Boston University Theses & Dissertations

2016

Essays on financial frictions, misallocation and development dynamics

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ESSAYS ON FINANCIAL FRICTIONS, MISALLOCATION AND DEVELOPMENT DYNAMICS

by

EI YANG

B.A., Renmin University of China, 2008
M.A., New York University, 2010

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2016

Approved by

First Reader ______________________________
Simon Gilchrist, PhD
Professor of Economics

Second Reader ______________________________
Jianjun Miao, PhD
Professor of Economics

Third Reader ______________________________
Robert G. King, PhD
Professor of Economics

Acknowledgments

First, I would like to thank my advisor Professor Simon Gilchrist, who provides me with endless support and guidance. The two most important things I learn from Simon are the way to think as a superior macroeconomist and the taste of excellent research. A superior macroeconomist focuses on the real world and an excellent research combines beautiful insights with empirical support. I would also like to thank Professor Jianjun Miao and Professor Robert G. King. Jianjun keeps pushing me for better research, which is challenging but a valuable experience for me. Bob always encourages me to pursue what I'm interested and most importantly, he gives me endless support for this journey. I'm so grateful to all my advisors.

Second, I would like to thank all my other professors and friends. I learn a lot from Professor Alisdair McKay both from his classes and personal discussions. Professor Dilli Mookherjee always encourages me on the topic and provides me many useful insights. I'm very glad to have Weize Chen as a true friend in the last 3 years. I have enjoyed so many conversations with him.

I'm so indebted to my parents Weiming Yang and Liangguo Huang. It is so amazing to raise a child from a small city in southwest China, who now earn a Ph.D. degree in the US. They sacrificed a lot. My mother, especially, shows me the spirit of perseverance.

I'm also very grateful to my parents-in-law Shenggang Lei and Chaoying Zhao. Without any hesitation, they offered countless help to take care of my son, Ethan Xinchen Yang, since he was born during my Ph.D. studies. I cannot finish my Ph.D. degree in time without them.

The last special thanks go to my lovely wife, Ying Lei. It is her infinite love that makes my life become colorful and full of joy. We have supported each other for almost 10 years. And I feel so happy that we finished every degree, from B.A. to Ph.D., together, no matter how hard it was.

ESSAYS ON FINANCIAL FRICTIONS, MISALLOCATION AND DEVELOPMENT DYNAMICS

EI YANG

Boston University, Graduate School of Arts and Sciences, 2016

Major Professor: Simon Gilchrist, Professor of Economics

ABSTRACT

This dissertation consists of three chapters on financial friction, misallocation and development dynamics.

The first chapter considers how financial frictions and mobility distortions generate the persistence of post-reform development dynamics. I build a general equilibrium model and calibrate it to China. The mobility distortion is an occupation distortion that restricts a proportion of agents to the low-productive sector. A removal of distortions triggers the transition of the economy. Using a calibrated version of the model, the transition path displays slow convergence and mimics the patterns observed in data. The mobility distortion creates high-ability, but poor, agents before the reform. This provides a channel for financial frictions to have longer effect after the reform. Compared with the literature that uses tax distortions, the economy with mobility distortions generates slower convergence.

The second chapter is a welfare analysis of the well-documented depressed migrant wage in China from a dynamic perspective. The depressed migrant wage per se attracts fewer migrant workers and lowers the migrants' consumption and the aggregate output. However, it encourages urban entrepreneurs to substitute capital for labor, relaxing the effect of financial frictions. The net effect on output and consumption depends on the stage of development. Initially, it benefits the economy by speeding up TFP growth and capital accumulation in the urban sector. In the later stage, owing to low consumption of migrants, policy intervention can increase aggregate consumption and output.

The third chapter investigates why the intergenerational income mobility decreases and the inequality increase for China over the past 30 years. I propose a theoretical overlapping generation model with missing capital markets, increasing the return to human capital and increasing education cost to explain these facts. After the economic reform happens, all levels of wages go up and all families accumulate and update human capital. However, the increasing education cost and credit constraint prevent the children from rural families from accumulating human capital quickly. The urban families accumulate human capital faster than the rural families. These predictions from the model are verified in the census data. Whether this process continues or not depends on the subsidy of education. Government education policy can improve the allocation of education in the economy.

Contents

List of Tables

List of Figures

Chapter 1

The Persistence of Development Dynamics: Financial Frictions and Mobility Distortions

1.1 Introduction

When a developing country takes on successful economic reforms, it starts to grow for decades. For example, the Asian miracle countries and China go through persistent growth after their economic reforms. There is little controversy on the fact that economic transition is triggered by economic reforms. However, the long-lasting transitional process itself is puzzling. What are the factors that lead to the persistence of the development dynamics?

The literature approaches this question using the combination of financial frictions and tax distortions. Both financial frictions and tax distortions depress the output and the total factor productivity (TFP) in an economy. The tax distortions depend on the agents' abilities. Agents with higher ability will have a higher probability of being taxed and will have higher tax rates. The tax distortions create initial misallocation in the economy, and the financial frictions slow down the resource reallocation after the economic reform removes the tax distortions. The models in this line of research generate faster converging speed than the data. In other words, the persistence of transitional dynamics is still low. Therefore, we need to analyze other elements that contribute to the persistence of development dynamics.

In this paper, I consider a mobility distortion in a two-sector economy. The mobility distortion restricts a part of the population to the rural sector. These agents cannot freely move to the urban sector and choose the better urban occupations. The mobility distortion is different from the tax distortion used in previous literature, because it does not depend on agents' abilities. The group of agents who suffer from mobility distortions has the same ability distribution as the unrestricted group. Examples for such mobility distortion include the caste system in India, the north-south division of African-Americans in US history, and the rural-urban division under the Hukou policy in China. In particular, I use the Hukou policy of China as the example throughout this paper.

The mobility distortion increases the persistence of transitional dynamics through its impact on the joint distribution of assets and abilities before the reform. Because a significant amount of agents are restricted to the low-productivity sector and cannot freely choose occupations, the mobility distortion directly creates a large distortion on occupation choice. Among those restricted agents, the high-ability ones are influenced the most. Given the low earnings in the rural sector, these restricted high-ability agents save less and are poor. When the economy takes on reforms to remove the distortion, it takes a long time for these high-ability, but poor, agents to build up wealth, engage in entrepreneurship, and reach an efficient scale of production.

There are two major contributions in this paper. First, it is the first paper that analyzes the role of mobility distortion in generating persistence of development dynamics in a general equilibrium model. It shows that mobility distortions, combined with financial frictions, can generate additional persistence in the transitional dynamics. The convergence speed is slower than the one in an economy without mobility distortion. Second, this paper analyzes rural-urban migration, which is missing in the previous literature that discusses the persistence of development dynamics. In this paper, the rural-urban migration is more than a supply of urban labor force. With the assumption of heterogeneity in ability and the occupation choice structure, the process of the rural-urban migration generates an inflow of both high-ability workers and potential entrepreneurs. Thus, the urbanization process contributes to the persistence of the development dynamics after economic reforms.

The model is a continuous-time heterogeneous agents model and I calibrate it to China. Specifically, agents are heterogeneous in their asset holding and abilities. Each agent's ability evolves according to a diffusion process and determines the optimal occupation for the agent. There are three occupations in the economy: farmers in the rural sector, and workers and entrepreneurs in the urban sector. The earnings of workers are proportional to their abilities, whereas the earnings of farmers are constant and independent of their abilities. The entrepreneurs in the urban sector hire workers and rent capital to produce. The individual production technology is in the form of the span-of-control model, so the entrepreneurial profit is also proportional to the ability. Financial frictions are imposed on the entrepreneurship as a collateral constraint in capital renting.

The economy starts from a stationary equilibrium under distortions. In the benchmark, the initial distortions include the mobility distortion and a lump-sum tax distortion. Mobility distortions are the main distortion in the benchmark and the lump-sum tax distortions are used to capture all the other distortions for the entrepreneurs. Due to the mobility distortion, a proportion of agents is forced to work in the rural sector as farmers. The policy reform happens unexpectedly and it removes all the initial distortions once and for all. As a result, the economy begins to grow and it evolves to its terminal state. The endogenous occupation choices and the reallocation of factors generate the endogenous TFP process and other development dynamics. Along the transitional path, the output per capita, capital-output ratio, and urbanization level grow over time.

To show the effects of mobility distortion, I construct two alternative economies. The two economies are different from the benchmark only in the initial distortion. The first one has only mobility distortions and the second one has only entrepreneurial revenue tax distortions. Both economies start at their respective stationary equilibria respectively and the transitional dynamics are triggered by the removal of the initial distortions. The transitional paths are compared with the ones from the benchmark. To make the comparison reasonable, the initial degrees of distortions are chosen such that the rural employment shares are about the same across three economies. The results show that the mobility distortion creates more persistence than the revenue tax distortion. For example, if we use the time that the TFP covers half of the distance to the terminal level as a measure of convergence speed, it is 3 years in the economy with only revenue tax distortion, but 9 years in the benchmark, and 10 years in the economy with only mobility distortion. The urbanization speeds are also slower in the latter two economies than in the economy with only revenue tax distortion.

This paper uses the continuous-time modeling techniques and solves the model numerically with a finite difference method. The foundation of the model is a continuous-time version of the incomplete market models as in Aiyagari [1994], Bewley [1986] and Huggett [1993]. In this type of model, individual agents choose optimal actions based on their idiosyncratic states, and their expectations are rational. The aggregate state is the joint distribution of the individual state variables. When the economy is on the transitional path, we need to track the evolution of the joint distribution. The continuous-time method has an advantage of describing the evolution of the joint distribution using the Kolmogorov Forward (KF) equation (or Fokker-Plank equation). The KF equation is also easy to solve numerically. A detailed reference is Achdou et al. [2015].

1.2 Related Literature

My paper is part of the large theoretical and empirical literature that study the role of financial friction on economic development. Early contributions are Banerjee and Newman [1993]; Galor and Zeira [1993]; King and Levine [1993]; and Rajan and Zingales [1998]. See Banerjee and Duflo [2005] and Levine [2005] for recent surveys.

The recent related literature is the strand that focuses on the macroeconomic implication of micro misallocation. Restuccia and Rogerson [2008] use an implicit tax method to argue that resource misallocation shows up as a low level of TFP for developing countries. Hsieh and Klenow [2009] empirically show that China and India can gain large TFP improvement if the distortion in the economy is removed. Different from their implicit tax approach, my paper identifies one particular type of distortions: the mobility distortion. Also, this paper studies the transitional dynamics and it illustrates that the types of initial distortions matter for the transitional dynamics, and the steady state analysis is only one side of the whole story.

One closely related paper is Buera and Shin [2013] who document the stylized growth facts of successful Asian economies and quantitatively analyze the role of financial frictions and resource misallocation. They compare their model with the neoclassical growth model and emphasize that their model can generate a slower convergence and an endogenous hump-shaped TFP path. In their model, the initial distortions responsible for resource misallocations are modeled as taxes and subsidies on entrepreneurship. This paper builds on their insights, but it introduce mobility distortions and focuses on the rural-urban migration. Although not completely comparable due to the modeling method, my model generates longer transitional dynamics and a monotonically increasing TFP over time. Midrigan and Xu [2014] use producer-level data and emphasize the role of financial frictions on entry and entrepreneurial technology adoption. In contrast, I focus on the extensive distortion generated by the mobility distortion, and the mechanism of persistent transitional dynamics in my paper relies on this distortion.

My paper is also a part of the literature on rural-urban migration. Lewis [1954] builds a two-sector model with unlimited rural labor supply. Lewis's idea of reallocation of the labor force between sectors remains critical in my model. The difference is that the rural immigrants are now heterogeneous and limited in supply. Therefore, this paper can be thought of as describing the economy turning over the Lewis point. Todaro [1969] and Harris and Todaro [1970] address the force for migration by equating the expected wage from unemployment to the rural wage. Unemployment is abstracted away in my paper. The pulling force of the rural-urban migration comes from the increasing urban wage, which is endogenously determined by the reallocation in the economy.

