\\ \text{-}u\left((1\text{-}\pi)\,w_R+\pi\,(w_L\text{-}\sigma)\text{-}\mathbf{I}^k(X)\,(1-\pi)\,(X+\sigma)\text{-}\phi^j\right)\end{array}\right\}. \tag{10}$$

To proceed further, we define $n\equiv\frac{(N_H+\psi)h}{N_L}$. The high-skilled and low-skilled effective wage in (2) and (3) can be rewritten as:

$$\begin{aligned} w_H(n) &= \frac{\alpha A}{1+\tau}\left[\frac{\alpha n^\rho+(1-\alpha)}{n^\rho}\right]^{\frac{1}{\rho}-1}, \\ w_L(n) &= (1-\alpha)\,A\left[\alpha n^\rho+(1-\alpha)\right]^{\frac{1}{\rho}-1}. \end{aligned}$$

It can be shown that w_H is decreasing in n while w_L is increasing in n.[11] The skilled-unskilled wage ratio is then given by:

$$\frac{w_H h}{w_L}=\frac{1}{1+\tau}\frac{\alpha}{1-\alpha}n^{\rho-1}h,$$

[11] These are standard results of CRTS production functions. Suppose we the following CRTS production function of urban output:

$$Y_U=F\left[(N_H+\psi)\,h,N_L\right].$$

This can be written in terms of $n\equiv\frac{(N_H+\psi)h}{N_L}$:

$$Y_U=N_L F\left[n,1\right]=N_L f(n)$$

where $f'>0>f''$. So we get

$$(1+\tau)\,w_H=f'(n)\,,w_L=f(n)-nf'(n)\,.$$

These in turn yield

$$(1+\tau)\frac{\partial w_H}{\partial n}=f''(n)<0,$$
$$\frac{\partial w_L}{\partial n}=-nf''(n)>0,$$

where our CES technology is a special case.

which is decreasing in n. We further define n_s such that $\frac{w_H h}{w_L} = 1$ when $n = n_s$. Thus we have:

$$n_s \equiv \left(\frac{1}{1+\tau} \frac{\alpha}{1-\alpha} h \right)^{\frac{1}{1-\rho}}.$$

Thus, we impose the following condition:

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

Only under Condition S rural parents will consider sending their children to urban area to attend college. To better understand Condition S, we plot in Figure 4 the high- and low-skilled wages against n. It can be easily seen from the figure that Condition S requires that urban wages, net of migration costs, are higher the rural wage. It is a *sufficient* condition to guarantee that rural agents have incentives to migrate to cities. If Condition S is violated, staying in rural may be a better option for them. We also impose an assumption on the probabilities of getting urban jobs: the relative probability of getting an urban job via education *cannot* be too low.

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

Assumption 1 states that the probability of getting an urban job after receiving college education cannot be lower than that for rural workers to find an urban job. Only when this assumption is satisfied, rural parents will have incentives to send their children to cities to attend college. Now, denote c_U^j as the consumption of children if they are sent to urban area and c_R^j as the consumption of children if they are kept in rural area. We have the following lemma:

Lemma 1 *Under Assumption 1, if Condition S is satisfied,* $u\left(c_U^j\right) > u\left(c_R^j\right)$.

Proof.

From (8) and (9) we have:

$$\begin{aligned} c_U^j &= \gamma_H w_H h + \gamma_L w_L + (1-\gamma_H-\gamma_L) w_R - \mathbf{I}^k(X)(1-\gamma_H-\gamma_L)(X+\sigma) - \phi^j, & (11) \\ c_R^j &= (1-\pi) w_R + \pi (w_L - \sigma) - \mathbf{I}^k(X)(X+\sigma) - \phi^j. & (12) \end{aligned}$$

By substracting (12) from (11) and rearranging terms, under Condition S, we have:

$$\begin{aligned} c_U^j - c_R^j &= \gamma_H w_H h + \gamma_L w_L + (\pi-\gamma_H-\gamma_L) w_R + \mathbf{I}^k(X)(\gamma_H+\gamma_L)(X+\sigma) - \pi(w_L-\sigma) \\ &= \gamma_H w_H h + \gamma_L w_L - \pi w_L + (\pi-\gamma_H-\gamma_L) w_R + \mathbf{I}^k(X)(\gamma_H+\gamma_L)(X+\sigma) + \pi\sigma \\ &> (\gamma_H+\gamma_L-\pi)(w_L-w_R) + \mathbf{I}^k(X)(\gamma_H+\gamma_L)(X+\sigma) + \pi\sigma \\ &> 0. \end{aligned}$$

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

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

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

We are able to present our first finding on parents' decision for their children's education choice.

Proposition 1 *Under Assumption 1, if Condition S is satisfied, parents will send their children to urban area to attend college.*

The intuition of the above proposition is straightforward. If the probability of finding an urban job is reasonably high (Assumption 1), it is worthwhile for parents to pay the educational and migration costs to send their children to urban area to receive education. Otherwise, sending children to urban area to attend college would not be a good "investment" from parents' perspective.

3.3 Comparative statics

In the following, we examine how $\Delta^i\left(\mathbf{I}^k, x^j\right)$ responds to changes in the parameterization, i.e. we examine whether the "marginal" parent (parent who is indifferent between sending children to attend college in urban area or keeping children in rural area) will send her children to urban area or keep the children in rural area. Denote $u_{c_U}^i = u_c\left(c_U^i\right)$ as the marginal utility, we have

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

We have the following proposition:

Proposition 1

Under Condition S and Assumption 1,

1. *More parents will be willing to send their children to urban area to attend college when their children become more talented ($x_j \downarrow$).*

2. *More parents will be willing to send their children to urban area to attend college when the chances for the children to obtain a job in cities are higher ($\gamma_H \uparrow$, $\gamma_L \uparrow$).*

3. *Less parents will be willing to send their children to urban area when the chance of being nongmingong becomes higher ($\pi \uparrow$).*

To examine how changes in migration cost (σ) affect parents' decision, we compute

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

Proposition 2

Under Condition S, parents will be more willing to send children to attend college in urban area even when the migration cost increases.

Proof.

Under Condition S', $c^j_U > c^j_R$ and $u^j_{c_U} < u^j_{c_R}$. Define $\Lambda \equiv u^j_{c_U}\left(1-\gamma_H-\gamma_L\right)\mathbf{I}^k\left(X\right) - u^j_{c_R}\left[\pi + \left(1-\pi\right)\mathbf{I}^k\left(X\right)\right]$. Also,

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

Therefore, if $-\beta\mathbb{E}_{\mathrm{X}}\Lambda > u^i_{c_U}$, $\frac{d\Delta^i\left(\mathbf{I}^k, x^j\right)}{d\sigma} > 0$.

3.4 Evolution of workers

Only adult agents supply labor to the market and each agent gives birth to only one child, and hence the whole adult population participates in the labor market. Denote $\left(N^t_H, N^t_L\right)$ as the skilled and unskilled workers in the city and N^t_R as the rural labor force 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 by generation-j urban worker with job J, working as a type K worker in urban area. Thus, δ_{JK} captures the job mobility across generations in urban area. In general, we shall expect that $\delta_{JJ} > \delta_{JK}$ for $J \neq K$. Under the assumption that the *hukou* of urban households are passed from one generation to another, we have

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

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

$$\begin{aligned} N_H^{t+1} &= \delta_{HH} N_H^t + \delta_{LH} N_L^t + N_R^t \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) \gamma_H dG(z^j) && (14) \\ N_L^{t+1} &= \delta_{HL} N_H^t + \delta_{LL} N_L^t + N_R^t \left\{ \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) \gamma_L dG(z^j) + \int \left[1 - \mathbf{I}^j \left(z^j, \mathbf{I}^k\right)\right] \pi dG(z^j) \right\} && (15) \\ N_R^{t+1} &= \left(1 - \delta_{HH} - \delta_{HL}\right) N_H^t + \left(1 - \delta_{LH} - \delta_{LL}\right) N_L^t \\ &\quad + N_R^t \left\{ \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) \left(1 - \gamma_H - \gamma_L\right) dG(z^j) + \int \left[1 - \mathbf{I}^j \left(z^j, \mathbf{I}^k\right)\right] \left(1 - \pi\right) dG(z^j) \right\} && (16) \end{aligned}$$

with the initial labor force in the city and in the countryside, denoted by $\left(N_H^0, N_L^0\right)$ and N_R^0, given, respectively. With the assumption given in (13), equations (14) - (16) can further be simplified as

$$\begin{aligned} N_H^{t+1} &= \delta_{HH} N_H^t + \left(1 - \delta_{LL}\right) N_L^t + N_R^t \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) \gamma_H dG(z^j), && (17) \\ N_L^{t+1} &= \left(1 - \delta_{HH}\right) N_H^t + \delta_{LL} N_L^t + N_R^t \left\{ \pi + \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) \left(\gamma_L - \pi\right) dG(z^j) \right\}, && (18) \\ N_R^{t+1} &= N_R^t \left\{ \left(1 - \pi\right) - \int \mathbf{I}^j \left(z^j, \mathbf{I}^k\right) \left(\gamma_H + \gamma_L - \pi\right) dG(z^j) \right\}. && (19) \end{aligned}$$

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

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

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

In the quantitative analysis below, we will focus on the case where $\delta_{HH} = 1$, $\delta_{HL} = 0$, $0 < \delta_{LL} < 1$ and $\delta_{LH} + \delta_{LL} = 1$.

3.5 Equilibrium

First, all labor markets must be clear under the factor prices $\{w_H, w_L, w_R\}$ given by (2), (3) and (5):

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

where N_J^{dt} denotes labor demand of type J. Finally, 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.

We are now ready to define the equilibrium for our model.

Definition 1 *A competitive equilibrium of the model consists of consumption, migration choice and wage rates, $\{c^i, c^j, \mathbf{I}^j, w_H, w_L, w_R\}$, such that $\{w_H, w_L, w_R\}$ are given by (2), (3) and (5), and the parents (generation i) solve (6) subject to (7), (8) and (9), given initial distributions of population, N_H^0, N_L^0 and N_R^0, in the two locations. In addition, $\{w_H, w_L, w_R\}$ are such that labor markets clear in all locations as given by (20) and (21).*

4 Numerical Analysis

The numerical analysis focuses on the periods that China has started its economic reform, 1980-2007. In addition, there was a policy reform on the government job assignment for college graduates in 1994: from 1951 to 1994, government assigned jobs to all college graduates; after 1994 obtaining a college degree does not imply that a job is guaranteed. To consider the policy change, the period of 1980-2007 is further divided into two regimes: the first regime spans from 1980-1994 and the second regime is for the period of 1995-2007.

The model is calibrated to fit the data from China for the two regimes. Based on the calibrated parameters, the urban TFPs for the period of 1981-2007 is backed out. Then the model is simulated from 1981 to 2007 to be our benchmark model. Using the benchmark model, we decompose the contributions of *zhaosheng* and working migration to per capita output, urban shares, and skill premium. Finally, various policy experiments are performed.

4.1 Calibration

The model period is 25 years. In every period, total population is normalized to be one so that there is no population growth for the economy. Using the ratios of rural and urban residence to total population in data, the rural and urban population are computed.[12] High skilled workers in the numerical analysis are defined as workers with educational attainment of college or above. Then,

[12] Source: China Population and Employment Statistics Yearbook.

using the data of employment by educational attainment in cities, the stocks of high and low skilled workers are computed accordingly.[13]

The preset common parameters in both regimes are described as follows. Following Liao (2013), the annual time preference rate for China is set at 4.43 percent and hence the parental altruistic factor on children β is equal to 0.3384. The utility function is assumed to be the CRRA form:

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

where ε is the inverse of the elasticity of intertemporal substitution. The real business cycle literature conventionally sets ε to be one. Hall (1989) argued that the inverse of the elasticity of intertemporal substitution is very high. Balvers and Bergstrand (2002) estimated that the value of ε is 2.04. Chang (1998) chose to set ε at 2. Here we set ε to be 2. There is no nationwide survey on child-rearing cost for rural China. Zhu and Zhang (1996) estimated that the average child-rearing cost for rural villages in Xianyang, located in Shaanxi province of China, was about 17.4% of family income for an age 0-16 child in 1995. Therefore, we set ϕ to be 0.174, i.e., 17.4% of rural household income.[14] We assume that children's talents follow a Pareto distribution because in the literature the distribution is commonly associated with wealth and income, which are correlated with talents. The cdf for talent z^j is 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 parameter of Pareto distribution, respectively. Following the literature studying on firm size, productivity, and international trade, the location parameter $z_{\min}$ is set to be 1.[15] In addition to the tuition and migration cost, rural parents also have to pay for their own consumption and the child-rearing cost. Thus, setting $z_{\min}$ to be one implies that not all rural parents are affordable for college education. Only rural parents having relatively talented children are able to send their children to college in urban areas. Another parameter of the Pareto

[13] We only have data on 2002-2007 urban employment by educational attainment. For 1982, 1990, 1995-1999 and 2001, we compute N_H and N_L using the whole country data with adjustments based on 2002-2007 urban-to-whole country employment by educational attainment data. For 1980, 1985 and 2000, we do the similar adjustments using Barro-Lee (2000) data. For years with no available data, we compute N_H and N_L with intrapolation.