The third strand of closely related literature is the one on China's economy. Recently, there is increasing interest in understanding the behavior of China's economy. Song et al. [2011] focus on the coexistence of large trade surplus and high return to capital. They build an overlapping generation model and emphasize the reallocation between the state-owned firms and privately-owned firms within the manufacturing sector. Chang et al. [2015] document the trend and cycle patterns for China and propose that the preferential credit policy for the heavy industries accounts for these patterns. Both papers focus on the reallocation in the manufacturing sector. In contrast, my paper focuses on the mobility distortion motivated by the Hukou policy and studies the rural-urban migration, a reallocation in the labor market, from the perspective of occupation choice. I view my paper as complementary to theirs in improving a broad understanding of the transitional dynamics of China.

There is other literature on the Hukou policy. For example, Dollar and Jones [2013] apply a search-and-match model and treat Hukou policy as an exogenous restriction on numbers of immigrant workers searching for a job in the city. The analysis is built on the steady state of the economy and there are no entrepreneurs. Compared with their paper, I focus on the effects of the Hukou policy on the transitional path. More importantly, my paper models the Hukou policy in an abstract way.

The idea of the mobility distortion is beyond the Hukou policy in China. It is an occupation distortion based on group characteristics. For example, Hayashi and Prescott [2008] document the prewar patriarchy in Japan and study how it induced a sectoral distortion and a depressed output level in a standard neoclassical two-sector growth model. The patriarchy forces the son designated as heir to stay in agriculture and this can also be thought of as a mobility distortion.

In an abstract sense, the idea that mobility distortion creates a persistent transitional dynamics, can be applied to other similar occupation distortions. A recent paper by Hsieh et al. [2013] documents an increasing share of black men, black women and white women in the high-skilled occupation distribution between 1960 and 2008 in the US. With an augmented static Roy model, they infer that the barriers to occupation are decreasing over time and account for 15 to 20 percent of the growth in aggregate output per worker during that period. The mechanism of mobility distortions analyzed in my paper can be used to understand the dynamic effects of these occupation distortions.

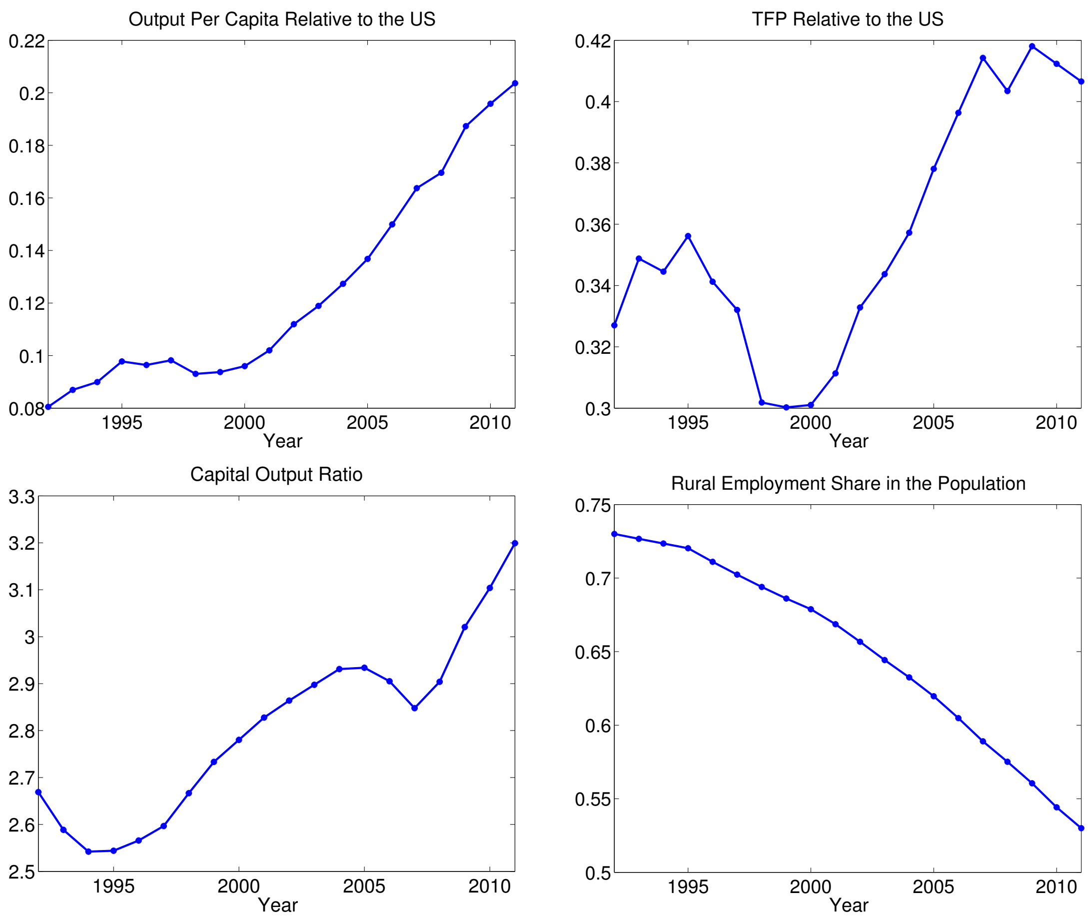
Various papers have modeled the transitional dynamics and investigate the main factors and mechanism. King and Levine [1993] find that it is hard to use the neoclassical growth model to generate slow transitional dynamics. Imrohoroglu et al. [2006] improve the neoclassical model for Japan by feeding in an exogenous calculated TFP. Buera and Shin [2013] point out the importance of interaction between financial frictions and initial distortions to generate endogenous TFP dynamics and the slow convergence. Moll [2014] analyzes theoretically the role of the persistence of idiosyncratic productivity shock in determining the speed of convergence and steady-state productivity losses in a continuous-time model. My paper is complementary to these papers and illustrates the importance of mobility distortions and rural-urban migration.

1.3 Motivation Evidence from China

There are two major features of China: a long-lasting growth after reform and the Hukou policy.

Over the last 20 years, China has gone through a persistent economic transition after economic reforms, and has achieved huge success in economic outcomes. The output per capita relative to the US increases from 8.05% in 1992, to 20.36% in 2011, and the TFP level relative to the US increases from 32.70% to 40.66% (see Figure 1.1). The economy's capital-output ratio increases from 2.67 to 3.20 during the same period. The slow speed of the gradual urbanization is another salient feature. The rural employment share of the population keeps decreasing from 73% in 1992, to 53% in 2011.

Figure 1.1: Economic Variables for China from 1992 to 2011



Note: The starting year in the graph is 1992 and the end year is 2011. All data except rural employment share come from the Penn World Table 8.0 (PWT). See Feenstra et al. [2015] for the instructions on the PWT. The rural employment share is taken from the National Bureau of Statistics of China. See Appendix A for details on the data construction.

Behind these dynamics are the economic reforms in 1992, and changes in the implementation of the Hukou policy in China. In 1992, China started a complete market-oriented reform, and has kept itself on this track to the present. During the same period, the government has relaxed the implementation of the Hukou policy. The Hukou policy initially was created in the 1950s to strictly restrict rural-urban migration through a registration system. The population is divided into two groups. One with the rural Hukou and the other with the urban Hukou. During the most strictly-implemented period, agents with the rural Hukou could not stay in the city legally. However, after the economic reform took place in 1992, agents with the rural Hukou can go to the urban sector and find jobs there. After 20 years of the reforms, there are 274 million of immigrant workers in 2014, estimated by the National Bureau of Statistics of China. Compared with the labor force of 915 million in 2014, the immigrant workers account for about 23.39%.

The Hukou policy in China is a perfect example of mobility distortion. The original Hukou policy restricts a proportion of agents in the economy to stay in the rural area, and the division of agents is based on their rural-urban registration status. The direct impact of the Hukou policy is on the rural-urban migration and the supply of urban labor force. Another impact, usually ignored but critical, is on the allocation of talents: some smart agents with rural Hukou cannot utilize their talents when the Hukou policy is strictly implemented. Both impacts are important for the development dynamics after the economic reform.

1.4 Model

The core of the model is a continuous-time version of Aiyagari-Bewley-Huggett (ABH) incomplete market model. The model has no aggregate uncertainty as in ABH model. Different from the ABH model, this model embeds endogenous occupation choice, financial frictions, and initial distortions.

I first lay out an economy setting with occupation choice and financial frictions, and then describe the individual decisions, aggregation of the economy and the associated competitive equilibrium. The distortions are introduced into the economy at the end.

1.4.1 Environment

There is a continuum of agents of measure one. Each agent maximizes an individual discounted expected utility

There is a continuum of agents of measure one. Each agent maximizes an individual discounted expected utility

$$\max E_0 \int_0^\infty e^{-\rho t} u(c) dt$$

where $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ is the constant relative risk aversion (CRRA) utility function. The parameter ρ stands for the discounted rate.

Every agent is endowed with an idiosyncratic ability z. It follows a diffusion process modeled as

$$d \log z = \psi(\mu - \log z)dt + \sigma dW$$

where W is the standard Brownian motion. The parameter ψ determines the persistence of the process and μ controls the location of the long-run distribution. This process is a continuous-time version of a discrete AR(1) process in $\log z$. The long-run distribution generated by this process is a lognormal distribution $\log z \sim N(\mu, \frac{\sigma^2}{2\psi})$. By Ito's lemma, the process of z is derived as $dz = \left[\psi\left(\mu - \log z\right) + \frac{\sigma^2}{2}\right] z dt + \sigma z dW$. I denote this as $dz = \mu(z)dt + \sigma(z)dW$.

There are three occupations: rural farmer, urban worker and urban entrepreneur. The earnings of a farmer are u, independent of the individual ability z. The ability z is important because it enters the earnings of the last two occupations.

The earnings of an urban worker are wz^θ, where w is the urban wage rate. The ability z enters work's earning with a monotone transformation z^θ, which can be interpreted as an effective unit of labor for a worker with ability z. The parameter $\theta \geq 0$ also controls the curvature of the wage profile across the workers. Different from earnings of a farmer, the worker's earnings are proportional to the ability.

An entrepreneur's earnings are the net profits out of production. As an entrepreneur, the agent can run a firm with a production function

$$f(z,k,l) = z\left(k^{\alpha} l^{1-\alpha}\right)^{1-\nu}$$

by renting capital k and hiring labor l from factor markets. The entrepreneurial production function is specified as a span-of-control model as in Lucas [1978]. The parameter $1-\nu$ is called the span of control. Like the earnings of workers, the individual ability z is important, because it directly controls how productive the firm is and the profits of the firm.

The financial friction shows up as a collateral constraint in the entrepreneurial production. When the asset holding is low, the entrepreneur faces a collateral constraint for the capital used in production, $k \leq \lambda a$. The parameter $\lambda \geq 1$ represents the degree of financial frictions. When $\lambda = 1$, borrowing is shut down and no entrepreneur can borrow in order to produce. The capital amount used in production equals the asset holding. When $\lambda = \infty$, the economy is free from financial friction and entrepreneurs can borrow as much as they want. As a result, when the financial friction presents, it can affect the entrepreneur's production scale.

Given the urban wage rate w and the renting rate R, the net profit from being an entrepreneur is defined as

$$\pi(a,z;w,R) = \max_{k \leq \lambda a, l} f(z,k,l) - Rk - wl.$$

These are the instant earnings for being an entrepreneur.

The asset market is incomplete as the ABH model. There is only risk-free bond trading in the economy. It pays with interest rate r. All agents can save with bonds, but agents cannot borrow money to smooth their consumption. The borrowing constraint for a risk-free bond is set as $\underline{a} = 0$.

The asset market is competitive. The financial intermediaries receive deposits from savers and lend capital to entrepreneurs. The zero-profit condition implies that the rental rate of capital is the sum of the risk-free rate and the capital depreciation rate: $R(t) = r(t) + \delta$.

The urban labor market is also competitive. No entrepreneurs have the power to determine the urban wage. The urban wage is determined by the supply and demand of the workers' effective units of labor.

1.4.2 Individual Decisions

At the beginning of each period, all agents observe their ability z, and then choose their occupations. The production takes place and the economy provides earnings for all agents. At the end of each period, agents choose how much to consume and save.

The production is assumed to be intra-period and there is no cost for switching occupation. For entrepreneurial activity, there is no capital or labor adjustment cost. These assumptions simplify the occupation problem. The occupation is not a state variable but a choice variable in each period.

Agents will choose the optimal occupation by comparing the earnings from different occupations

$$M(a, z; w, R) = \max\left\{u, wz^{\theta}, \pi(a, z; w, R)\right\}.$$

Conditional on being entrepreneurs, agents will choose the optimal capital renting and labor hiring. These factor demand functions depend on the individual state (a, z), because the financial friction restricts the production scale of poor entrepreneurs. Given rental rate R and wage rate w, the optimal capital holding can be computed from the first-order condition as

$$\begin{aligned} k(a, z; w, R) &= \min\left\{\lambda a, (z)^{\frac{1}{\nu}}\left(\frac{(1-\nu)\alpha}{R}\right)^{\frac{1-(1-\nu)(1-\alpha)}{\nu}}\left(\frac{(1-\nu)(1-\alpha)}{w}\right)^{\frac{(1-\nu)(1-\alpha)}{\nu}}\right\} \\ &= \min\left\{\lambda a, (z(1-\nu))^{\frac{1}{\nu}}\left(\frac{\alpha}{R}\right)^{\frac{1-(1-\nu)(1-\alpha)}{\nu}}\left(\frac{1-\alpha}{w}\right)^{\frac{(1-\nu)(1-\alpha)}{\nu}}\right\}. \end{aligned}$$

The first term shows the effect of financial friction: Firms may be operated at an inefficient and small scale when the individual asset holding a is small. The second term is the capital holding for unconstrained entrepreneurs. Given the capital, the optimal labor hiring is computed from the first-order condition as

$$l(a, z; w, R) = \left(\frac{z(1-\nu)(1-\alpha)}{w}\right)^{\frac{1}{1-(1-\nu)(1-\alpha)}} k(a, z; w, R)^{\frac{(1-\nu)\alpha}{1-(1-\nu)(1-\alpha)}}.$$

Because the production function is complementary in capital and labor, the labor demand is proportional to the capital using in the firm.

Given the optimal occupation choice, the budget constraint is

$$\begin{aligned} da &= [M(a,z;w,R) + ra - c]\,dt \\ a &\geq 0 \end{aligned}$$

The optimal occupation choice $M(a,z;w,R)$ depends on the individual state (a,z) and prices (w,R). Given strikingly high rural earnings u, it is possible that all agents will choose to be farmers in an equilibrium. However, this equilibrium is not interesting. In a reasonable equilibrium, the occupation choice in the model should reflect the economic development from migration and entrepreneurship activity. Migration from rural sector to urban sector should increase the output in the economy because occupations in the urban sector represent high productivity occupations. In such a situation, a high-ability agent without financial friction will prefer entrepreneurship the most, and will prefer to be a worker second.

Suppose the economy has a set of reasonable parameters such that the urban sector represents high productivity. This will be the economy used throughout the paper. Agents with relatively low ability, consider being workers or farmers. There will be a worker cutoff $\underline{z}$, such that for all $z \leq \underline{z}$, the earnings from being a farmer are greater than the earnings from being a worker $u > wz^{\theta}$. The optimal occupation in this case is a rural farmer. Agents with relatively high ability, consider being workers or entrepreneurs. The entrepreneurial cutoff is determined by equation $wz^{\theta} = \pi(a,z;w,R)$. This cutoff is a function of (a,z). Any agent with (a,z) satisfying $wz^{\theta} \leq \pi(a,z,A;w,R)$ will choose to be an entrepreneur.

The individual agent's problem can be written in a recursive form using the Hamilton-Jacobin-Bellman (HJB) equation:

$$\begin{aligned} \rho v(a,z,t) &= \max_{c} u(c) + [M(a,z;w(t),R(t)) + r(t)a - c]\,\partial_a v(a,z,t) \\ &\quad + \mu(z)\partial_z v(a,z,t) + \frac{1}{2}\sigma^2(z)\partial_{zz} v(a,z,t) + \partial_t v(a,z,t). \end{aligned} \tag{1.1}$$

The HJB equation is a second-order partial differential equation. The value function $v(a,z,t)$ depends on t because the prices $r(t)$ and $w(t)$ may be changing along the equilibrium path. The first-order condition for the interior region $a > 0$ is $u'(c(a,z,t)) = \partial_a v(a,z,t)$. This equation links the value function with the optimal consumption function. For the constrained agent at $a = \underline{a}$, we have $u'(c(\underline{a},z,t)) > \partial_a v(\underline{a},z,t)$. Intuitively, the additional consumption provides more utility to the agent at the constraint.

1.4.3 Aggregation

The aggregate state in this economy is the joint distribution of abilities and assets $G(a,z,t)$. It is time-varying if the economy is on a transitional path. I denote the density function as $g(a,z,t)$.

The aggregate state is a distribution because prices depend on the aggregate variable generated from the distribution. The agents are heterogeneous in asset and ability, and the aggregation of their actions produces the real quantities in the economy. More importantly, agents are forward-looking. To make current consumption and saving decisions, agents need to predict the future prices $w(t)$ and $r(t)$. This requires them to keep track of the joint distribution $g(a,z,t)$ to infer the occupation choices of others, and the supply and demand in the factor markets.

Taking advantage of the continuous-time method, the evolution of the density function of the joint distribution is characterized by the Kolmogorov Forward (KF) equation:[1]

$$\partial_t g(a,z,t) = -\partial_a \left[s(a,z,t)g(a,z,t)\right] - \partial_z \left[\mu(z)g(a,z,t)\right] + \frac{1}{2}\partial_{zz}\left[\sigma^2(z)g(a,z,t)\right] \tag{1.2}$$

where $s(a,z,t) = M((a,z;w(t),R(t)) + r(t)a - c(a,z,t)$ is the saving function. Intuitively, the change in the aggregate distribution comes from two parts. The first part comes from change in asset holding. It naturally relates to individual's optimal saving function. This is summarized by the first term on the right-hand side of the KF equation. The other two terms describe the change from the evolution of ability. The last requirement for the density function is that the integral of $g(a,z,t)$ sums to one: $\int\int g(a,z,t)dadz = 1$ for all t.

[1]Please refer to Stokey [2008] for the rigorous derivation of the KF equation in general form.

Using the density function, all the aggregate quantities can be calculated by integration over the state space. Given the occupation choice of each agent, the urban labor supply is an integration of all workers' effective labor units. The urban labor demand is an integration of all entrepreneurs' labor demand. Similarly, given the saving decision of each agent, the total saving is the integration of all agents' saving. The total capital is the integration of each entrepreneurs' capital demand.

1.4.4 Equilibrium

The equilibrium in this economy is a competitive equilibrium. The equilibrium is the time paths for prices $r(t)$, $w(t)$, $t \geq 0$, and corresponding quantities such that given the initial distribution $g(a, z, 0)$:

(1) given the time paths of prices $r(t)$ and $w(t)$, each agent chooses an occupation based on individual state (a, z), and chooses how much to consume $c(a, z, t)$ and save $s(a, z, t)$ to maximize discounted utility;

(2) the urban labor market is clear at each period: the supply of urban labor equals the labor demand from urban entrepreneurs;

(3) the asset market is clear at each period: the supply of saving from all agents equals the capital demand from urban entrepreneurs.

1.4.5 The Initial Distortion in the Benchmark Model

In this subsection, I introduce the distortions in the benchmark model. The initial distortions are used to characterize the economy before policy reform. The economy is initially at a stationary equilibrium with these distortions. The transitional dynamics are triggered by the removal of these distortions.

The first distortion is the mobility distortion. It is assumed that a proportion of q_1 agents in the economy are restricted to the rural area. They cannot choose the optimal occupation according to their ability. Instead, they can only choose to be farmers. With mobility distortion, there is a large labor misallocation on both the urban labor supply and the quantity and quality of active entrepreneurs.

The second distortions are lump-sum taxes on the active entrepreneurs. They are used to capture the other initial distortions imposed on the urban sector. The lump-sum tax is positively correlated with the entrepreneurial ability. With the lump-sum tax, the profit function for an entrepreneur is

$$\pi(a, z, A; w, r, \kappa(z)) = \max_{k \leq \lambda a, l} f(z, k, l; w, r) - Rk - wl - \kappa(z)$$

where $\kappa(z)$ is the lump-sum tax, specified as $\kappa(z) = z^{q_2}$ and $q_2 > 0$. The lump-sum tax directly distorts the entry decision of urban labor force. It also distorts capital and labor allocation among entrepreneurs indirectly through general equilibrium prices.

1.4.6 Two Alternative Economies for Comparison

To illustrate the role of mobility distortions in the benchmark model, I create two alternative economies. The transitional dynamics from those two economies will be compared with the one from the benchmark model. These two economies are different from the benchmark model only in the initial distortions. After the removal of initial distortions, they evolve into the same terminal state of the benchmark model.

The first alternative economy is one with only mobility distortions. The initial distortion only consists mobility distortion. Because mobility distortion creates more persistence than other distortions, the transitional path from this economy provides a natural boundary.

The second alternative economy is one with only revenue tax distortions. The initial distortion here only consists of revenue taxes. The revenue taxes $\tau(z)$ are imposed on active entrepreneurs. The profit function for the entrepreneur is changed into

$$\pi(a, z, A; w, r, \kappa(z)) = \max_{k \leq \lambda a, l} (1 - \tau(z)) f(z, k, l; w, r) - Rk - wl$$

where $\tau(z) = 1 - \exp(-q_3 z)$ with $q_3 \geq 0$. As a result, agents with higher ability will suffer from a higher revenue tax if they choose to be entrepreneurs. The revenue tax tries to capture the distortions on the urban sector.

As long as the urban tax distortions do not put some high-ability agents into the occupation of farmers, they are different from mobility distortion, and the forms of lump-sum tax or revenue tax are not important from the point of transitional persistence.

1.4.7 An Economy without Financial Friction

To understand the economy better and provide a definition of the TFP, this subsection investigates an economy without financial friction.

Under no financial friction, the model is much simpler. First, the occupation choice for being an entrepreneur is independent of the individual's asset holding. Suppose the economy is at a meaningful equilibrium. The agents with the highest ability will choose to be entrepreneurs and the agents with the lowest ability will choose to be farmers. In this equilibrium, the cutoff to be entrepreneur will be determined by comparison between wage income wz^{θ} and the unconstrained profit $\nu z^{\frac{1}{\nu}}\left[(1-\nu)\left(\frac{\alpha}{r+\delta}\right)^{\alpha}\left(\frac{1-\alpha}{w}\right)^{1-\alpha}\right]^{\frac{1-\nu}{\nu}}$. Thus, every firm will be operated in its efficient scale. Let the cutoff of being an entrepreneur to be denoted as $\bar{z}$. Because occupation choice no longer depends on the asset holding, the wealth distribution affects the dynamics only through the aggregate quantity of the capital it generates.

Second, in this equilibrium, the aggregate urban output has a simple expression of a decreasing-return-to-scale production function

$$Y_t^{ur} = \left(\int_{\bar{z}_t} z_i^{\frac{1}{\nu}} dG(z)\right)^{\nu} L_t^{(1-\nu)(1-\alpha)} K_t^{\alpha(1-\nu)}.$$

The aggregate urban production function shares the same power coefficients in factor inputs as the production function at the individual level, but the quantities of inputs are aggregate effective labor and capital in the urban sector. Another key feature is the technology factor. In the aggregate urban production function, the technology level is linked to an expression of the individual abilities from the active entrepreneurs. This technology factor increases when more entrepreneurs enter or when active entrepreneurs have the higher individual ability.

The aggregate urban production function motivates the urban TFP definition as

$$TFP_t = \frac{Y_t^{ur}}{L_t^{(1-\alpha)(1-\nu)} K_t^{\alpha(1-\nu)}}.$$

Throughout this paper, I focus on the urban TFP. I treat it as the data counterpart because there is no aggregate production function for the model economy.

1.5 Quantitative Analysis

Three economies are calibrated in this section: the benchmark economy, the economy with only mobility distortions, and the economy with only revenue tax distortions. Each of them initially is at stationary equilibrium. They are different in the initial conditions created by different distortions. The benchmark economy has two types of distortions while the other two counterfactual economies have only one type of distortion. The distortions are removed at the beginning of reform. Even though the initial conditions are different across economies, the ending stationary equilibria are same.

To calibrate the parameters for these economies, I first set some standard parameters as values used in the literature, and then calibrate parameters that determine the common terminal state. By design, these parameters are invariant for all economies and across time. In the end, I calibrate distortion parameters for the initial state. For the benchmark model, the parameters are chosen such that the initial state of the economy is close to the status of China in 1992. The distortion parameters for the other two economies are selected to make them comparable with the benchmark model.