[14] There are estimates for child-rearing costs in urban China. Ye and Ding (1998) study the child-rearing cost for age 0-16 children in Xiamen Special Economic Zone in 1996. They find that average child-rearing cost in Xiamen accounted for 34% of annual family income, and the number for Beijing in 1995 was roughly 20%.

[15] See, for example, Melitz and Ghironi (2005), Melitz and Redding (2014), Mayer, Melitz, and Ottaviano (2014), Bernard et al (2003), and Eaton and Kortum (2002).

distribution is θ. Because talents are believed to be correlated with income levels, we attempt to use the estimate in the literature of income inequality. However, there is no such estimate for China. We thus borrow the estimate for the U.S. economy: Feenberg and Poterba (1993) estimated that the average shape parameter in Pareto distribution for the U.S. income inequality in 1980-1990 is 1.92.[16] Thus θ is set to be 1.92. The last preset common parameter is the elasticity of substitution between high and low skilled labor in the production function $1/(1-\rho)$. The estimate for developed countries in the literate ranges from 1 to 3.[17] For developing countries, the estimated value is larger, from 2 to 7. [18] We thus choose the elasticity of substitution between high and low skilled labor to be 2 so that ρ is equal to 0.5.

Other preset parameters are described as follows. The rural production technology is linear, so the rural technology scaling factor B is equal to the rural wage rate. Because we are interested in learning the relative economic positions of rural and urban areas in China and how the regional technological disparities shape individuals' migration decisions, we normalize rural per capita income of 2007 to be one, i.e., B of 2007 is 1. Then, using the time series data of rural per capita income, we obtain the average rural per capita income and B in regime 1 is 0.3685 and the ratio of rural per capital income in regime 2 to that in regime 1 is 1.947. Therefore, the average rural per capita income and B in regime 2 is 0.7177.

Because of the GJA policy, college graduates were assigned jobs immediately before 1994. Specifically, the assigned jobs were either in the government sector or in the state-owned enterprises. Thus, the job finding rates for college graduates are set to be $\gamma_H = 1$ and $\gamma_L = 0$ in the first regime. In the second regime, the average employment rate in city districts in 1995-2007 is 0.9627.[19] We do not have data on job mismatching rate for college graduates, but we believe that the job mismatching rate should be low over the period of 1995-2007. Thus, we choose to set γ_L at 0.05 and γ_H is

[16] The average estimate of the shape parameter for the U.S. income inequality in 1950-1990 is 2.11 in Feenberg and Poterba (1993).

[17] See, for example, Autor, Katz and Krueger (1998), Acemoglu (2003), Ciccone and Peri (2005), and O'Rourke, Rahman and Taylor (2013).

[18] For example, Toh and Tat (2012) estimated that the value for Singapore is 4.249. te Velde and Morrissey (2004) used data from Singapore, Hong Kong, Korea, Philippines and Thailand to estimate and obtained the value of 2.78. The results in Gindling and Sun (2002) imply that the value in Taiwan is between 2.3 and 7.4.

[19] In the model, $\gamma_H + \gamma_L$ refers to the employment rate for college graduates who migrated from rural area. However, there is no counterpart in the data. We thus have to use urban employment rate, computed by the formula: 1−total registered unemployed persons (with urban non-agricultural hukou)/total work force in urban areas. We are aware that the actual employment rate in urban area could be lower.

solved to be 0.9127. The calibrated results are not sensitive to the value of γ_L that we choose for regime 2. The next preset parameter is the probability of working migration π. First, there is no available nationwide survey on rural to urban migration in China. Second, our model abstracts from endogenous fertility and mortality. Third, China adopted a much serious one-child policy in urban area than in rural area. Thus, we believe that the change in urban population can be a good proxy for the amount of rural to urban migration. The change in urban population is then multiplied by working-related migration reasons (column 3-5 in Table 1, the average of 1985 and 2000) to obtain the total number of migrant workers.[20] Finally, the total number of migrant workers is divided by the stock of rural population to obtain the probability of a rural worker to migrate to cities for working. As reported in Table 2, the average probabilities for regime 1 and regime 2 are 0.0036 and 0.0083, respectively.

The human capital possessed by high-skilled workers h relative to low-skilled workers is computed based on the Mincerian method. To do this, we first compute average years of schooling for educational attainment of below college and college or above. In the first regime, the average years of schooling below college was 8.02 and 14.1 for college and above. For the second regime, they were 8.95 and 14.52 for below college and for college and above, respectively.[21] In addition, the education return coefficients reported by Zhang et al. (2005) for regime 1 and regime 2 are 0.0479 and 0.0835, respectively. Thus, the computed h for the first and second regime are 1.3381 and 1.5922, respectively. The last preset parameters are those for intergenerational mobility. In the numerical analysis, we assume that the residence of urban households are passed from one generation to another and no reverse migration. Besides, in both regimes, we focus on the case of upward mobility. Therefore, $\delta_{HH} = 1$, $\delta_{HL} = 0$ and $\delta_{LH} + \delta_{LL} = 1$ in both regimes. The method of calibrating δ_{LL} will be described later.

Other parameters are calibrated to match data moments or are solved by model equations. They are discussed as follows. To calibrate the intergenerational mobility from low skilled worker to low skilled worker δ_{LL}, we first compute the zhaosheng flow from data following the same way

[20] Data on migration by reasons are taken from the 1990 and 2000 census. The 1990 census reports information on the immigration by reasons (by residence of 1985) and the 2000 census reports information on emigration by reasons. To match the notion of migration in our model, we categorize immigration (emmigration) flow due to *study or train* as migration through zhaosheng, and categorize immigration (emmigration) flow due to *work and business*, *job transfer*, *and job assignment* as "working migration" or "migration due to work".

[21] The details are summaried in Table A.1 and A.2.

that we used to compute the number of working migrants (as reported in Table 2). Based on the computed zhaosheng flow, the stocks of rural workers, high skilled workers and low skilled workers in urban area implied by our model are computed by the evolution of workers equations, (17)-(19). Then the ratio N_H/N_L implied by the model is constructed. Finally, δ_{LL} is solved so that the model ratio N_H/N_L matches the ratio of N_H/N_L in data for both regimes. δ_{LL} is thus equal to 0.9996 and 0.9885 for regime 1 and regime 2, respectively. The decrease δ_{LL} shows that the upward intergenerational mobility in China has been improved over 1995-2007. The wedge τ on the return of high skilled workers is solved to match the regime average of skill premium $w_H h/w_L$ 1.2296 in the first regime and 1.6576 in the second regime. The calibrated wedge is equal to 2.6960 and 1.7960 in regime 1 and 2, respectively. Urban TFP A is calibrated to match the regime average of urban premium w_L/w_R 1.7781 and 2.0076 in regime 1 and 2, respectively. The calibrated A is equal to 2.7365 in regime 1 and 5.5010 in regime 2. We assume that the parameters in urban production function α and ψ are deep parameters. Thus, they are calibrated together to match 1 minus labor income share in data. The data of average labor income share in regime 1 is 0.4614 and in regime 2 is 0.5002. Thus, the calibrated α and ψ are equal to 0.6762 and 0.0562, respectively.

The last parameter is the migration cost σ. The indifference boundary equation (10) is used to pin down σ. Rural parents decide whether to send their children to college or not according to the expected wages in urban area, the probability of finding good jobs and the migration cost. Denote $\hat{z}$ as the threshold of children's talent such that rural parents are indifferent between sending their children to college or not. In other words, when children are talented such that $z^j \equiv 1/x^j \geq \hat{z}$, their parents will send them to college ($\Delta^i(\mathbf{I}^k, x^j) \geq 0$). The endogenous threshold $\hat{z}$ therefore dichotomizes the "destiny" of rural children, deciding whether they belong to the "sent" group or the "stay" group. More specifically, define $\tilde{N}_E^t$ as the *zhaosheng* flow at time t. Because children's talents follow the Pareto distribution, $\tilde{N}_E^t$ 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 directly computed using the data of *zhaosheng* flow. The computed average $\hat{z}$ for the first and the second regime are equal to 42.36 and 28.61, respectively. Setting the indifference boundary equation to be zero, now we are ready to solve the migration costs. The calibrated σ are 0.0268 and 0.0807 for regime 1 and 2, respectively. All parameters are summarized in Table 3.

Table 4 provides the calibration results for the two regimes. The migration cost as a percentage of rural wage rate is 7.3% in regime 1 and 11.2% in regime 2. They are similar to the setting in the literature.[22] As expected, the relative TFP of urban to rural area from regime 1 to regime 2 slightly increases. Based on the above parameters, the next step is to calibrate the annual urban TFP for the period of 1981-2007 and do simulation to be our benchmark model. Once the benchmark model is obtained, we are able to decompose the effects of *zhaosheng* and working migration to macroeconomic variables. Besides, policy experiments are also ready to conducted.

4.2 Calibration for the annual urban TFP and simulation

To obtain annual urban TFP, we calibrate a series of urban TFP A and a series of wedge τ, so that skill premium $w_H h/w_L$ and urban premium w_L/w_R exactly match the time series data. The procedure is described as follows.

First, in the two-regime calibration, we use Mincerian method to compute the human capital for high-skilled workers h. They are equal to 1.3381 and 1.5922 in the first and second regime, respectively. To obtain a series of h, we further assume that the annual growth rate of human capital is constant during the period of 1980-2007. Then the series of h is computed so that the average human capital in the period of 1981-1994 is exactly equal to the regime average 1.3381 and the average of those in the period of 1995-2007 is equal to 1.5922 in the second regime. Second, the simulation procedure of population goes as follows. Similar to the two-regime calibration, we use the change in urban population as proxy for the amount of rural to urban migration. The change in urban population is then multiplied by the migration reasons of studying or training (column 6 in Table 1, average of 1985 and 2000) to obtain the total number of *zhaosheng*. The total number of *zhaosheng* is divided by the stock of rural population to obtain a time series of *zhaosheng* flow. Given the *zhaosheng* flow, the 1980 data on N_R, N_H and N_L, and the calibrated parameters, we are able to compute a series of the threshold $\hat{z}$, working migration, N_R, N_H, and N_L according to the evolution of workers equations, (14)-(16). Finally, importing the time series data of rural per capita income (2007 is normalized to be one) and using the skill premium equation and urban premium equation, a series of urban TFP A and wedge τ are solved so that the skill premium $w_H h/w_L$ and urban premium w_L/w_R implied by the model exactly match the time series data.

Figure 5 plots the calibrated urban TFP and the rural TFP during 1981-2007. It can be noted that, since 1985, the urban TFP relative to rural TFP exhibits a slightly upward trend.

[22] Based on the estimates in Zhao (1999b), Liu (2011) sets the migration costs at the level of 15% of annual income.

This may correspond to China's economic reform, the privatization of state-owned enterprises, and deregulation of price controls. As reported in Table 4, the average annual growth rate of the calibrated urban TFP is about 5.3% and the average growth rate of the calibrated relative TFP is lower, about 0.2% per year.

Figure 6 provides a comparison between the model and the data for urban per capita output and total output level.[23] Define the urbanization rate as the ratio of urban workers to total workers. Figures 7 compares the model performances to the data from the perspective of urbanization rates and the stocks of urban high- and low-skilled workers. Our model shows a lower urbanization rate than that in the data. Besides, the discrepancy between the model and the data widens as time goes by. The gap could be explained by the migration reasons we included when the model is calibrated: In the model, only educational and working-related reasons are included. They only account for about 50% of total migration. Migrants due to other reasons could also go to work as low-skilled workers later. However, we do not account for these migrants in the calibration. As a result, our model underperforms the urbanization rate and the stock of urban workers. Since there are fewer workers in urban area, especially mainly due to fewer low-skilled workers, the urban per capita output in the model is slightly higher than that in the data. Because more workers stay in rural area in the model and rural technology is less productive, total output level in the model is slightly lower than that in the data.

This simulation is our benchmark model for decomposition and in comparison with policy experiments later. Table 5 summarizes macroeconomic variables in the benchmark model for the whole period (1981-2007), the period in the first regime (1981-1994), and the period in the second regime (1995-2007). As expected, total output per capita in regime 2 is more than double to that in regime 1. Urban output as a percentage of total output increases. In addition, the share of urban employment goes up. These implies that urban production becomes more important in the second regime. In urban area, the ratio of high-skilled employment in regime 2 is about four times as large as that in regime 1 and skill premium increases. These are all consistent with the experience of China's development.

[23] Based on equations (1) and (4), we use the data on N_H, N_L and N_R to obtain the computed data on urban per capita output and output levels.

4.3 Decomposition: The contribution of *zhaosheng* and working migration

To study the contribution of zhaosheng and working migration to macroeconomic variables, we shut down the migration channel of zhaosheng, the channel of working migration, and finally all channels of rural-urban migration is completely shut down. Then the above three scenarios are compared with the benchmark model to examine the contribution of zhaosheng and working migration to the development of China during 1981-2007.

Figure 8 depicts urban per capita output, total output level, and the urbanization rate under the decomoposition. The benchmark model and the three scenarios are plotted for comparison: (1) only the channel of zhaosheng is allowed; (2) only working migration is allowed; and (3) no rural-urban migration is allowed, i.e., both migration channels are shut down. In the first scenario that only the channel of zhaosheng is allowed, urban per capita output is higher than that in the benchmark model because all migrants are high-skilled workers and they are productive. But, the percentage of migrants through zhaosheng is not large, so the urbanization rate in this scenario is far away from the benchmark model. For the second scenario that only working migration is allowed, urban per capita output is slightly lower than that in the benchmark model since working migrants are less productive. In contrast, urbanization rate is now very close to that in the benchmark model. Again this indicates the fact that working is the main reason for the rural-urban migration.