1.5.1 Calibration

Parameters Invariant across Time and Economies

The time length is one year. The relative risk aversion coefficient γ is set to equal 1.5 as standard. The depreciation rate δ and the production parameter α are chosen as 10% and 0.5, respectively.

These numbers are based on estimates for China from Bai et al. [2006]. The ability process parameter μ is set to equal 0 as a normalization.

I associate the economy without financial frictions, distortions or the rural sector to the US economy, because the US economy is a financially developed economy and it had only 1.5% employment in the agriculture sector in 2012 (U.S. Bureau of Labor Statistics). This idea provides moments to calibrate the span of control ν, the persistence of $\log z$ process ψ, the standard deviation of shock σ and wage function parameter θ. The parameters are chosen to jointly match the following moments: share of entrepreneurs 7.5% calculated from Survey of Consumer Finances (SCF) (Cagetti and Nardi, 2006), the wealth share of the top 10 percent household 76.7% in 2010 calculated from SCF (Wolff, 2012), the employment share of top 16% establishments (US census 2012), and the top5 earning share 30% in 1998 (Buera and Shin, 2013). When the persistence parameter ψ is larger, the high ability stays longer and the wealth share of the top 10% will be large. For a given ψ, the higher is the risk parameter σ, the lower is the employment ratio of the top 16% of firms. The higher is the revenue share ν, the higher is the proportion of entrepreneurs in the population. The wage profile parameter θ translates into the top5 earning share of the economy. After fixing these parameters, I choose discounted rate ρ to match the risk-free rate of 4%. The summary is in Table 1.1

Table 1.1: Calibrated Invariant Parameters

		Variable	Target	Data	Model
ρ	0.0730	discounted rate	risk-free interest rate	4%	4.00%
θ	0.07	worker's wage function	top5 earning share	30%	29.55%
ψ	$e^{-\psi} = 0.87$	persistence of $\log z$ process	top10 wealth share	76.7%	76.7%
σ	$\frac{\sigma^2}{2\psi} = 0.252$	sd of $\log z$ process	16% employment share	51.6%	51.2%
ν	0.16	revenue share of entrepreneur	Share Entrepreneurs	7.5%	7.49%

Parameters of Terminal State and Initial Distortions

The next two elements to be calibrated are the financial friction and the rural earnings. These two parameters will determine the terminal state of the economy after the reform. After the terminal state is pinned down, I choose parameters for initial distortions across three economies.

The parameter of financial frictions has a long-run effect on the GDP and TFP in the stationary equilibrium. The higher is λ, the lower are GDP and TFP in the stationary equilibrium. Following Buera and Shin [2013], I choose λ such that the external finance-to-GDP ratio is 0.79 (China in 2002) in an economy without any other distortion. The external finance-to-GDP ratio is defined by the sum of two ratios: private credit by deposit money banks and other financial institutions to GDP and private bond market capitalization to GDP.[2] This gives $\lambda = 1.446$, which suggests that entrepreneurs can borrow 40% against their assets. The financial parameter is kept constant over the transition. This is a simplification of reality. In reality, the financial market reform comes later and is implemented more slowly.

The rural income u is set to make the rural employment share equal to 10% in the terminal state under the financial friction. The number 10% is close to the mean of the rural employment share of Japan (7%) and Korea (18%) in 2014 (World Bank Database). This choice of parameter implies that without any distortions other than the financial friction, the economy has 10% rural employment (See Talbe 1.2). Up to now, all parameters relevant to the terminal state are pinned down.

Table 1.2: Calibrated Parameters for Terminal State

		Variable	Target	Data	Model
λ	1.446	financial friction	external finance / GDP	0.79	0.79
u	1.40	rural income	terminal rural employment	10%	9.73%

The initial condition in the benchmark model is modeled by a stationary equilibrium with two additional distortions. The parameter q_1 describes how many agents are restricted to stay in the rural area, and the parameter q_2 describes the degree of output tax correlated with ability. Those two are chosen jointly to match the two moments from the data on China: the initial rural employment rate of 73.0% in 1992, and a 26.1% increase in TFP relative to the US from 1992 to 2011. By choosing $q_1 = 0.424$ and $q_2 = 0.6$, these two moments are matched exactly (see Table 1.3). Given these parameters, the initial equilibrium of the benchmark economy is pinned down.

[2] These data are from the Financial Development and Structure Dataset constructed and maintained by researchers in the World Bank. See Beck et al. [2009] and Cihak et al. [2012] for reference.

Table 1.3: Initial Condition for the Benchmark

		Variable	Target	Data	Model
q_1	0.424	mobility distortion	initial rural employment	73.0%	73.0%
q_2	0.6	urban distortion	increase in TFP after 20 years	26.1%	26.1%

To make the alternative two economies comparable with the benchmark one, I choose the parameters of distortion such that the initial rural employment is as close to the benchmark as possible. In this sense, the three economies have a similar distortion in terms of rural employment share. Table 1.4 gives the numbers of initial rural employment share for three economies.

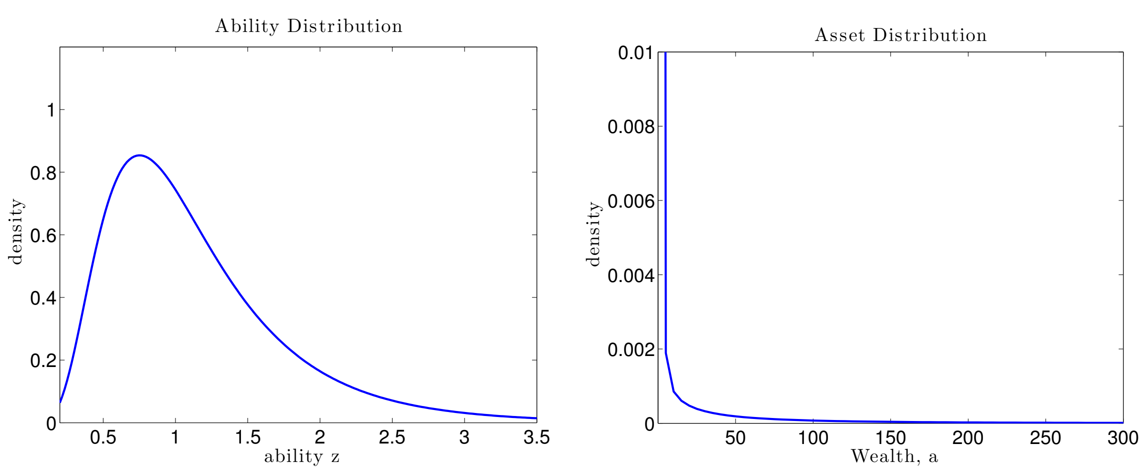
Table 1.4: Calibrated Parameters for Initial Conditions

	Mobility	Lump-sum Tax	Revenue Tax	Initial Rural Employment
Benchmark	$q_1 = 0.424$	$q_2 = 0.6$	$q_3 = 0$	73.0%
Mobility Distortion Only	$q_1 = 0.701$	$q_2 = 0$	$q_3 = 0$	72.69%
Revenue Tax Only	$q_1 = 0$	$q_2 = 0$	$q_3 = 0.068$	72.61%

1.5.2 Initial Stationary Equilibrium of the Benchmark

This subsection shows important features of the stationary equilibrium of the benchmark before the mobility distortion is removed. Both the financial frictions and the initial distortions are critical to understand the agents' behavior and the allocation of the resources in the economy.

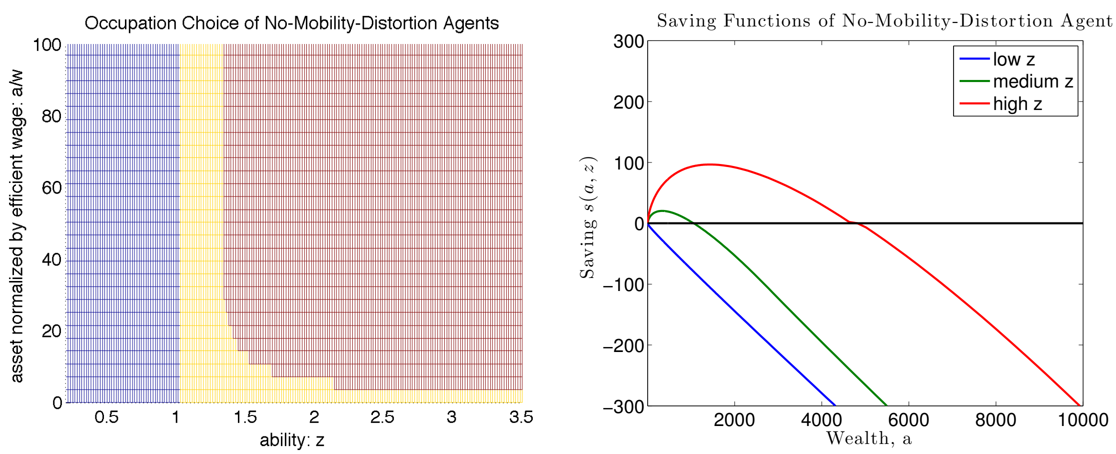
Figure 1.2: Ability Distributions and Asset Distribution



The diffusion process of the ability generates a stationary log-normal ability distribution (see Figure 1.2). Most of the agents are endowed with low abilities. Although the individual's ability is changing over time, the distribution of ability for the whole economy is stationary. The distortions do not change the stationary distribution of ability.

Compared with the stationary ability distribution, the stationary wealth distribution is highly left-skewed. Most agents have very little wealth. The wealth distribution is a consequence of agents' saving decisions. In equilibrium, only the high-ability agents without mobility distortion have strong incentives to save and accumulate large wealth.

Figure 1.3: Policy Functions for the Agents without Mobility Distortion



For the agents not restricted by mobility distortions, their occupation choices and saving functions are affected by financial frictions and the lump-sum tax (see Figure 1.3). In current calibration, if there are no financial frictions and distortions, the high-ability agents will choose to be entrepreneurs. With financial constraint, the decision of whether to be entrepreneur depends on the asset holding. This feature is salient in the occupation-choice figure. The high ability, but poor, agents choose to be workers. There is a more subtle issue. Even if one is an entrepreneur, the scale of the firm will also depend on the asset holding. The firm's scale under financial friction can be smaller than the efficient scale.

The financial friction has important effects on the saving functions. In an economy without financial friction, the saving function will be decreasing in the asset holding for all abilities. For the high-ability agents who earn more, they save to build buffering wealth when poor, and they spend to smooth consumption when rich. Their saving functions start positive and decrease as assets increase. For the low ability agents, they earn less and use assets to smooth the consumption. Their saving starts close to zero and declines as assets increase.

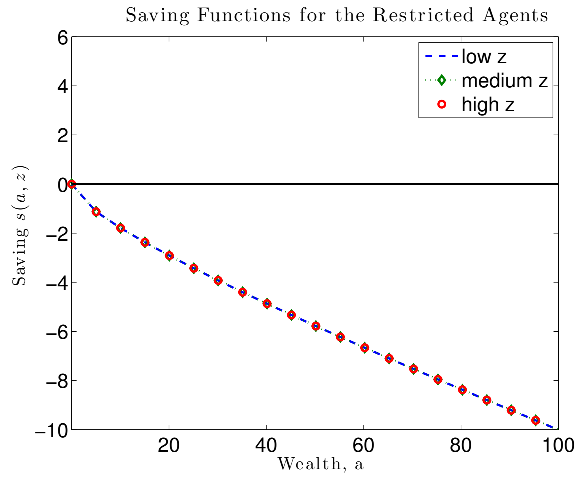
When the financial friction appears, the saving becomes highly nonlinear depending on ability. The high ability, but poor, agents cannot save a lot because their earnings are low. If they are entrepreneurs, their current production scales are limited by the financial constraint. The financial friction can make things even worse and these agents have to become workers. To overcome the collateral constraint and take advantage of high productivity, the high-ability agents will save more as their firms are expanding from small scale. As the firms become larger, the return to saving decreases. At some level of asset holding, the firm will reach its unconstrained scale. The return from saving equals the prevailing return on the risk-free bond. Passing that level of asset holding, the impatience motive dominates the high return and their saving becomes negative. The agents with the low ability have low earnings so that the consumption-smooth motive always dominates. Their saving functions remain to be monotone decreasing in asset holding.