To further study the contribution of each migration channel to various macroeconomic variables, total per capita output, the share of urban output, the share of urban employment, the share of high-skilled employment, and skill premium are calculated in all scenarios. Table 6 reports the level values. Percentage changes relative to the benchmark model are summarized in Table 7. The results show that the contribution of zhaosheng to total output per capita is 3.4% of the benchmark model, while the contribution of working migration is 4.9% during the whole period. Compared with the magnitude of working migration, the effects of zhaosheng cannot be ignored. We also find that in the first regime the contribution of zhaosheng is slightly more important than that of working migration. In contrast, in the second regime, working migration matters. For the share of high-skilled employment, zhaosheng contributes more than one-fourth of that in the benchmark and thereby lowering skill premium. This is because most of the migrants through zhaosheng are high-skilled workers. In contrast, working migration has negative impacts on the share of high-skilled employment and skill premium goes up.

The results also suggest the complementarity between high- and low-skilled workers, although the complementary is more important for urban production in the second regime, explaining 0.6% of the share of urban output in the model. When both channels of migration are allowed, there is a larger stock of low-skilled workers in urban area, so the complementarity between high and low-skilled workers in production leads to a higher high-skilled wage. This makes college education more attractive, and rural parents are more willing to send their children to attend college. Therefore, the share of urban output goes up.

4.4 Counterfactual analysis

Based on the benchmark model, we are now ready to conduct counterfactual experiments. Three types of policy experiments are provides: (1) policies on GJA; (2) policies on working migration; and (3) other policies. The details are discusses as follows. Table 8 reports the level values of all experiments. The percentage changes relative to the benchmark model are summarized in Table 9.

(1) Policies on GJA

Two policies experiments are conducted here. First, we deal with a scenario that the Chinese government continued the GJA policy after 1994. In other words, γ_H in the second regime is set to be one in the experiment. Because high-skilled jobs is guaranteed for college graduates, college education becomes more attractive, the share of urban employment slightly goes up, and the share of high-skilled employment slightly increases. The policy has a slightly positive impact on total output per capita and on the share of urban output. The magnitude is small because in the benchmark model the job finding rate for college graduates is already very close to one. Thus, even if the government continues the GJA policy, the influences are small. These results provide a justification for the abolishment of the GJA policy in 1994.

The second scenario is what China would be if the Chinese government had not introduced the GJA policy in the first regime. To conduct this experiment, γ_H in the first regime is set to be equal to that in the second regime. The results show that China's total output per capita would be lower and the share of urban output and urban employment would be smaller. This is mainly due to fewer high-skilled workers. However, again the magnitude is small because the job finding rate for college graduates in the benchmark model is close to one. Figure 9 and 10 plot the simulated series for these two experiments. As shown in the figures, the differences between the benchmark model and the counterfactual experiments are all minor.

(2) Policies on working migration

China has introduced the *hukou* registration system to control rural-urban migration since the early 1950s. In the 1980s and earlier, it was hard to live in urban area without urban *hukou*. Thus, the *hukou* system has become a restriction on working migration. In the counterfactual experiments here, two scenarios are explored. First, we consider the case that if China had relaxed the control over the *hukou* system and allowed more rural migrants to work in urban area since 1980. Specifically, π in the first regime is set to be equal to that of the second regime, meaning that the first regime has the same regulation on rural-urban migration as that in the second regime. The second scenario is the opposite way: what China would be if Chinese government had not loosened the control over rural-urban migration in the 1990s? To conduct this case, π in the second regime is set to be equal to that of the first regime, implying that the regulation on rural-urban migration in second regime is still as strict as that in the first regime. The results are reported in Table 8 and Table 9. Figure 11 and Figure 12 plot the simulated series for these two cases.

We find that a more relaxing migration policy in the first regime is good for total output, total output per capita, the share of urban output, and the share of urban employment. More migrants result in a higher urbanization rate. However, because the relaxing migration policy encourages migrants of low-skilled workers, the share of high-skilled employment in urban area declines and urban per capita output is lower. In the second scenario, the strict regulation on working migration in the second regime results in fewer migrants. Therefore, urbanization rate declines, total output is lower, and the share of urban output is smaller. Fewer low-skilled migrants lead to a higher urban per capita output. The results also indicate that when the proportion of low-skilled workers in urban area is smaller, the marginal productivity of high-skilled workers will be lower, leading to a lower high-skilled wage and depressing rural parents' incentive to send their children to college.

(3) Other policies

The results in the two-regime calibration show that the migration cost as a percentage of rural wage goes up from 7.27% to 11.24%. The increase in the migration cost could be attributed to the increasing living cost in urban area as pointed out by Zhang and Song (2003). However, in the *hukou* reform in the 1990s (for example, the blue-print *hukou*), several local government actually set up favorable rules to encouraged high-skilled and professional workers to move to cities. To consider this case, ,we therefore perform a counterfactual experiment by assuming that the migration cost (as a percentage of rural wage) remains unchanged in the second regime. Specifically, σ/w_R in the second regime is exactly equal to the value in the first regime. The results are summarized in Table 8 and 9. Simulated series are plotted in Figure 13 and 14. We find that lower migration cost provides incentive for rural parents to send their children to college. Therefore, the policy encourages the migration through zhaosheng and has positive influences on the economy. Our model assumes that working migration is determined by an exogenous probability. There is no decision rules for working migration, so the lower migration cost has no or very minor impacts on working migration.

Finally, we examine the effects of TFP. First, we study the scenario that the relative TFP of urban to rural remains unchanged. To conduct this experiment, we reset urban TFP in regime 2 according to the average relative TFP in the first regime. In other words, urban TFP in regime 2 becomes lower than that in the benchmark but it is still increasing so that the relative TFP is constant. In the second scenario, we reset both urban and rural TFP in regime 2 to be the values in regime 1. The results are summarized in Table 8 and 9. Simulated series are plotted in Figure 15-18. In the first experiment, urban TFP is lower so educational migration becomes less attractive to rural parents. As a result, the share of high-skilled employment declines and the share of urban employment decreases. They both have negative effects on urban production. The story of the second experiment is slightly different. Because both urban and rural TFP return to the level of regime 1, staying in rural area may not be a better choice. Thus, educational migration is still attractive and the influence on urbanization rate is minor. However, both urban and rural TFP decline, thus we observe a large drop of urban per capita output and total output level in Figure 18.

5 Conclusions

The rapid economic growth and the associated fast urbanization in China have aroused great interests in academia recently. In this paper, we have explored the economic impacts of the migration on the Chinese economy, with a special focus on the nonconventional migration channel: education-based migration, or zhaosheng. Besides education-based migration, we study the conventional work-based migration in China as well. More specifically, we develop a dynamic equilibrium model with parental educational choice on children's higher education. We conduct decomposition analysis to study the effects of education-based and work-based migration on the development process of China, and perform policy experiments on various migration regulations, such as the shifts in policies regarding education-based migration, reductions in the regulations on work-based migration, reductions in urban living and housing costs, as well as setback of the urban TFP in the post 1995 regime.

We find that migration indeed contributes greatly to the development of China: rural-urban migration accounts for near 10 percent of per capita output during 1981-2007, and the effects of education-based migration and work-based migration are roughly comparable, explaining 3.4 and 4.9 percent of output per capita throughout 1981-2007, respectively. Rich interactions between education- and work-based migration are found as well. We find that education-based migration always depresses the skill premium, while work-based migration always boosts up skill premium. Our results thus illustrate the important role played by education migration in the development of China.

Our counterfactual analysis shows that the continuation of the government job assignment scheme for college graduates would not harm the economy much, and thus our results provide justification for the abolishment of the government job assignment scheme in 1994.

Along thess lines, it would be interesting to extend the framework developed here to study various issues on migration for developing countries. A natural extension is to allow for urban unskilled workers to accumulate their human capital as in Lucas (2004). This will further enhance the importance of the education-based channel of migration. Another exntension is to examine different underlying channels of the work-based migration. In particular, the early sample stage of zhaogong channel into the state-owned enterprises and the later stage blue-stamp channel into the private sector. This requires a generalization of the urban production into two sectors, namely, state-owned and private.

References

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

[2] Arnadottir, Helga, 2012, "Hukou and Educational Barriers: Discrimination for Migrant Children," BA Degree of Chinese Studies Thesis. University of Iceland: Iceland.

[3] Autor, David H., Lawrence F. Katz and Alan B. Krueger, 1998, "Computing inequality: have computers changed the labor market?" *The Quarterly Journal of Economics*, Vol. 113, No. 4, pp. 1169-1213.

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

[5] Balvers R. J. and J. H. Bergstrand, 2002, "Government expenditure and equilibrium real exchange rates," *Journal of International Money and Finance*, 21, pp. 667–692.

[6] Barro, R. J. and J. Lee, 2001, "International data on educational attainment updates and implications," *Oxford Economic Papers* 53(3), pp.541-563.

[7] Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum, 2003, "Plants and productivity in international trade," *American Economic Review*, 93(4), pp. 1268-1290.

[8] Brandt, L., J. Huang, G. Li and S. Rozelle, 2002, "Land rights in rural China: facts, fictions and issues," *The China Journal*, 47, pp. 67-97.

[9] Chan, K.W. and W. Buckingham, 2008, "Is China abolishing the *hukou* system?" *The China Quarterly*, 195, pp. 582-606.

[10] Chan, K.W. and L. Zhang, 1999, "The *hukou* system and rural-urban migration in China: processes and changes," *The China Quarterly*, 160, pp. 818-855.

[11] Chang, R., 1998, "Political party negotiations, income distribution, and endogenous growth," *Journal of Monetary Economics*, 41, pp. 227-255.

[12] Ciccone, A. and G. Peri, 2005, "Long-run substitutability between more and less educated workers: evidence from U.S. states, 1950-1990," *The Review of Economics and Statistics*, 87(4), pp. 652-663.

[13] Eaton, J. and S. Kortum, 2002, "Technology, geography and trade," *Econometrica*, 70, pp. 1741–1779.

[14] 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 J. Poterba, NBER, MIT Press.

[15] Ge, Suqin and Dennis Tao Yang, 2014, "Changes in China's wage structure," *Journal of the European Economic Association*, 12(2), pp. 300-336.

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

[17] Hall, Robert E., 1989, "Intertemporal substitution in consumption," *Journal of Political Economy*, 96(2), pp. 339-357.

[18] Hare, Denise, 1999, " 'Push' versus 'pull' factors in migration outflows and returns: Determinants of migration status and spell duration among China's rural population," *The Journal of Development Studies*, 35(3), 45-72.

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

[20] Hertel, T. and F. Zhai, 2006, "Labor market distortions, rural-urban inequality and the opening of China's economy," *Economic Modelling*, 23, pp. 76-109.

[21] Hu, Dapeng, 2002, "Trade, rural-urban migration, and regional income disparity in developing countries: a spatial general equilibrium model inspired by the case of China," *Regional Science and Urban Economics*, 32, pp. 311-338.

[22] Hu, Feng, Zhaoyuan Xu, and Yuyu Chen, 2011, "Circular migration, or permanent stay? Evidence from China's rural–urban migration," *China Economic Review* 22(1), pp. 64–74.

[23] Knight, J. and R. Gunatilaka, 2010, "Great expectations? The subjective well-being of rural-urban migrants in China," *World Development*, 38(1), pp. 113-124.

[24] 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.

[25] Li, Hongbin, Lei Li, Binzhen Wu, and Yanyan Xiong, 2012, "The end of cheap Chinese labor," *Journal of Economic Perspectives* 26(4), pp. 57–74.

[26] Liao, Pei-Ju., 2013, "The one-child policy: a macroeconomic analysis," *Journal of Development Economics*, 101, pp. 49-62.

[27] Liu, Z., 2005, "Institution and inequality: the *hukou* system in China," *Journal of Comparative Economics*, 33, pp. 133-157.

[28] Liu, Zhiqiang, 2008, "Human capital externalities and rural-urban migration: evidence from rural China," *China Economic Review*, 19, pp. 521-535.

[29] Liu, Jialu, 2011, "Human capital, migration and rural entrepreneurship in China," *Indian Growth and Development Review*, 4(2), pp. 100-122.

[30] Lu, Z. and S. Song, 2006, "Rural-urban migration and wage determination: The case of Tianjin, China," *China Economic Review*, 17, pp. 337-345.

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

[32] 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.

[33] Mayer, T., M. J. Melitz , and G. Ottaviano, 2014, "Market size, competition, and the product mix of exporters," *American Economic Review,* 104 (2), pp. 495-536.

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

[35] Melitz, Marc J. and Stephen J. Redding, 2014, "Missing Gains from Trade?" NBER Working Paper No. 19810.

[36] Mullan, Katrina, Pauline Grosjean and Andreas Kontoleon, 2011, "Land tenure arrangements and rural–urban migration in China," *World Development* 39(1), pp. 123–133.

[37] O'Rourke, K. h., A. S. Rahman and A. M. Taylor, 2013, "Luddites, the industrial revolution, and the demographic transition," *Journal of Economic Growth*, 18, pp. 373-409.

[38] Rozelle, S., Taylor, J. E., and A. deBrauw, 1999, "Migration, remittances, and agricultural productivity in China," *American Economic Review,* 89(2), pp. 287-291.

[39] Stark, Oded. *The migration of labor.* Cambridge, MA: Basil Blackwell, 1991.

[40] Tan, Guangyu, 2010, "Under the same blue sky? inequity in migation children's education in China," *Current Issues in Comparative Education*, 12(2), pp. 31-40.

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

[42] Toh, Ruby and Hui Weng Tat, 2012, "Trade liberalization, labor demand shifts and earnings inequality in Singapore," *Review of Urban and Regional Development Studies*, Vol. 24, No. 3, pp . 65-82.

[43] Wang, H., J. Tong, F. Su, G. Wei and R. Tao, 2011, "To reallocate or not: reconsidering the dilemma in China's agricultural land and tenure policy," *Land Use Policy*, 28, pp. 805-814.

[44] Whalley, J. and S. Zhang, 2007, "A numerical simulation analysis of (Hukou) labour mobility restrictions in China," *Journal of Development Economics*, 83, pp. 392-410.

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

[46] Xia, Qingjie, 2004, "How the Chinese rural workers choose occupation: A case study on nine villages in the northeast China," UK: Department of Economics, University of Bath.

[47] Yang, Dennis Tao, 1997, "China's land arrangements and rural labor mobility," China Economic Review, 8(2), pp. 101-115.

[48] Ye, W. Z. and Y. Ding, 1998, "The cost of child care in xiamen special economic zone in China," *Population and Economics*, No.6, pp. 24-57 (in Chinese).

[49] Zhang, J., Y. Zhao, A. Park, and X. Song, 2005, "Economic returns to schooling in urban China 1988 to 2001," *Journal of Comparative Economics*, 33, pp. 730-752.

[50] Zhao, Y., 1997, "Labor migration and returns to rural education in China," *American Journal of Agricultural Economics*, Vol. 79, No. 4, pp. 1278-1287.

[51] Zhao, Y., 1999a, "Leaving the countryside: Rural-to-ruban migration decisions in China," *American Economic Review*, 89(2), pp. 281-286.

[52] ——— 1999b, “Labor migration and earnings differences: The Case of Rural China,” *Economic Development and Cultural Change*, 47(4), pp. 767-782.

[53] Zhao, Y., 2003, “The role of migrant networks in labor migration: The case of China,” *Contemporary Economic Policy*, 21(4), pp. 500-511.

[54] Zhao, Q., X. Yu, X. Wang and T. Glauben, 2014, “The impact of parental migration on children’s school performance in rural China,” *China Economic Review*, 31, 43-54.

[55] Zhu, N., 2002, “The impacts of income gaps on migration decisions in China,” *China Economic Review*, 13, pp. 213-230.

[56] 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).

Appendix A: Data sources and data computation

1. Population

(1) Rural and urban population

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

(2) High-skilled and low-skilled workers

China Labour Statistical Yearbook reported educational attainment composition of urban employment (as a percentage of total urban employment). Thus, workers whose educational attainment are college and above are defined to be high-skilled workers. However, urban data is only available during 2002-2007. Thus, we first use 2002-2007 data to compute an urban to national-wide ratio (a ratio of educational attainment composition of urban employment to that of the whole country). The ratio is about 2.457. Second, for the year of 1982, 1990, 1995-1999, and 2001, the fraction of high-skilled workers as a percentage of total urban employment is computed using the national-wide data and is adjusted by the urban to national-wide ratio. The national-wide data of 1996-1999 and 2001-2007 are also from *China Labour Statistical Yearbook*. The data of 1982 is from *1 Percent Sampling Tabulation on the 1982 Population Census of the People's Republic of China*. The data of 1990 is from *China Population Statistical Yearbook* 1994. The data of 1995 is available in *1995 China 1% Population Sampling Survey Data*. For the year of 1980, 1985, and 2000, the educational attainment for total population in Barro and Lee (2000) is adjusted by the urban to national-wide ratio to obtain the N_H/N_U. Third, we do intrapolation for the years that 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 is the simple average of 1980-1994 and for regime 2 is the average of 1995-2007.

(3) The flow of rural to urban migration

There is no available national-wide survey on rural to urban migration in China. Here, we use changes in urban population as a proxy for the total amount of rural to urban migrants. The total amount of rural to urban migrants is then divided by the stock of rural population to obtain the flow of migrants (as a percentage of rural population). In the calibration, we take the simple average on the flow of migrants from 1981-1994 to be the flow of migrants in the first regime. The second regime is the average of 1995-2007. Finally, the average flows of migrants are multiplied by working-related reasons and studying or training (average of 1985 and 2000) to obtain the probability of working migration and *zhaosheng* flow, respectively.

(4) Migration reasons

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 the 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. *Tabulation on the 2000 Population Census of the People's Republic of Chin*a only reported the number of emigration and reasons for emigration. We thus choose the number of emigration from town and county to be the rural to urban migration in 2000. Then, the fraction of migrants due to each reason as a percentage of total rural to urban migration is computed. Finally, we categorize migration due to job transfer, job assignment, and work or business as working-related reason. The migration due to study or training is categorized to be migration via *zhaosheng*.

2. Human capital

Table A.1 summarizes average years of schooling for college and above or below college. The urban employment by education in 1995 is from *China Statistical Yearbook* 1998. 2002 and 2009 data are from *China Labour Statistical Yearbook* 2002 and 2009, respectively. We further assume the years of schooling for graduate is equal to 18 years, 16 years to college, 14 years to junior college, 12 years to senior high, 9 years to junior high, 6 years to primary school, and 1 year to semi-illiterate or illiterate. Then, weighted average years of schooling for college or above and for below college is computed. Table A.2 is the average years of schooling for 1981, 1988, 1995, and 2002. For those 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 return coefficients in China reported by Zhang et al. (2005) are 0.0479 and 0.0835 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 is available. Thus, we use urban employment rate as a proxy. To obtain urban employment rate, we first compute the urban unemployment rate. The urban unemployment rate is the total registered unemployed persons (only for those who held non-agricultural hukou) divided by total labor force. Only unemployed persons and labor force in city districts are included. Then, urban employment rate is calculated. The urban employment rate in the calibration is the simply average of 2000-2007. These data are all from *China City Statistical Yearbook*.

4. Labor income shares

Bai and Qian (2010) reported the sectoral labor share in GDP in 1978-2004 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 are belong to 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 in regime 1 is the average of 1980-1994, and regime 2 is the average of 1995-2004.

5. Rural income

China Statistical Yearbook 2011 reported rural real income per capita for rural household from 1978 to 2011. However, during the periods before 1990, only 1978, 1980, and 1985 are available. We thus use intrapolation 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 be one. The rural real income per capita of other years are adjusted accordingly. In the calibration, the rural income of regime 1 is the average of rural real income per capita during 1980-1994. 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 1988-2001. Ge and Yang (2014) estimate that for 1992-2007. Using the ratio of skill premium in Zhang et al. (2005) to Ge and Yang (2014), we construct a time series of skill premium during 1988-2007 based on Zhang et al. (2005). Furthermore, Lee (1999) estimate 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 the survey of SOEs. Therefore, we first compute the growth rate of skill premium from 1980 to 1988 in Lee (1999). Then, using the estimate of skill premium in 1988 in Zhang et al. (2005) and the growth rate computed from Lee (1999), the skill premium of 1980 is obtained. Finally, curve fitting with polynomial 3 is used to construct a series of 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

Urban premium is defined as the ratio of low-skilled wage to rural wage. *China Statistical Yearbook 2011* also reported urban real income per capita during 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 high-skilled wage and low-skilled wage, we are now able to compute urban premium using the data of skill premium, the ratio of urban to rural income per capita, and the relative stock of high-skilled workers to low-skilled workers. However, during the periods before 1990, only 1978, 1980, and 1985 are available. We thus use intrapolation to compute the urban premium for 1981-1984 and 1986-1989.

8. Computed 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, the computed data for urban per capita output is the computed data for urban output divided by the time series data of high- and low-skilled workers.

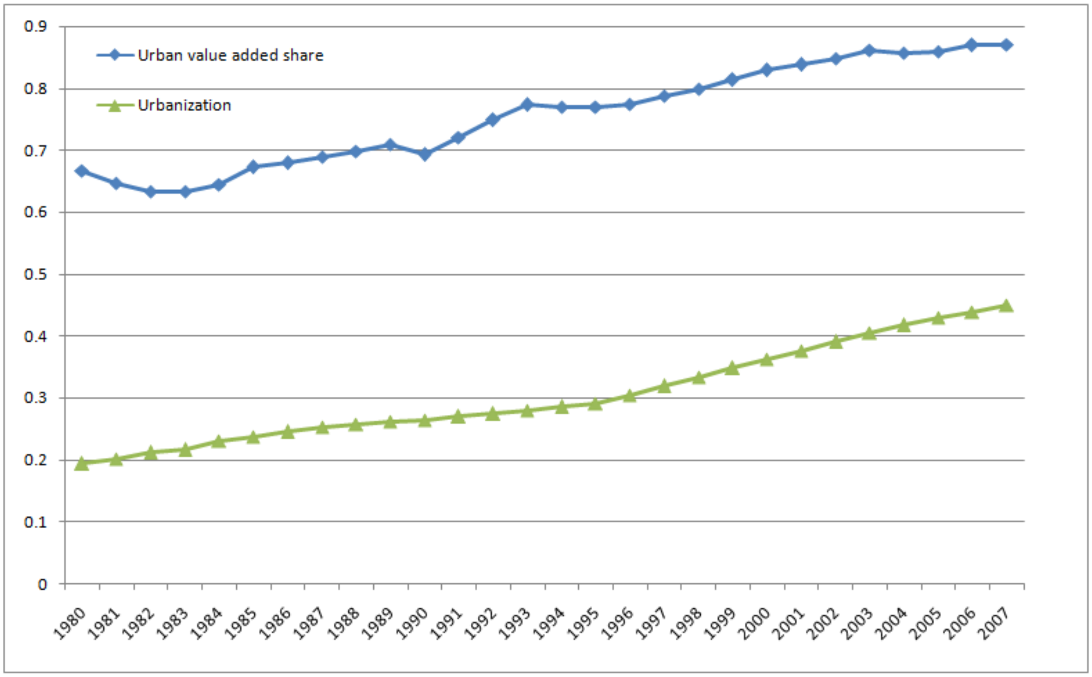
(2) Rural output

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

(3) Total output

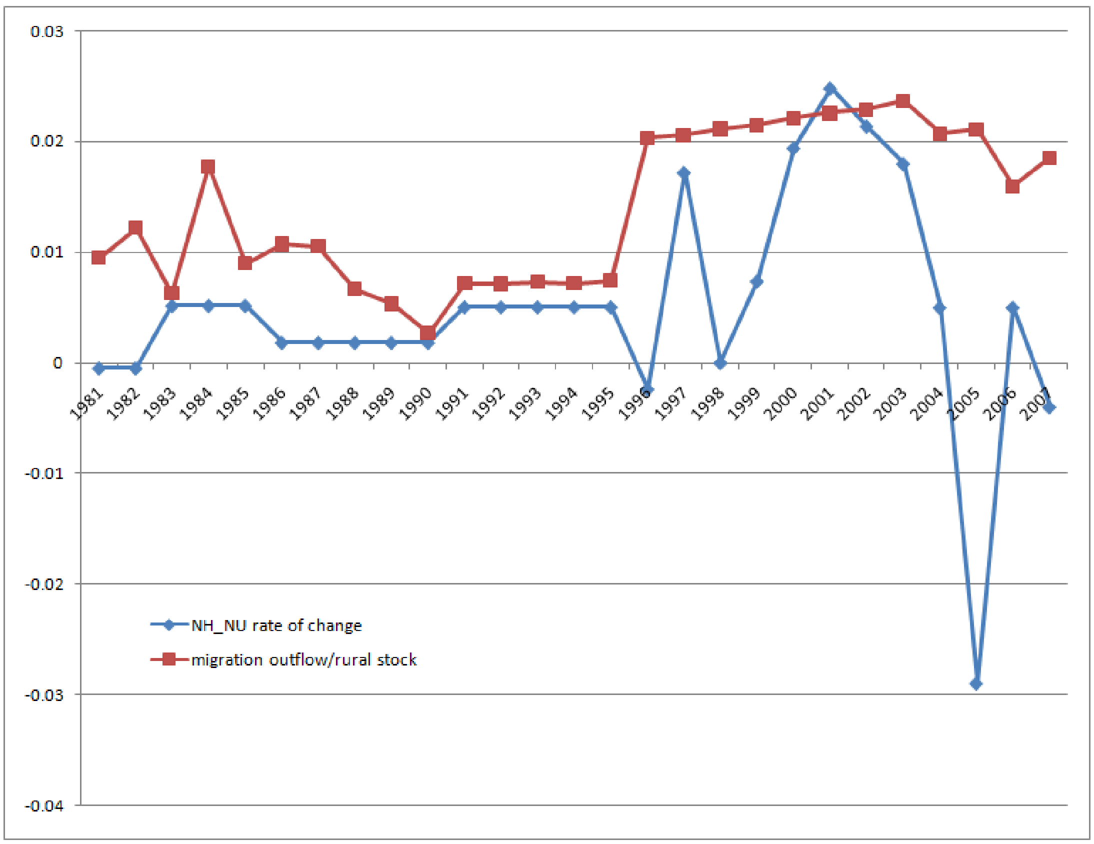
The computed total output is the sum of the computed data for urban output and rural output.