The agents under mobility distortions cannot choose their occupation, and only are able to become as farmers. Their saving functions will be decreasing over asset holding no matter what abilities they have. The saving functions for agents with different abilities are the same (see Figure 1.4).

1.5.3 The Long-Run Effects of Financial Frictions and Distortions

In this subsection, I evaluate the long-run effects of financial frictions and distortions. I calculate the stationary equilibria under different levels of financial frictions or distortions. The experiments show that both financial frictions and distortions can negatively affect the economy, and the effects are highly nonlinear in the degrees of frictions and distortions.

Figure 1.4: Policy Functions for the Agents without Mobility Distortion



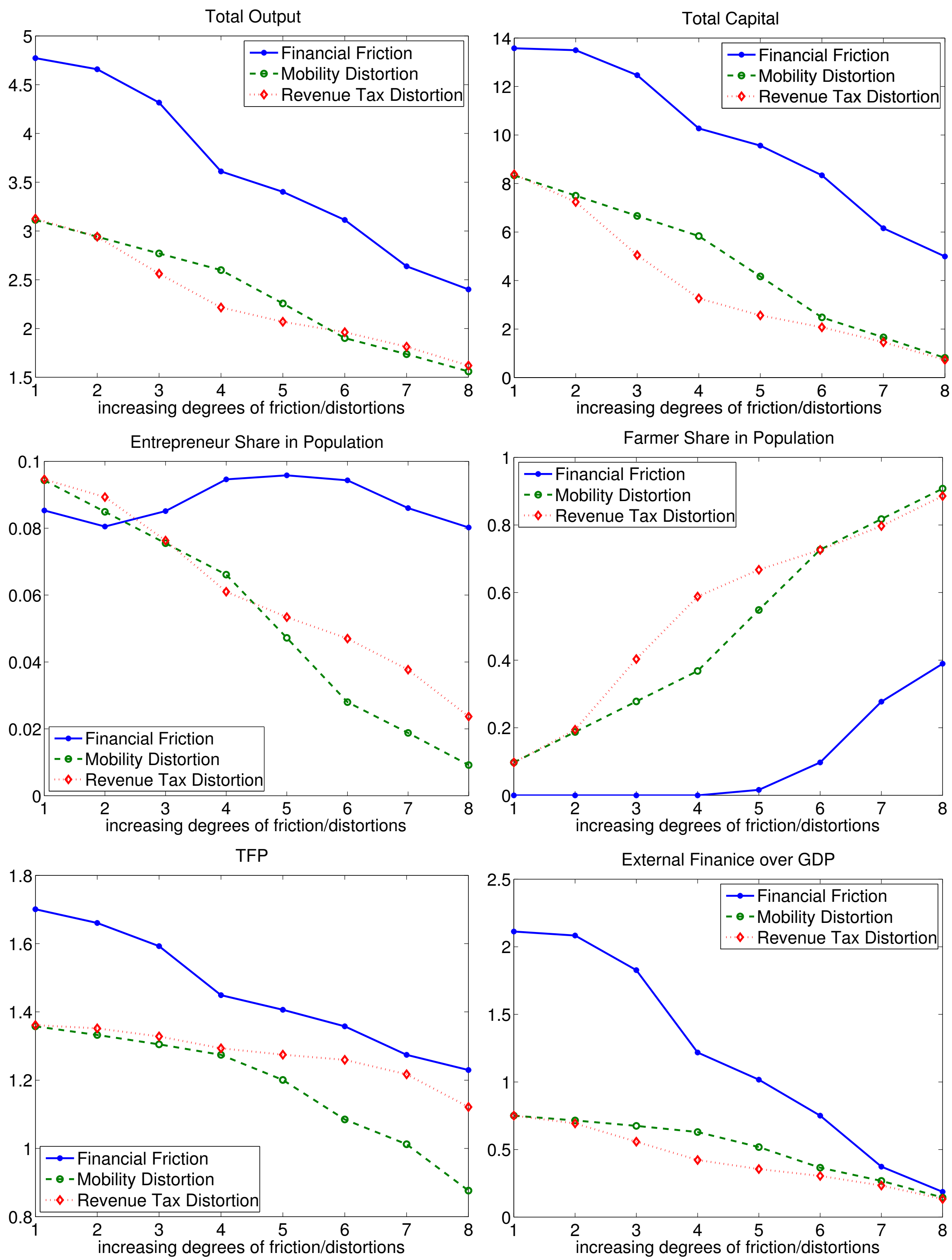
The Long-Run Effects of Financial Frictions

The first dimension under investigation is financial friction. The results confirm the insight from the literature that financial frictions greatly depress economy outcomes (see Figure 1.5).

The experiments start from an economy with no financial friction to a sequence of ones with increasing degrees of financial frictions. All the economies are at thier stationary equilibria and they are free of other distortions. Therefore, all the negative effects are coming from financial frictions. With a stricter collateral constraint, the economy experiences drops in total output, total capital, entrepreneurship, and external finance-to-GDP ratio as in Buera and Shin [2013]. This is the emphasis of the literature and it is confirmed in my model.

Financial frictions lead to a low level of entrepreneurship. The steep drop in the external finance-to-GDP ratio reflects the channel financial friction affecting the economy. Given the collateral constraint, some of the talented, but poor, agents cannot be entrepreneurs, and some are operating on a small scale. The lower level and low quality of entrepreneurship generate the decrease in output, TFP and capital level. In addition, some low ability, but rich, agents enter as entrepreneurs when financial frictions increase. The entry of low-ability entrepreneurs can be larger than the number of exits of high-ability entrepreneurs during some degree of financial frictions. This produces a hump-shaped path for the numbers of entrepreneurs.

Figure 1.5: Long-run Effect of Financial Friction and Distortions



Note: the numerical number of the parameters for financial friction, mobility distortion and revenue tax rate are set as $\lambda = 3 \times 10^5, 10^3, 5, 2, 1.7, 1.446, 1.2, 1.1$, $q_1 = 0, 0.1, 0.2, 03, 0.5, 0.7009, 0.8, 0.9$, and $q_3 = 0, 0.01, 0.03, 0.05, 0.06, 0.068, 0.08, 0.1$.

Different from the literature, my model has an additional rural sector. This factor creates more non-linearity in occupation choice. When the financial fractions are low, no farmers show up in the model. When the financial frictions increase to a certain degree, the low levels of TFP and capital lead to a low urban wage. The low-ability agents begin to prefer being farmers.

The Long-Run Effects of Distortions

The effects of distortions are evaluated under a fixed degree of financial friction $\lambda = 1.446$, which is the calibrated value used in the benchmark. Both mobility distortions and revenue tax distortions can change outcomes of the stationary equilibrium.

The mobility distortions lower the output by restricting more agents in the low-productivity sector. Under no lump-sum taxes $q_2 = 0$, the stationary equilibria are computed under the different levels of mobility distortions. The proportions of agents restricted to the rural sector increase from 0 to 90%. The mobility distortion allocates agents regardless of their ability. This creates a direct impact on the TFP. The huge amount of farmer tracks the degree of mobility distortion closely.

The last distortion under scrutinization is the revenue tax distortion. It is evaluated under the financial friction level of $\lambda = 1.446$, and no mobility distortion $q_1 = 0$. When q_3 increases from 0 to 0.1, the revenue tax rates change from a zero and flat tax rate scheme to an upward sloping nonzero one. The larger is q_3, the higher is the tax rate for all agents, and the larger is the slope of the tax rate. That is, the tax rates increase and they increase more for the agents with higher ability. The revenue tax distortion directly creates large distortions on the entry of entrepreneurs among the high-ability agents. An indirect effect comes from the misallocation of the resource. The larger is q_3, the more the economy is distorted.

Even though both distortions create a drop in GPD, TFP, capital and labor supply, the channels are different. The difference in transitional dynamics from different types of initial distortion is the one of key points of the next subsection.

1.5.4 Transitional Dynamics

In this subsection, I first analyze the transitional dynamics of the benchmark to show why mobility distortion creates persistence in the transitional dynamics. Then I compare the benchmark with two alternative economies in order to show the difference in the persistence under the mobility distortion and revenue tax distortions.

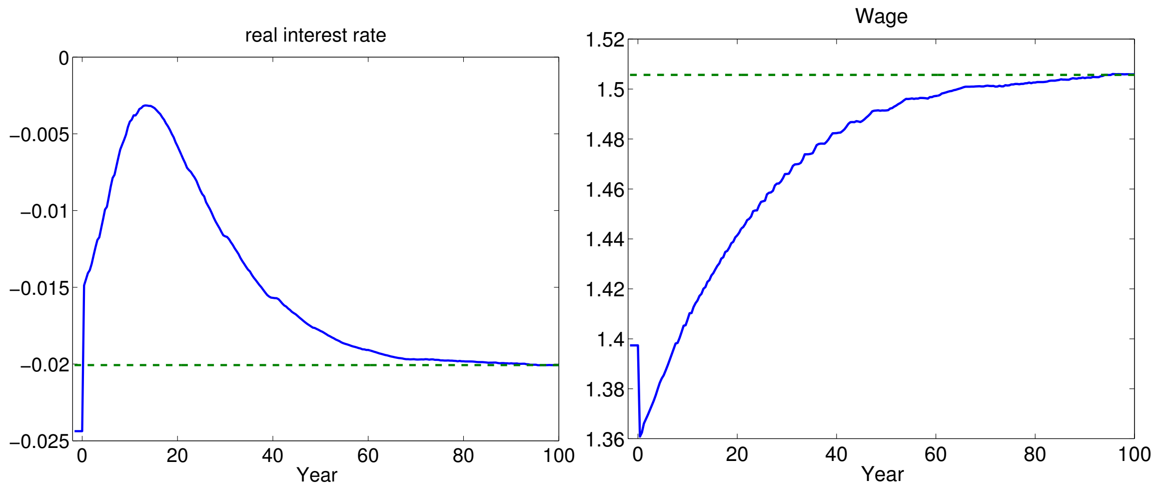
Transitional Dynamics of the Benchmark

The benchmark is at stationary equilibrium but is imposed with mobility distortion and lump-sum tax distortion. Both distortions are removed unexpectedly. The agents realize that it is permanent policy reform, and choose the best occupation according to their abilities and current asset holding. The economy starts its transition to a new stationary equilibrium after this distortions are removed. Along this transitional path, the economy goes through a complicated reallocation of factors between sectors and among entrepreneurs.

To understand the transition better, I analyze the allocation in the economy along two margins of the entrepreneurship: the extensive margin and the intensive margin. The extensive margin describes the entry and exit of entrepreneurs. The intensive margin refers to the capital and labor used by each active entrepreneur.

In the extensive margin, three different groups change their occupations. The first group consists of the high-ability agents who were initially restricted by the mobility distortion. After the mobility distortion is removed, they can utilize their talents. They move from rural to urban area and become workers. This surge in urban labor supply gives rise to the initial drop in urban wage (see Figure 1.6).

Figure 1.6: Price Evolutions in the Benchmark



The high-ability agents want to become entrepreneurs immediately, but they cannot do so. This is because their asset holding pre-reform is very low, and collateral constraints make the entrepreneurial profit low. Although they are not entrepreneurs right after the reform, they expect to become entrepreneurs in the future and begin to accumulate wealth. Because these agents are initially poor, it will take a long time to be productive and reach the efficient scale of the firm, even if their high abilities remain. This is the crucial reason why mobility distortion contributes to the persistence of transition.

Those who are initially relative poor workers but now choose to become entrepreneurs are the second group of agents. Facing the lump-sum tax distortion, the pre-reform profits for their small firms would be low, so they do not enter as entrepreneurs. When the lump-sum tax distortion is lifted, the profit is high enough. Due to the existence of financial friction, the scale and profit of their firms are still below their unconstrained counterparts.

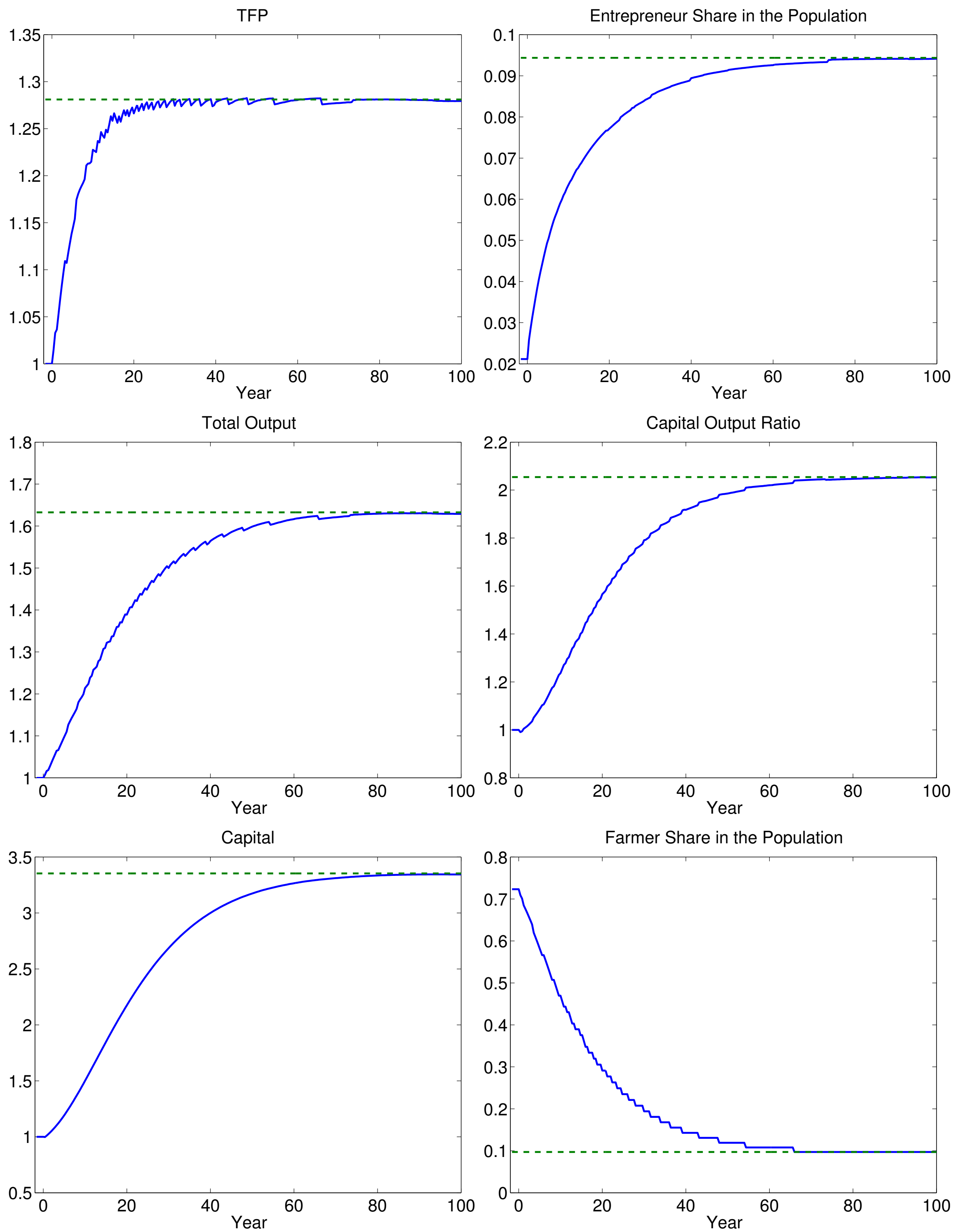
The last group in the extensive margin is a quitting group. Among this group, agents initially are entrepreneurs. Although their abilities are low, they are rich enough to operate a large enough firm such that the entrepreneurial profits are larger than the earnings from being workers. As the economy grows, the interest rate is driven up by capital demand from more productive agents. Additionally, the wage is rising too. The increase in both prices squeeze the profit for these entrepreneurs. At the same time, they could earn more as workers. Thus, along the transitional path, their firm scales will decrease and they gradually quit from engaging in production.

The intensive margin describes the reallocation of factors among the active entrepreneurs. The reallocation of the resource takes place along the transitional path all the time. The low-ability entrepreneurs gradually scale down their firms, while the high-ability entrepreneurs use more and more capital and labor in their production. The resources in the economy gradually shift to the entrepreneurs with high ability.

The factor prices also reflect the above entrepreneurial activities. The entry of the first two groups of new high-ability entrepreneurs drives up the demand for capital. Consequently, the interest rate jumps up at the beginning (see Figure 1.6). As more wealth is shifting to the high-ability agents, the demand for capital is higher and drives up the interest rate. Over time, the reallocation of capital will slow down and the diminishing return to capital begins. Then the interest rate begin to decreases.

The dynamics of the wage rate also reflect and influence the reallocation in the economy. The reallocation drops at the beginning because a significant amount of workers migrate to the urban sector and become workers. The wage rate is increasing over time because the improvements in the TFP dominates the effect from the increase in urban labor supply. This is the continuing pulling force for the rural-urban migration in the model. The increasing wage also pushes out the low-ability entrepreneurs and leads the resources to the high-ability entrepreneurs.

Figure 1.7: Transitional Dynamics in the Benchmark



Note: The dotted line denotes the level of the variable in the new stationary equilibrium. The total output, TFP, capital, and capital output ratio are normalized by their pre-reform level.

Given the economy-wide reallocation along the transitional path, the economy shows slow convergence and the simulation outcomes mimic the patterns in the data (see Figure 1.7). The TFP grows gradually. It takes more than 20 years for the TFP to be close to its terminal level. The other variables grow even more slowly than the TFP, and it takes almost 60 years for them to approach to their terminal levels. This feature is because capital accumulation and the rural-urban migration will continue even after the distortion on entrepreneurship is close to ending. The total output is equivalent to output per capita because there is no population growth in the model.

The rural-urban migration in this model contains high-ability workers and potential entrepreneurs. This feature does not depend on the mobility distortion. As long as there are heterogeneities in ability and occupation choice, this channel exits. The mobility distortions only strengthen this channel at the beginning of the reform. This channel is important because the potential entrepreneurs matter in the future even if they are workers now. In an economy with same-ability agents, the rural-urban migration only supplies workers to the urban sector. There will be no large degree of persistence of development dynamics.

The financial friction takes an important role in generating the persistence. It contributes to the persistence of transitional dynamics through two channels. First, it delays the entry of entrepreneurship. It raises the asset level required for entry as long as the agent is financially constrained, and also potential entrepreneurs need to save for the collateral. Second, the financial friction limits the speed of expanding. It increases the time for the high ability, but poor, entrepreneur to operate in full scale.

Comparisons with Two Alternative Economies

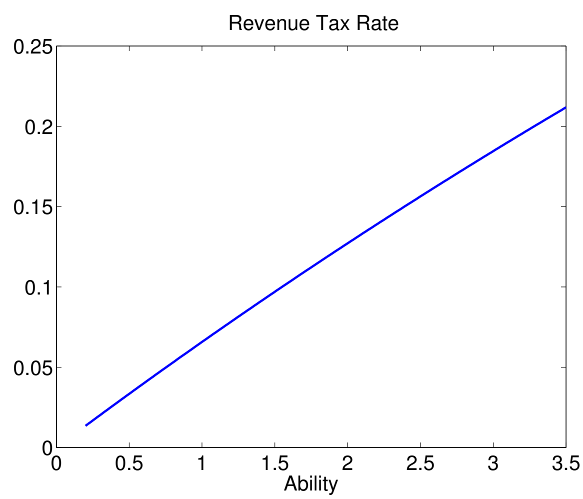
The mobility distortions and the rural-urban migration generate more persistence in the transitional dynamics compared with the revenue tax distortion. This subsection explains the mechanism and shows the quantitative differences through the comparisons between three economies.

Three economies are presented here: the benchmark, the economy with only mobility distortions, and the economy with only revenue tax distortions. All economies face the same level of financial friction throughout the transition. They also share the same terminal state. Therefore, after the distortions are removed in the economies, they all evolve to the same destination.

The benchmark economy has both the mobility distortion and the lump-sum tax distortion. The lump-sum tax distortion is used to represent other distortions in the urban sector. The other two economies only have one distortion. The calibration of parameters are in Table 1.4 in Section 1.5.1.

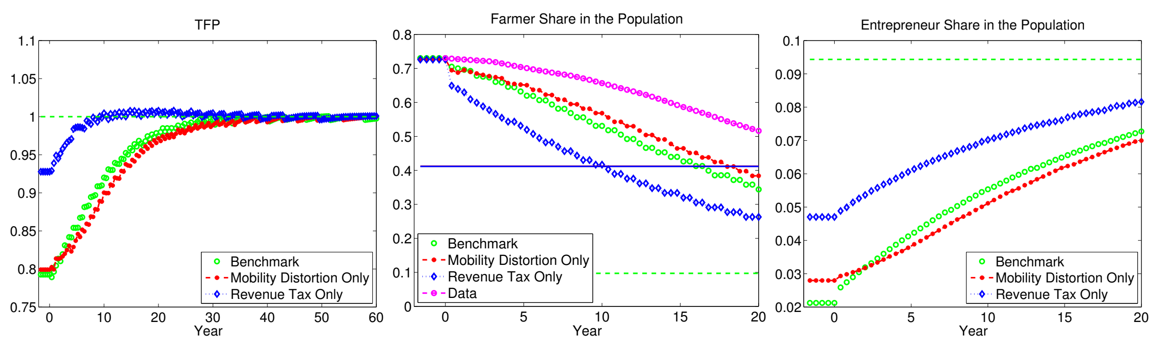
It is not reasonable to directly compare an economy with one initial distortion to an economy with two initial distortions. To make the comparison meaningful, I choose the parameters of distortions such that the initial rural employment share in the economies are the same. The economies are distorted relatively to the same degree in terms of the misallocation of employment in the rural sector. The economy with only mobility distortions initially restricts 70% of the population in the rural sector, while the benchmark restricts 42.4%. The revenue tax rate is increasing from 1.35% to 21.18% in the economy with only revenue tax distortions. The revenue taxes are levied on the active entrepreneurs and they are increasing in the abilities (see Figure 1.8).

Figure 1.8: Revenue Tax Distortions



Compared with the economy that has only revenue tax distortions, both the benchmark and the economy with only mobility distortions have a slower convergence of speed. Figure 9 shows the transitional paths for TFP in three economies. Because they are converging to the same terminal state, the TFP sequences are normalized by the TFP level in the terminal state. The figure on the right is the same as the one on the left, but with a 20-year period. This time period corresponds to the years 1992 to 2011 in China. We can use the time to cover the half distance between the initial level and the terminal level as a measure of convergence speed. The time increases from 3 years in the economy with only revenue tax distortions to 9 years in the benchmark, and 10 years in the economy with only mobility distortions. This is a huge increase in the benchmark model.

Figure 1.9: Comparisons between Economies



Note: The solid lines are the middle points of the initial level and the terminal level.

The mobility distortion is the key to understanding the difference in convergence. First, there are fewer entrepreneurs in the urban sector. Second, some agents with high ability in the rural area cannot use their ability even as workers. Without knowing that the mobility distortion will be removed in the future, these agents consume all their low income in equilibrium and save basically nothing. As a result, they are poor. When the mobility distortion is removed permanently, they move to urban sector to first become workers and then they become entrepreneurs. As initially being very poor, they need to accumulate a large amount of wealth to become entrepreneurs.

The high-ability agents have a different situation in the economy with only revenue tax distortions. Although revenue taxes have impacts on the decisions of being entrepreneurs or not, the high-ability agents can either be urban workers or entrepreneurs with small firms. Their earnings are higher than the farmers and they are not as poor as farmers. When the revenue taxes are removed, they are close to their efficient scales in production.

The lower numbers of entrepreneurs and the lower entry rate of high-ability entrepreneurs are the reasons for the slow converging speed of TFP. The paths of the entrepreneur share in the population are in Figure 10, and they verify these two points.

Another essential difference is the speed in urbanization (see Figure 1.9). The rural employment share in the population decreases more slowly in the benchmark model than in the economy with only revenue tax distortions. Twenty years after the reform there is still nearly 40% rural employment in the benchmark and 27% in the later economy. The number for the employment share in the rural sector is around 53% in the data. Economies with mobility distortions produce a closer simulation to the data.