Figure 1: Urbanization rates and urban output shares over 1980-2007



Note: Urbanization rate is defined as urban population shares out of total population. Urban output shares are computed based on Bai and Qian (2010), excluding agricultural sector.

Figure 2: Migration flows over rural population and changes in urban tertiary education employment shares



Note: The series NH_NU rate of change refers to the changes in the urban tertiary education employment shares. As there is no good data on migration, we use changes in urbanization as a proxy for migration outflow.

Figure 3: Timeline of the model

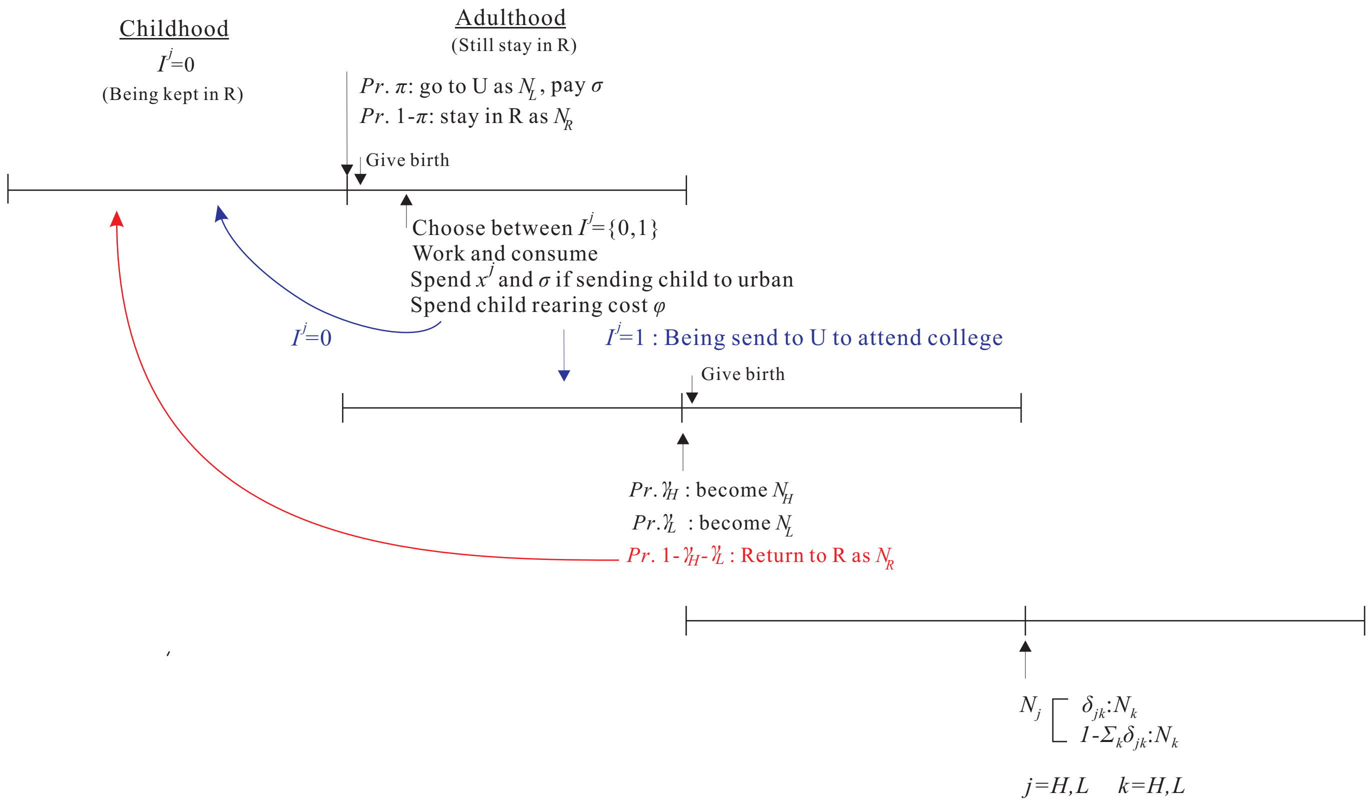


Figure 4: High- and low-skilled wages and rural wage rate