The benchmark assumes a once-and-for-all removal of mobility distortion. With such an assumption, the impact of mobility distortion works only through the distortion on the initial condition. In reality, the removal of mobility distortion is complicated and nonlinear. It is even different across the provinces within the country. If embedded with gradual removal of the mobility distortion, the simulation outcome can be closer to the data.

In sum, the mobility distortion creates more persistence in the transitional dynamics compared with the economy with only revenue tax distortions. The difference stems from the initial joint distribution of assets and abilities. The mobility distortion restricts some high-ability agents in the rural sector and makes them poor. These high-ability agents need more time to engage in production and produce at efficient scales. As a result, the economy grows more slowly.

The diverse post-reform performances also highlight the need to investigate the types of distortion. We need to understand not only the direct effect of the distortion but also the endogenous behavior that comes with it. In terms of the mobility distortion, it has a more direct impact on the extensive margin of entrepreneurship. More importantly, it also produces endogenous asset-holding behavior, which has a long effect even after the distortion is removed.

1.6 Conclusion

This paper analyzes how financial frictions and mobility distortions can generate the persistent development dynamics after economic reforms in a heterogeneous-agent occupation choice model. The mobility distortion restricts a proportion of agents to the low-productive occupation. Being calibrated to China, the model produces the slow convergence of the economy after the economic reforms remove the initial distortions.

The paper highlights the role of mobility distortion in generating persistence of development dynamics. The mobility distortion contributes to the slow convergence by creating a large proportion of high-ability, but poor, agents. After the removal of mobility distortions, it takes time for these agents to fully utilize their abilities in the entrepreneurship because of the financial friction.

This paper generates a slow rural-urban migration and also slow convergence of other macroeconomic variables. The urbanization process is accompanied with the occupation change from the rural farmer to the urban worker or entrepreneur. It provides not only workers but also potential entrepreneurs to the economy. Therefore, the rural-urban migration itself is a source for the persistence of development dynamics.

One limitation in the model is that all the initial distortions are removed together at the beginning of the period. It is a simplification to illustrate the idea and solve the model. A slow and gradual removal pace is closer to the reality and should produce a slower urbanization result than the current paper.

There are other extensions interesting to investigate. One example is to see what will happen when the distortions are removed one by one. The results will provide policy advice when distortions are varied in the degree of political objection. Another interesting issue is to identify, from the data, the degree of mobility distortion in China. Hsieh et al. [2013] recently propose a framework to measure the occupation distortions in the US. It is worthwhile to extend their framework to China. All of these are questions for future research.

1.7 Appendix A: Data Description

The date sources are the Penn World Table 8.0 and the National Bureau of Statistics of China.

The time period in Figure 1 is from 1992 to 2011. The rural employment shares of the population are taken from the National Bureau of Statistics of China.

The data of the output per capita, TFP and the capital output ratio come from the Penn World Table 8.0. A detailed construction process is described by Feenstra et al. [2015].

The output per capita relative to the US is constructed as following. Use the expenditure-side real GDP at chained PPPs in 2005 US dollar for outputs. Divide this sequence by the population in the country. This produces the output per capita for China and the US respectively. The final step is to divide the numbers of China by the ones of the US.

The TPF levels are one constructed at current PPPs. The US is 1 by construction.

The sequence of capital output ratio for China is constructed by using the capital stock at constant 2005 national prices and the real GDP at constant 2005 national prices.

1.8 Appendix B: Numerical Method

A nice and detailed reference for the numerical method used here is Achdou et al. [2015]. In this appendix, I summarize the main points of the numerical method.

The Hamiltonian-Jacobi-Bellman (HJB) equation, which characterizes the optimal individual policy function, is solved with value function iteration. When approximating the HJB equation on discrete grids, implicit approximation method, and upwind scheme are used to speed up the convergence[3]. The resulting system is a linear system in the unknown value function, and this system can be solved efficiently with a sparse matrix in the computer. Using the same approximation method for Kolmogorov Forward (KF) equation, it is a linear system in the unknown distribution. A nice shortcut is that the coefficient matrix in the KF equation is a transpose matrix already solved in the HJB equation iteration. This produces efficiency in solving the model.

In my model, the grids of ability and asset are same for all economies. The ability is divided evenly among $[0.2, 3.5]$ with grid number 200. The asset range is chosen to be $[0.01, 15000]$ such that the saving function of the highest ability intersects the zero line before reaching the maximum asset level. I use uneven grids in asset holding to capture the nonlinear part of saving function with more accuracy.

[3]Please refers to Barles and Souganidis [1991] for theoretical proof for convergence

1.8.1 Stationary Equilibrium

I solve the stationary equilibrium based on the method of Achdou et al. [2015]. The difference is that I have to iterate on both real interest rate r and wage w until capital and labor markets clear.

1. Guess the interest rate r^l in the stationary equilibrium.

2. Guess the wage $w^{l,m}$.

3. Given the prices r^l and $w^{l,m}$, solve the HJB equation and then the Kolmogorov forward equation. Both equations are solved by the finite difference method combined with implicit method and Upwind Scheme.

4. Check whether the labor market is clear. If not, update $w^{l,m}$ according the excess labor demand with bisection method. Repeat step 3-4 until $w^{l,m+1}$ clears the labor market under r^l.

5. Check whether the capital market is clear under r^l and $w^{l,m+1}$. If not, update r^l according the excess capital demand with bisection method. Repeat step 3-5 until r^{l+1} clears the capital market.

HJB equation

The boundary condition for z implies $0 = \partial_z v(a, \underline{z}) = \partial_z v(a, \bar{z})$ for all a.

The state constraint boundary condition $\partial_a v(\underline{a}, z) \geq u'(c)$ for all z. Let $v(a_i, z_j) = v_{i,j}$.

The discrete version of HJB equation (implicit method):

$$\begin{aligned} \frac{v_{i,j}^{n+1} - v_{i,j}^{n}}{\Delta} + \rho v_{i,j}^{n+1} &= u(c_{i,j}^{n}) + \partial_a v_{i,j}^{n+1}\left[M_{i,j} + ra_i - c_{i,j}^{n}\right] \\ &\quad + \mu_j \partial_z v_{i,j}^{n+1} + \frac{1}{2}\sigma_j^2 \partial_{zz} v_{i,j}^{n+1} \end{aligned}$$

Use Upwind Scheme. The idea is to use the forward difference approximation whenever the drift of the state variable is positive and the backward difference approximation whenever it is negative.

$$\begin{aligned}
\partial_{a,B} v_{i,j} &= \frac{v_{i,j} - v_{i-1,j}}{\Delta a} \\
\partial_{a,F} v_{i,j} &= \frac{v_{i+1,j} - v_{i,j}}{\Delta a} \\
\partial_{z} v_{i,j} &= \frac{v_{i,j+1} - v_{i,j}}{\Delta z} \\
\partial_{zz} v_{i,j} &= \frac{v_{i,j+1} - 2v_{i,j} + v_{i,j-1}}{(\Delta z)^2}
\end{aligned}$$

The discrete HJB now is

$$\begin{aligned}
\frac{v_{i,j}^{n+1} - v_{i,j}^{n}}{\Delta} + \rho v_{i,j}^{n+1} &= u(c_{i,j}^n) + \partial_{a,F} v_{i,j}^{n+1} (s_{i,j,F}^n)^+ + \partial_{a,B} v_{i,j}^{n+1} (s_{i,j,B}^n)^- \\
&\quad + \mu_j \partial_z v_{i,j}^{n+1} + \frac{1}{2}\sigma_j^2 \partial_{zz} v_{i,j}^{n+1}
\end{aligned}$$

where $s_{i,j,F} = M_{i,j} + ra_i - c_{i,j,F}^n$ and $s_{i,j,B} = M_{i,j} + ra_i - c_{i,j,B}^n$ and $x^+ = \max\{x, 0\}$ and $x^- = \min\{x, 0\}$.

Substitute the definition in and simplify, we have

$$\begin{aligned}
\frac{v_{i,j}^{n+1} - v_{i,j}^{n}}{\Delta} + \rho v_{i,j}^{n+1} &= u(c_{i,j}^n) + x_{i,j} v_{i-1,j}^{n+1} + y_{i,j} v_{i,j}^{n+1} + z_{i,j} v_{i+1,j}^{n+1} + \chi_j v_{i,j-1}^{n+1} + \zeta_i v_{i,j+1}^{n+1} \\
x_{i,j} &= -\frac{(s_{i,j,B}^n)^-}{\Delta a} \\
y_{i,j} &= -\frac{(s_{i,j,F}^n)^+}{\Delta a} + \frac{(s_{i,j,B}^n)^-}{\Delta a} - \frac{\sigma_j^2}{(\Delta z)^2} - \frac{\mu_j}{\Delta z} \\
z_{i,j} &= \frac{(s_{i,j,F}^n)^+}{\Delta a} \\
\chi_j &= \frac{\sigma_j^2}{2(\Delta z)^2} \\
\zeta_j &= \frac{\mu_j}{\Delta z} + \frac{\sigma_j^2}{2(\Delta z)^2}
\end{aligned}$$

With $x_{1,j} = z_{I,j} = 0$ for all j, $v_{0,j}^{n+1}$ and $v_{I+1,j}^{n+1}$ are never used.

This a system of $I \times J$ equations. In matrix form

$$\frac{1}{\Delta}(v^{n+1} - v^n) + \rho v^{n+1} = u^n + A^n v^{n+1}$$

KF equation

Discretize the KF equation and use Upwind Scheme, we have a linear equation

$$A^T g = 0$$

Note, the coefficient matrix is a transpose of the A matrix in the HJB equation calculation.

1.8.2 Transition Dynamics

I need to find out the entire transition path. This needs an iteration on both the price functions $r(t)$ and $w(t)$ for $t = 0, ..., T$.

Guess price functions $r^l(t)$ and $w^{l,m}(t)$, iterate them with the following algorithm.

1. Given $r^l(t)$ and $w^{l,m}(t)$, solve HJB with terminal condition $v(a, z, T)$ backward to get value function $v(a, z, t)$, associated occupation choice and the saving policy function $s^l(a, z, t)$ for $t = 0, ..., T$.

2. Given the policy functions, solve KF equation with initial condition $g(a, z, 0)$ forward to get $g^{l,m}(a, z, t)$ for $t = 0, ..., T$.

3. Check if the labor market is clear each period. If not, take $g^{l,m}(a, z, t)$ and $r^l(t)$ as given, construct $\tilde{w}^{l,m}(t)$ to clear the per period labor market. Update the wage function $w^{l,m+1}(t) = \xi_w w^{l,m}(t) + (1 - \xi_w)\tilde{w}^{l,m}(t)$. Repeat 1-3 until wage function converges.

4. Given $g^{l,m}(a, z, t)$ and $w^{l,m+1}(t)$, check if the capital market is clear for each period. If not, construct $\tilde{r}^l(t)$ to clear capital market each period. Update interest rate function $r^{l+1}(t) = \xi_r r^l(t) + (1 - \xi_r)\tilde{r}^l(t)$.

5. Repeat step 1-4 until $r^{l+1}(t)$ is sufficiently close to $r^l(t)$ for $t = 0, ..., T$.

Solving the Time-Dependent HJB

The key approximation equation is

$$\rho v^n = u^{n+1} + A^{n+1} v^n + \frac{1}{\Delta}(v^{n+1} - v^n)$$

The consumption is approximated by the future value function.

Solving the Time-Dependent KFE

The KF equation evolves as following using an explicit method

$$\frac{g^{n+1} - g^n}{\Delta t} = (A^n)^T g^n \Rightarrow g^{n+1} = g^n + \Delta t (A^n)^T g^n$$

or an implicit method:

$$\frac{g^{n+1} - g^n}{\Delta t} = (A^n)^T g^{n+1} \Rightarrow g^{n+1} = \left(\mathbb{I} - \Delta t (A^n)^T\right)^{-1} g^n$$

Chapter 2

The Welfare Analysis of Depressed Migrant Wage in China: A Dynamic View

2.1 Introduction

The transition of a poor country to a modern, rich country is a joint process of capital accumulation, urbanization and TFP growth. When current developing countries begin this process, like China and India, they face two challenges. They have productive entrepreneurs, but those entrepreneurs are suffering from borrowing constraint. As a result, productive firms can not expand quickly. They have sufficient labor resources, but most of the labor are located in the low productive rural sector. The economy can produce more and the welfare will be higher if most of the labor force can engage in the productive urban sector. In 1960, China has 83.3% population lived in rural area. After 50 years, the numbers decreases to 50.8% and the majority is still poor. There could be more migrant workers if there were no financial frictions or labor market frictions. This paper focuses one particular labor market friction, that is, the depressed migrant wage in the urban sector.