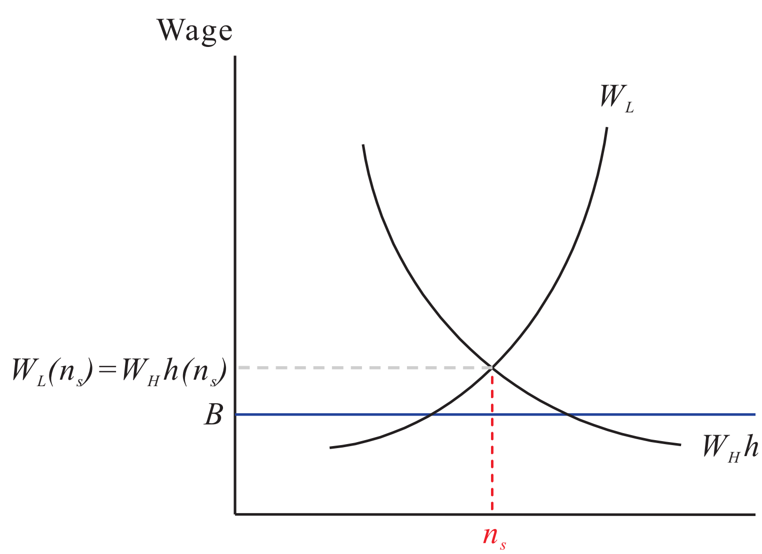


Figure 5: Calibrated urban and rural TFP over 1981-2007

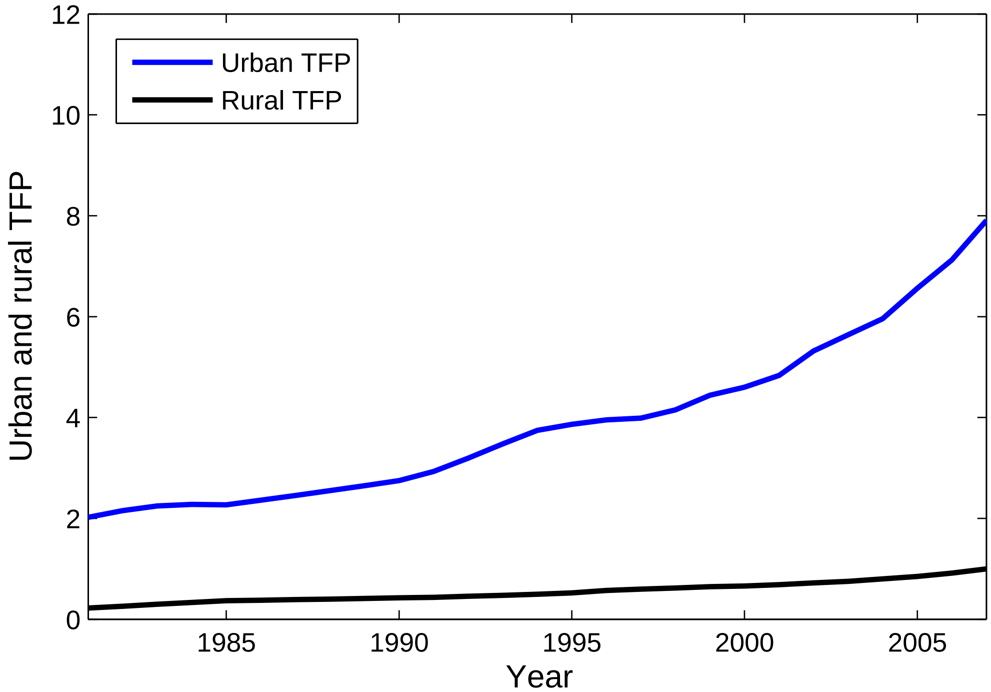


Figure 6: Benchmark model - output

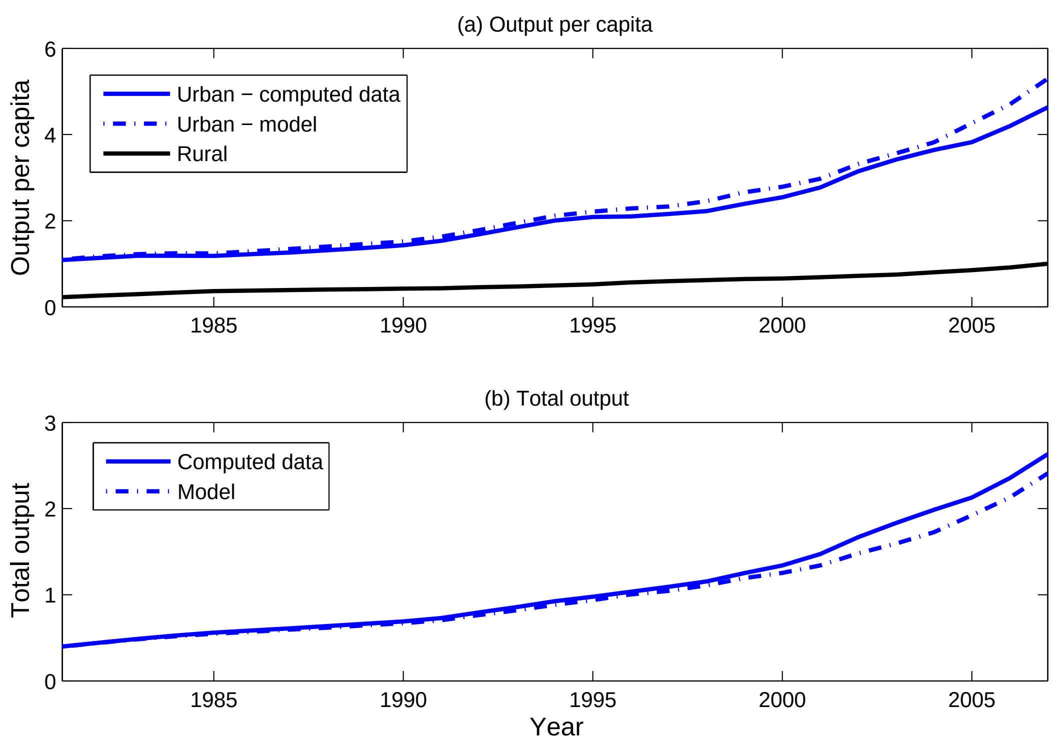


Figure 7: Benchmark model - urbanization rate and labor share

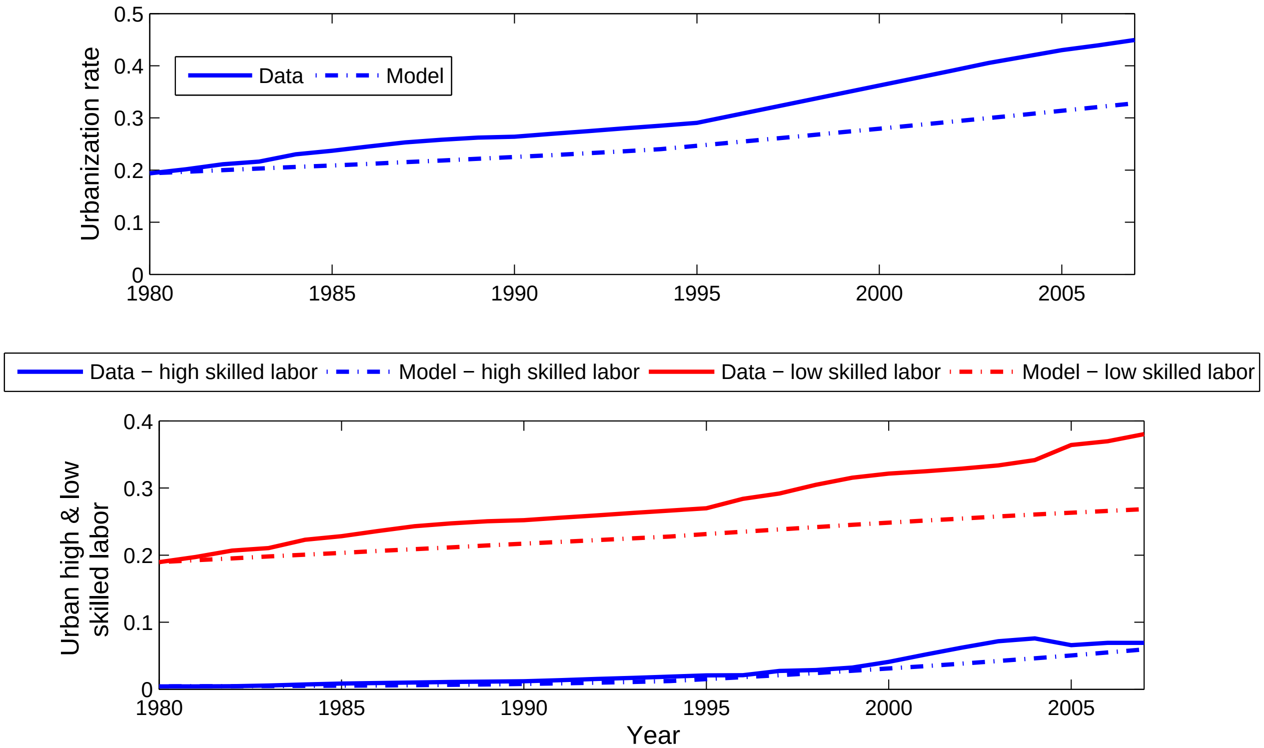


Figure 8: Decomposition

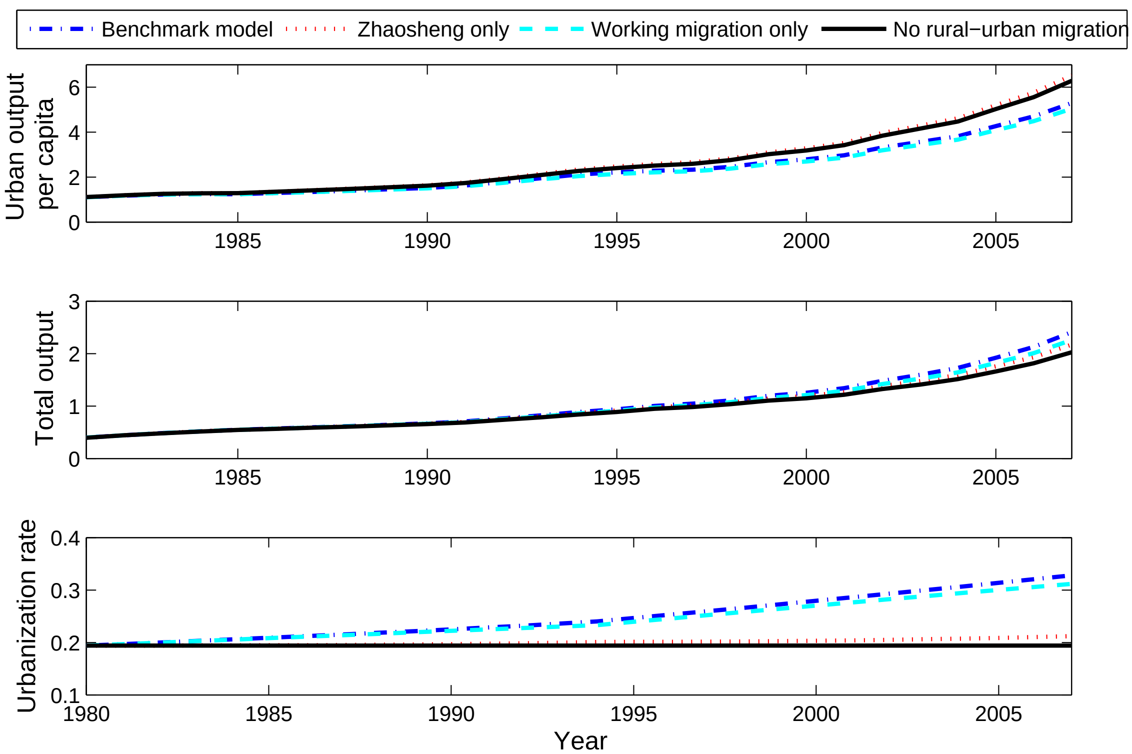


Figure 9: GJA policy - output

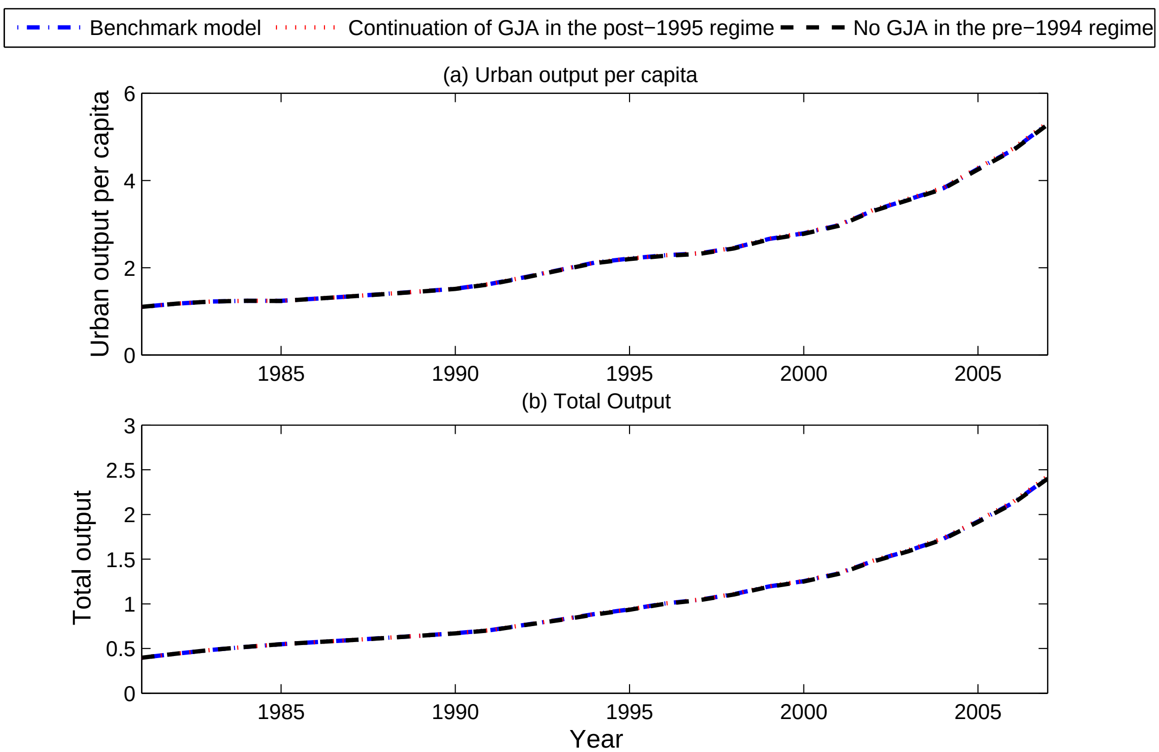


Figure 10: GJA policy - urbanization rate and labor share

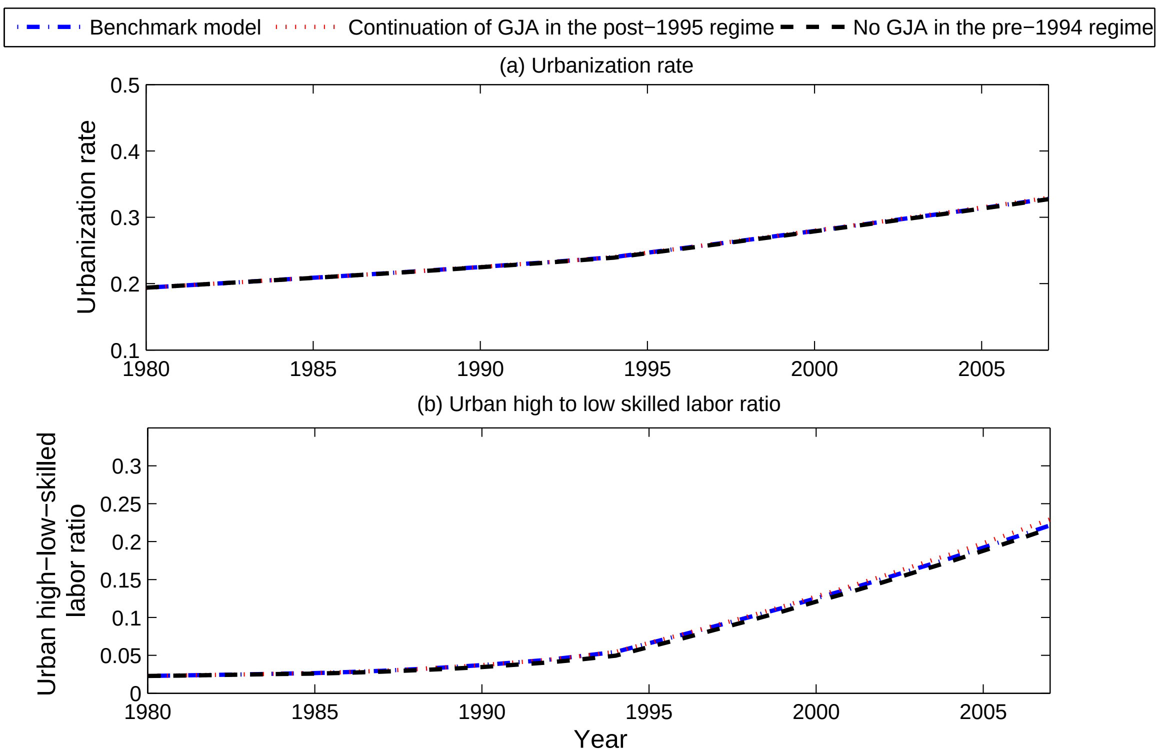


Figure 11: Policy on working migration - output

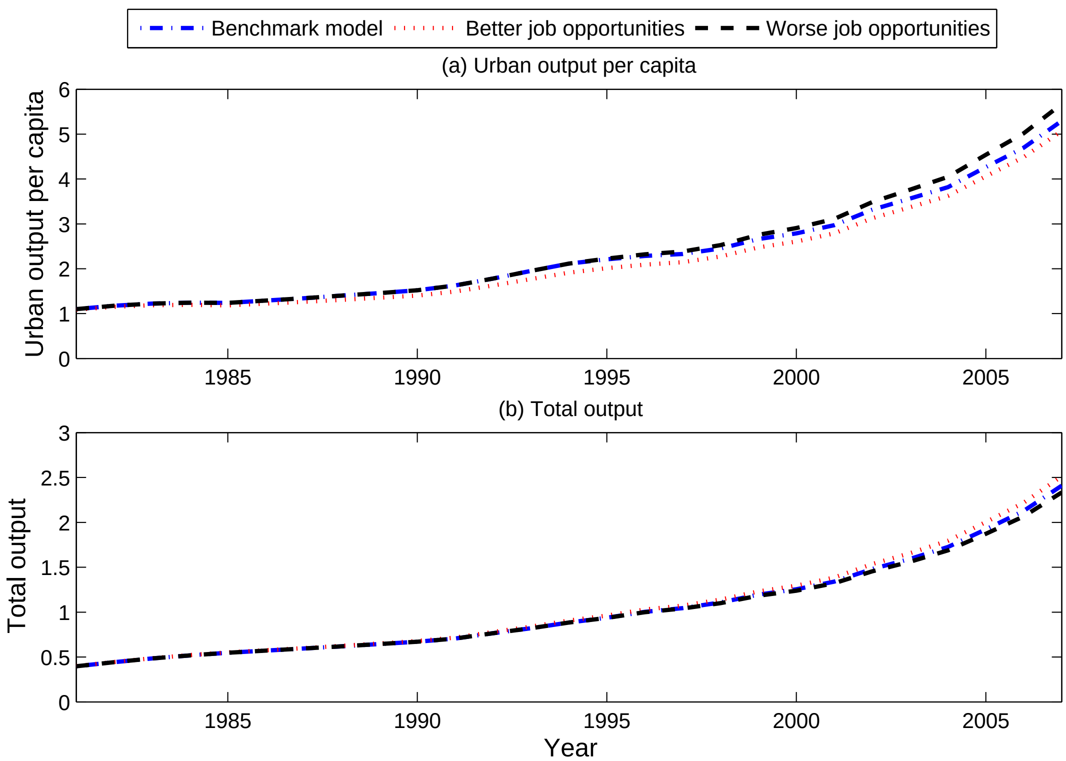


Figure 12: Policy on working migration - urbanization rate and labor share

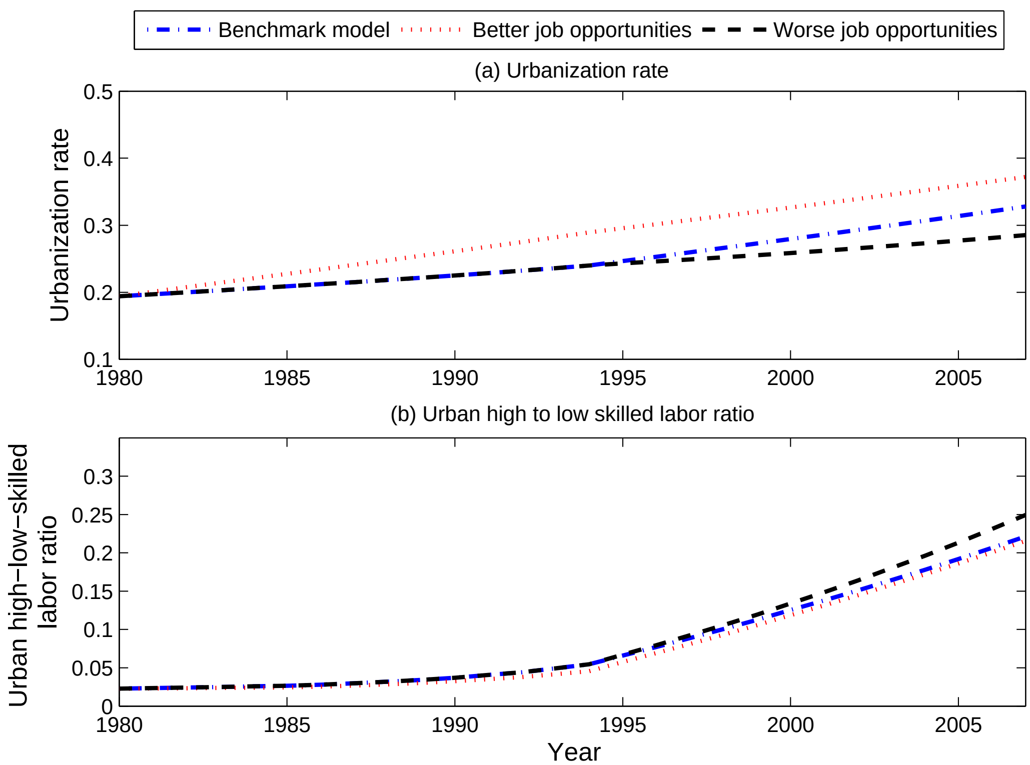


Figure 13: Policy on migration cost - output

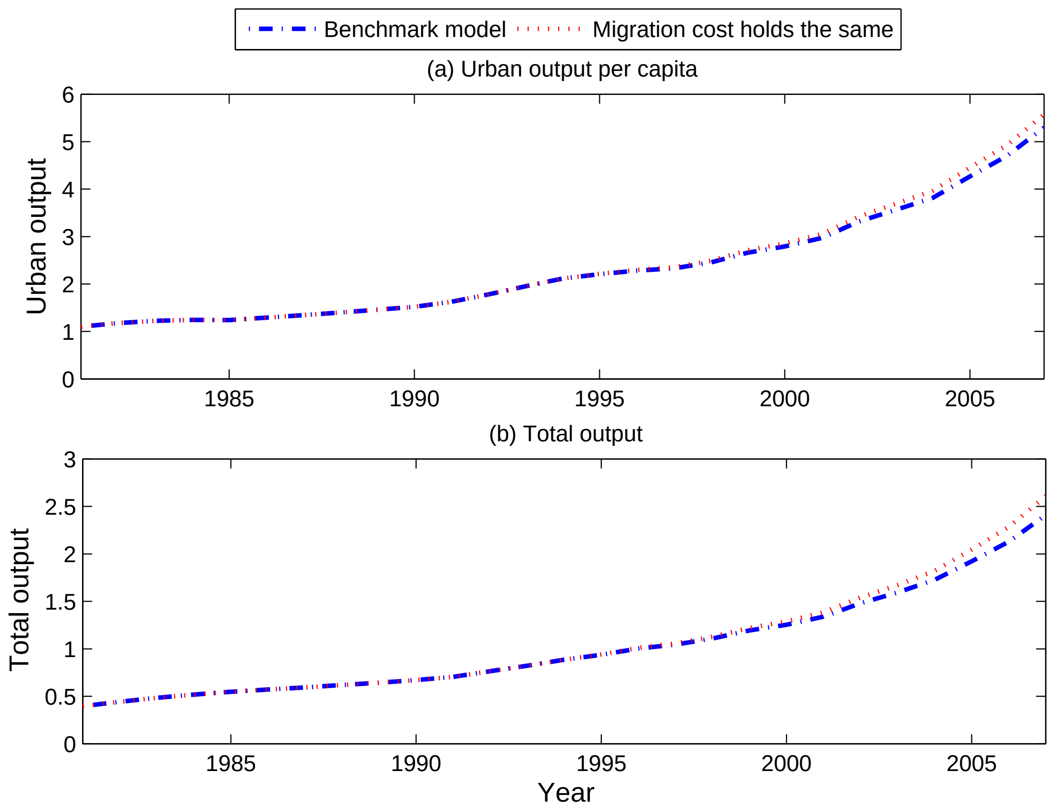


Figure 14: Policy on migration cost - urbanization rate and labor share

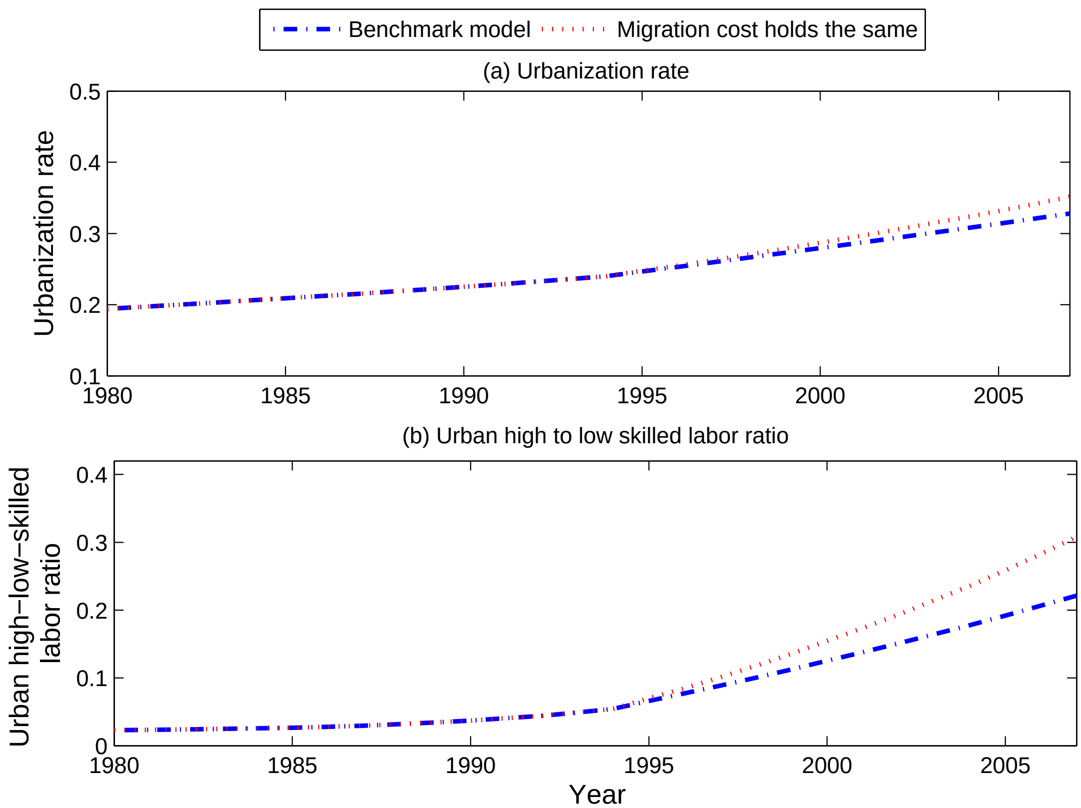


Figure 15: Relative TFP remains unchanged - output

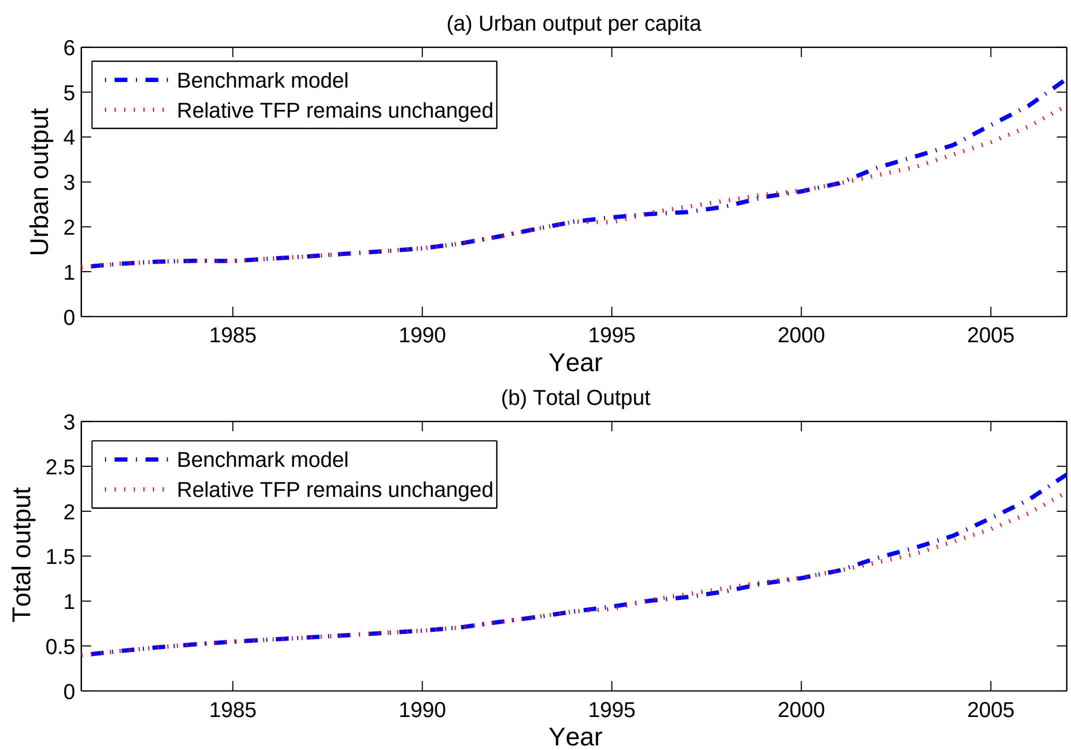


Figure 16: Relative TFP remains unchanged - urbanization rate and labor share

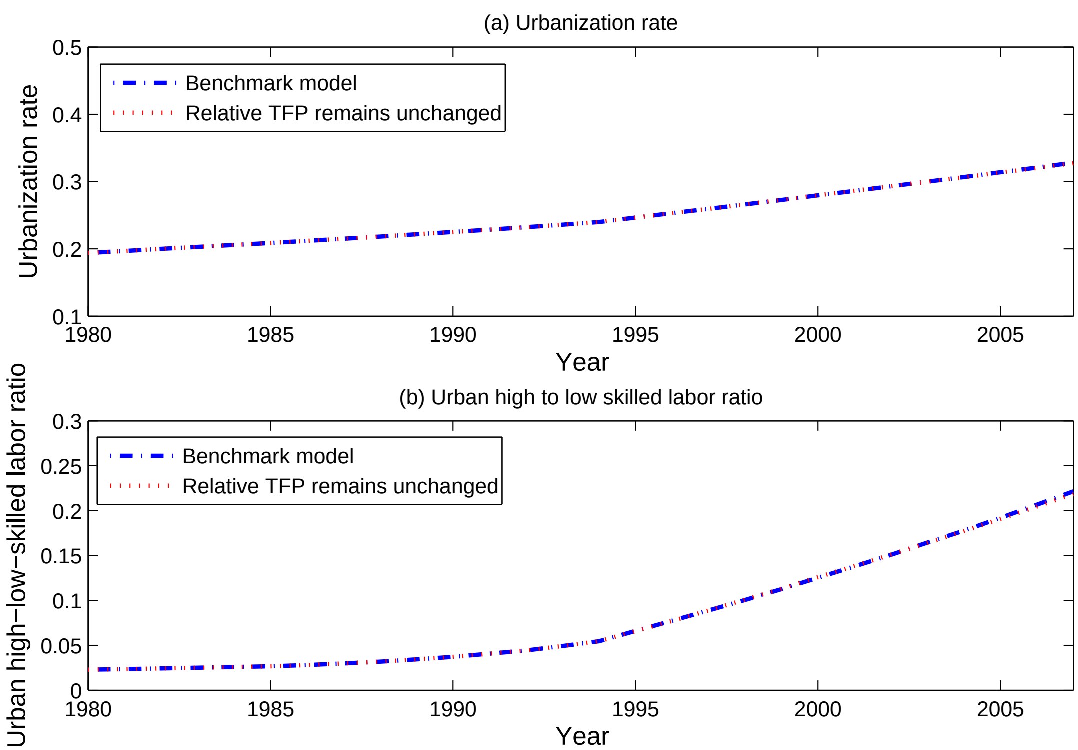


Figure 17: TFP in both areas remain unchanged - output

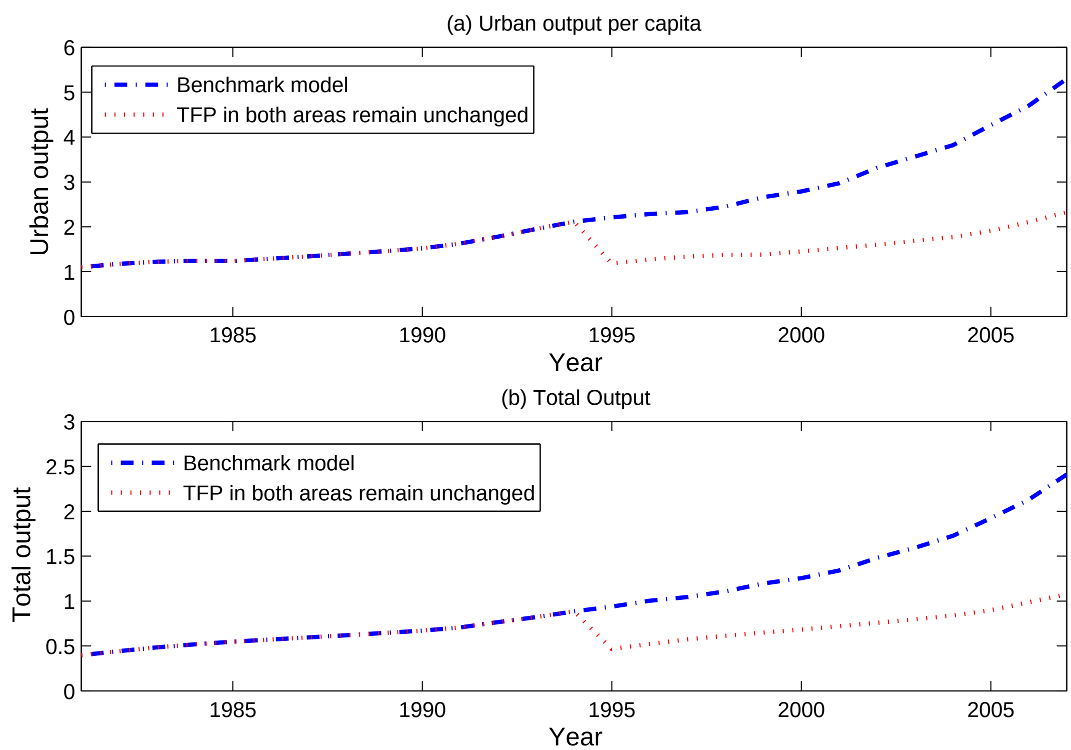


Figure 18: TFP in both areas remains unchanged - urbanization rate and labor share

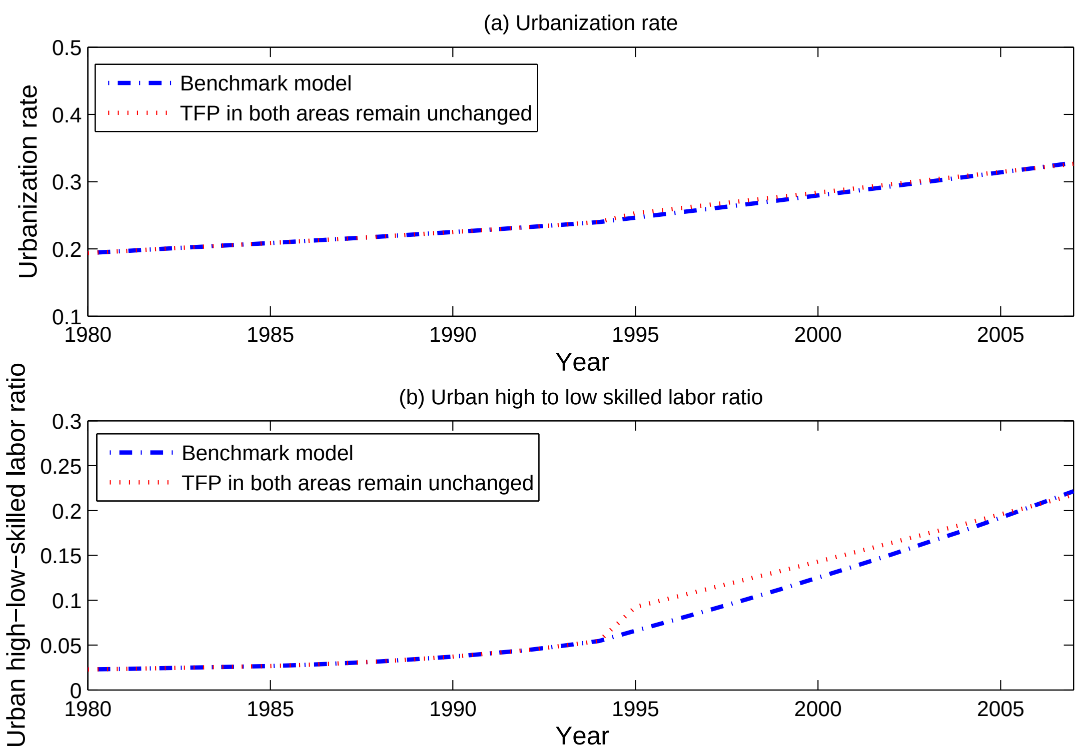


Table 1: Migration by reasons

Cause of Migration	Total	Job Transfer	Job Assignment	Work or Business	Study or Training	Others
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%

Note:

1. Source: Authors' computation based on 1990 census and 2000 census data.

2. In the census, 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 working migration, and migration due to study or training as migration via zhaosheng.

Table 2: Zhaosheng flow and the probability of working migration

	Zhaosheng flow	Prob. of working migration
Pre-1994	0.00058946	0.003554486
Post-1995	0.00114381	0.008281515

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

Table 3: Parameters

	Regime 1	Regime 2	Target
Preset parameters			
β	0.3384	0.3384	Liao (2013)
ε	2	2	assumption
ϕ	0.174	0.174	Zhu and Zhang (1996)
z_{min}	1	1	assumption in the literature
θ	1.92	1.92	Feenberg and Poterba (1993)
ρ	0.5	0.5	assumption
B	0.3685	0.7177	$w_R = 1$ in 2007 and use the growth rate of w_R in data
γ_L	0	0.05	assumption
γ_H	1	0.9127	urban employment rate in regime 2 is 0.9627
δ_{HH}	1	1	assumption
π	0.0036	0.0083	see Table 2
h	1.3381	1.5922	Mincerian rate of return equation
Calibrated parameters			
α	0.6762	0.6762	match (1-labor income share) in two regimes
ψ	0.0562	0.0562	match (1-labor income share) in two regimes
δ_{LL}	0.9996	0.9885	match regime average N_H/N_L
σ	0.0268	0.0807	solve from the indifference boundary equations
A	2.7365	5.5010	match regime average w_L/w_R
τ	2.6960	1.7960	match regime average $w_H h/w_L$

Note: Calibrated results are not sensitive to the value of γ_L in regime 2.

Table 4: Calibrated results

	Regime 1	Regime 2
σ/w_R	0.0727	0.1124
$\hat{z}$	42.36	28.61
A/B	7.425	7.665
Annual growth rate of A, 1981-2007		5.31%
Annual growth rate of A/B, 1981-2007		0.24%

Table 5: 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	1.0299	0.5757	0.2507	0.0754	1.4759
Regime 1: 1981-1994	0.6189	0.5162	0.2174	0.0327	1.2678
Regime 2: 1995-2007	1.4724	0.6399	0.2866	0.1213	1.6999

Table 6: Decomposition - level

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 $(w_H h/w_L)$
Shut down working migration but still allow zhaosheng					
Whole: 1981-2007	0.9793	0.5356	0.2004	0.0905	1.3224
Regime 1: 1981-1994	0.6121	0.4984	0.1962	0.0362	1.2016
Regime 2: 1995-2007	1.3748	0.5757	0.2050	0.1490	1.4525
Shut down zhaosheng but still allow working migration					
Whole: 1981-2007	0.9946	0.5629	0.2445	0.0538	1.5426
Regime 1: 1981-1994	0.6117	0.5102	0.2151	0.0227	1.2937
Regime 2: 1995-2007	1.4070	0.6195	0.2761	0.0874	1.8107
No migration					
Whole: 1981-2007	0.9459	0.5209	0.1939	0.0621	1.3818
Regime 1: 1981-1994	0.6054	0.4924	0.1939	0.0250	1.2249
Regime 2: 1995-2007	1.3126	0.5516	0.1939	0.1021	1.5509

Table 7: Decomposition - 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	3.4%	7.0%	20.0%	28.6%	-4.5%
Regime 1: 1981-1994	1.2%	1.2%	1.1%	30.6%	-2.0%
Regime 2: 1995-2007	4.4%	3.2%	3.7%	28.0%	-6.5%
Working migration					
Whole: 1981-2007	4.9%	7.0%	20.0%	-20.1%	10.4%
Regime 1: 1981-1994	1.1%	3.4%	9.7%	-10.7%	5.2%
Regime 2: 1995-2007	6.6%	10.0%	28.5%	-22.8%	14.6%
Interactive migration					
Whole: 1981-2007	-0.2%	-4.4%	-17.4%	9.0%	0.5%
Regime 1: 1981-1994	-0.1%	0.0%	0.0%	3.6%	0.2%
Regime 2: 1995-2007	-0.2%	0.6%	0.2%	10.6%	0.7%
Non-migration factors					
Whole: 1981-2007	91.8%	90.5%	77.3%	82.5%	93.6%
Regime 1: 1981-1994	97.8%	95.4%	89.2%	76.5%	96.6%
Regime 2: 1995-2007	89.1%	86.2%	67.7%	84.2%	91.2%

Table 8: Counterfactual analysis - level

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 $(w_H h/w_L)$
Benchmark model					
Whole: 1981-2007	1.0299	0.5757	0.2507	0.0754	1.4759
Regime 1: 1981-1994	0.6189	0.5162	0.2174	0.0327	1.2678
Regime 2: 1995-2007	1.4724	0.6399	0.2866	0.1213	1.6999
Policy on GJA					
1. Continue the GJA in regime 2: $\gamma_{H,2}$=1					
Whole: 1981-2007	1.0324	0.5764	0.2510	0.0766	1.4720
Regime 1: 1981-1994	0.6189	0.5162	0.2174	0.0327	1.2678
Regime 2: 1995-2007	1.4778	0.6412	0.2872	0.1238	1.6918
2. No GJA in regime 1: $\gamma_{H,1}$=$\gamma_{H,2}$					
Whole: 1981-2007	1.0265	0.5744	0.2502	0.0728	1.4837
Regime 1: 1981-1994	0.6177	0.5153	0.2171	0.0310	1.2724
Regime 2: 1995-2007	1.4668	0.6381	0.2859	0.1178	1.7113
Policy on working migration					
1. Better job opportunities in regime 1: π_1 increases to π_2					
Whole: 1981-2007	1.0595	0.6035	0.2872	0.0710	1.5661
Regime 1: 1981-1994	0.6267	0.5379	0.2445	0.0293	1.3469
Regime 2: 1995-2007	1.5257	0.6742	0.3331	0.1160	1.8022
2. Worse job opportunities in regime 2: π_2 reduces to π_1					
Whole: 1981-2007	1.0170	0.5677	0.2393	0.0796	1.4410
Regime 1: 1981-1994	0.6189	0.5162	0.2174	0.0327	1.2678
Regime 2: 1995-2007	1.4458	0.6232	0.2628	0.1301	1.6276
Other policies					
1. Migration cost in regime 2 lowered to as in regime 1					
Whole: 1981-2007	1.0615	0.5835	0.2557	0.0885	1.4370
Regime 1: 1981-1994	0.6189	0.5162	0.2174	0.0327	1.2678
Regime 2: 1995-2007	1.5382	0.6561	0.2969	0.1487	1.6192
2. Set regime 2 urban TFP according to the average relative TFP ratio as that in 1981-1993					
Whole: 1981-2007	1.0080	0.5725	0.2506	0.0752	1.4763
Regime 1: 1981-1994	0.6189	0.5162	0.2174	0.0327	1.2678
Regime 2: 1995-2007	1.4271	0.6332	0.2865	0.1210	1.7008
3. Set regime 2 urban and rural TFP to those in regime 1					
Whole: 1981-2007	0.6757	0.5746	0.2523	0.0808	1.4603
Regime 1: 1981-1994	0.6189	0.5162	0.2174	0.0327	1.2678
Regime 2: 1995-2007	0.7368	0.6375	0.2899	0.1325	1.6676

Table 9: Counterfactual analysis - 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 $(w_H h/w_L)$
Policy on GJA					
1. Continue the GJA in regime 2: $\gamma_{H,2}$=1					
Whole: 1981-2007	-	-	-	-	-
Regime 1: 1981-1994	-	-	-	-	-
Regime 2: 1995-2007	0.36%	0.20%	0.22%	2.04%	-0.48%
2. No GJA in regime 1: $\gamma_{H,1}$=$\gamma_{H,2}$					
Whole: 1981-2007	-0.32%	-0.23%	-0.20%	-3.40%	0.53%
Regime 1: 1981-1994	-0.19%	-0.18%	-0.13%	-5.22%	0.36%
Regime 2: 1995-2007	-0.38%	-0.28%	-0.25%	-2.87%	0.67%
Policy on working migration					
1. Better job opportunities in regime 1: π_1 increases to π_2					
Whole: 1981-2007	2.88%	4.83%	14.56%	-5.74%	6.12%
Regime 1: 1981-1994	1.26%	4.20%	12.50%	-10.56%	6.24%
Regime 2: 1995-2007	3.61%	5.37%	16.24%	-4.34%	6.02%
2. Worse job opportunities in regime 2: π_2 reduces to π_1					
Whole: 1981-2007	-1.25%	-1.39%	-4.57%	5.62%	-2.36%
Regime 1: 1981-1994	0.00%	0.00%	0.00%	0.00%	0.00%
Regime 2: 1995-2007	-1.81%	-2.60%	-8.30%	7.26%	-4.26%
Other policies					
1. Migration cost in regime 2 lowered to as in regime 1					
Whole: 1981-2007	-	-	-	-	-
Regime 1: 1981-1994	-	-	-	-	-
Regime 2: 1995-2007	4.46%	2.53%	3.60%	22.56%	-4.75%
2. Set regime 2 urban TFP according to the average relative TFP ratio as that in 1981-1993					
Whole: 1981-2007	-	-	-	-	-
Regime 1: 1981-1994	-	-	-	-	-
Regime 2: 1995-2007	-3.08%	-1.04%	-0.04%	-0.22%	0.05%
3. Set regime 2 urban and rural TFP to those in regime 1					
Whole: 1981-2007	-	-	-	-	-
Regime 1: 1981-1994	-	-	-	-	-
Regime 2: 1995-2007	-49.96%	-0.38%	1.14%	9.24%	-1.90%

Table A.1: 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 A.2: 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.