How do financial friction and depressed urban migrant wage jointly affect the factor allocation and the development dynamics? How do they dynamically affect the welfare of the economy, measured by the total output? This paper answers these questions by building a dynamic heterogeneous agent model and quantifying it with data from China.

To understand the effect of depressed migrant wage, we need to build a full dynamic general equilibrium model with financial frictions. This is because the reallocation of capital and labor is a dynamic process. The depressed migrant wage, a friction on the labor market, has interaction with financial frictions on the urban entrepreneurs. And the effect is highly nonlinear. A model evaluating the effect of depressed wage with comparative statics will underestimate the effect and misses the dynamics during the transition.

The model economy starts from some distorted initial condition, and it is moving towards its stationary equilibrium but facing both financial and labor market frictions all the time. The initial condition can be seen as a point in a transitional path or a starting point after some distortion policy is removed. The financial market friction takes the form of collateral constraint for the entrepreneurs. The entrepreneurs have time-varying entrepreneurial productivity and different levels of wealth. At each point of time, some high productivity entrepreneurs may operate in small scale because they don't have enough collaterals while some low productivity entrepreneurs can still be active as the user cost of capital is low. This generates capital misallocation and a low level of urban sector productivity.

The friction in the labor market is characterized as a separated urban labor market between urban worker and migrant worker, and a depressed disposal income for the migrant workers. As a result, the incentive to migrate for rural labor force is lower. It creates a distortion on sectoral labor allocation. If migrant workers can accumulate new skills when working in the urban sector, the labor market friction also slows down the human capital accumulation used in the urban sector. As a result, the quantity and quality of urban worker are much lower in an economy with low urbanization. The impact from depressed wage is even bigger.

When those two frictions work together, financial friction produces a lower urban TFP and the output loss is strongly amplified by the lower urbanization rate and low capital accumulation. When rural labor with high working ability delay their migration due to the labor market friction, the equilibrium effective urban wage rate is higher and the profit for entrepreneurs is lower. Given the financial constraint, lower profit leads to a slower reallocation among the entrepreneur, so the urban TFP is lower and capital accumulates slower. Over time, lower urban TFP and capital level, in turn, generate a lower marginal product of urban labor. This is equivalent to a weaker pulling force for rural workers and a slower urbanization rate. As a result, a large output loss because the sectoral labor misallocation is severe in this case. The welfare loss is especially large if we think those rural labor force are majority poor in the developing country.

The model generates endogenous GDP growth, urban TFP growth,increasing investment to GDP ratio and urbanization. I calibrate it with the data from China and matched the transitional dynamics for China from 1986 to 2013. After the calibration, I do groups of counterfactual experiments. All experiments within each group use the same initial condition but differ in the frictions or economy structure. In the main experiment, I remove the labor market friction to predict what will happen to the economy. This simulation shows that using economy with both fractions as base, the urban TFP would be lower and shows a U-shape curve over time and a trough at -4% at around 6 years, while the output would be much lower with a trough at -38% at around 19 years.

Compare with aggregate growth accounting approach, the dynamic heterogeneous agent model used here has three advantages. First, it provides a detailed micro-foundation. The economy outcome from any policy change is a result of the response from rational agents. Second, the dynamic feature can provide accurate answers to policy change. For example, it worths to know how long the policy will take to have the effect and what the transitional dynamics will look like. Analysis base on steady state analysis will lose those dimension and may be misleading. Third, the productivity is endogenous. This is important because agents' behavior response to productivity change, and at the same time they affect the productivity process. A policy evaluation with fixed or exogenous productivity will have a bias.

2.1.1 Related Literature

The most related literature is the strand on misallocation due to financial frictions. Buera and Shin [2013] document the stylized facts of successful developing economies and quantitatively analyze how the financial frictions and resource misallocation can explain the observed development dynamics by emphasizing the endogenous TFP and increasing investment-output ratio. Moll [2014] analyzes theoretically how the persistency of idiosyncratic productivity shock under financial constraint determines the speed of convergence and steady state productivity losses. Midrigan and Xu [2014] use producer-level data and emphasize the distortion on entry and technology adoption rather than the intensive margin. Comparing with this literature, my model incorporates the labor market friction and emphasize its interaction with the financial friction and the resulting sectoral labor misallocation.

Restuccia and Rogerson [2008] and Hsieh and Klenow [2009] use implicit tax method to argue that resource misallocation shows up as a low level of TFP for developing countries. Gilchrist et al. [2013] use information on the dispersion in borrowing costs for a subset of US manufacturing firms and argue the efficiency loss is low. These papers focus on the steady state, while my model emphasizes the dynamic process.

There are other previous research focusing on the role of the labor market in urbanization. Todaro [1969] and Harris and Todaro [1970] address the relation between immigration and urban minimum wage within a two-sector model. In the equilibrium, the expected wage equals the rural wage. Lucas [2004] use a theoretical model to explain the urbanization by human capital accumulation. New migrants can accumulate the skills in the city and obtain urban jobs.

Some literature also focuses on labor market friction. Hayashi and Prescott [2008] discusses how the prewar patriarchy in Japan forced the son to stay in agriculture and induced a sectoral distortion and a depressed output level with a standard neoclassical two-sector growth model. And Jovanovic [2014] use a partnership setting to link mismatch in human capital and growth.

2.2 Stylized Facts for China

In recent years, there are many nice summaries on China's economy performance. Related papers are Zhu [2012], Meng [2012], Dollar and Jones [2013], and Chang et al. [2015]. China shares similar developing pattern like other Asian miracle countries, as documented by Buera and Shin [2013], but it has some particular features too.

The common features are increasing GDP per capita, TFP, investment rates and urbanization rate along the developing path. The unique features are relatively high investment rates as 40% of GDP and relatively low labor share. This paper doesn't focus on why these features arise, but focuses on the common features and how the labor market friction affects the economy.

The labor market in China is in the form of Hukou system, which is essentially a registration system. Each citizen has a Hukou, and is divided into urban Hukou and rural Hukou. Urban Hukou is obtained by born in the city and at least one parent of the child holds urban Hukou. There are other methods to get urban Hukou, but it is generally hard to get the urban Hukou through other ways. The Hukou system protects the interests of urban Hukou holders. First, urban workers are protected from migrant workers by having priorities in being hired and no additional hiring fees

Table 2.1: Education and Training

	no schooling	primary	secondary	tertiary	non-agricultural training
2012	1.5%	14.3%	78.5%	5.7%	25.6%
2014	1.1%	14.8%	76.8%	6.7%	32%

levied by the local government. This creates a separated urban labor market. Second, without urban Hukou, migrant workers cannot fully enjoy the social insurance and housing fund from local government. According to the National Bureau of Statistics of China (NBS), the participation rates of migrant worker in 2008 are 9.8% for the pension fund, 24.1% for work injury insurance, 13.1% for health insurance, 3.7% for unemployment insurance and 2.0% for childbirth insurance.

The scale of migrant workers in China is large and increasing over time. It increases from 225.4 million in 2008 to 273.9 million in 2014. Compared with US population in 2014, 318.9 million, the number of migrant workers is about 85.9% of US total population.

Among the migrant workers, the average education level and training experience are increasing over time. The data are shown in Table 2.1 and coming from a national survey of National Burea of Statistics (NBS) in China starting from 2012. Another fact worth noting is that the number of migrant workers is still increasing and migrant workers move back and forth between rural and urban sector. As a result, the average experience of working in the urban sector is increasing and these working experience and skills are diffused among the whole rural population.

2.3 Model

Time is continuous. The economy consists two sectors: rural sector and urban sector.

There is a continuum of entrepreneurs with measure one. They live in the city and make consumption-saving decisions.

There is a continuum of wage earners with measure one. They supply their labor inelastically. All workers are hand-to-mouth agents. According to their Hukou status, they are divided into two parts: L_c denotes those who have urban Hukou, and L_r denotes those with rural Hukou. The sum of those two is one, $L_c + L_r = 1$. I assume the government controls the amount of urban Hukou L_c

in the economy and take it as exogenous.

2.3.1 The problem of workers and the labor market frictions

Each worker is endowed with a work ability z. I assume $\log z$ is an Ornstein-Uhlenbeck (OU) process $d\log z = \theta_w(\mu_w - \log z)dt + \sigma_w dW_t$, A continuous time counterpart for AR(1) process. The stationary distribution of $\log z$ is a Gaussian distribution $N(\mu_w, \frac{\sigma_w^2}{2\theta_w})$. By Ito's lemma, the underlying process for z is a diffusion process $dz = \left(z\theta_w(\mu_w - \log z) + \frac{1}{2}z\sigma_w^2\right)dt + \sigma_w z dW = \mu_z dt + \sigma_z dW$.

I call workers with urban Hukou urban workers. For those workers without urban Hukou, if they work in rural sector, I call them farmers and if they work in urban sector, I call them migrant workers.

Incomes for the urban worker is set to wz, where w is the efficient wage in urban sector. For simplicity, I assume all the urban workers will be employed in city regardless of their working abilities[1]. The problem for urban workers is trivial. They are fully employed, hand-to-mouth agent without any occupation choice.

When working as farmers, workers earn a constant rural wage u, independent of their working ability. The migrant worker is paid at wz, but the disposal income for migrant worker is $\xi wz - \kappa$. Both ξ and κ are exogenous and capture the operational cost for migrant workers when they are living in the city. The difference is that one is proportion to wage and one is fixed over time. We adopt this implicit tax in the misallocation literature as a short cut for modeling the labor market friction. As most migrant workers do not have access to social insurance plan, they need to pay additional cost when they are temporary unemployed, ill or injured and additional cost for renting or buying house.

All the workers maximize their expected present discounted utility from consumption. I use $V(z(t), t)$ to denote the value function of workers with rural Hukou, $V_1(z(t), t)$ and $V_0(z(t), t)$ to denote the value function of being an migrant worker and a farmer respectively. Each period, given the current worker ability $z(t)$ and the wage rate $w(t)$, these workers need to choose their occupation: farmer or migrant worker. There is no uncertainty when they choose the occupation and there is no

[1] Unemployment is an interesting issue but not a focus here. I leave it for future work.

cost for switching occupation. The value functions satisfy the following equations:

$$\begin{aligned} V(z(t),t) &= \max\{V_0(z(t),t), V_1(z(t),t)\} \\ V_0(z(t),t) &= udt + e^{-\rho dt} EV(z(t+dt), t+dt) \\ V_1(z(t),t) &= (\xi w(t) z(t) - \kappa)\, dt + e^{-\rho dt} EV(z(t+dt), t+dt) \end{aligned}$$

The optimal policy is a cutoff policy summarized by $\bar{z}(t)$. For any $z(t) \geq \bar{z}(t)$, the occupation choice is an migrant worker. And for $z(t) < \bar{z}(t)$, they will be farmers. The cutoff $\bar{z}(t)$ is determined by

$$u = \xi w(t) \bar{z}(t) - \kappa$$

The Evolution of Density of Farmers and migrant Workers

Let $h(z,t)$ denote the working ability distribution for workers without urban Hukou. Given the cutoff $\bar{z}(t)$, I want to describe how the distribution evolves over time and figure out the corresponding part for migrant workers and farmers. For simplicity, the working ability distribution for urban Hukou workers is assumed to be the same as the one of rural Hukou workers.

The evolution of $h(z,t)$ can be described by the Fokker-Planck (FP) equation or Kolmogorov Forward (KF) equation[2]:

$$\partial_t h(z,t) = -\partial_z \left[\mu(z) h(z,t)\right] + \frac{1}{2}\partial_{zz}\left[\sigma^2(z) h(z,t)\right]$$

with $\int_0^{\bar{z}(t)} h_0(z,t)dz + \int_{\bar{z}(t)}^{+\infty} h_1(z,t)dz = 1 - L_c$, for any $t \geq 0$.

Let $h_0(z,t)$ and $h_1(z,t)$ denote the density functions of the farmers in rural sector and migrant workers in city. Given the cutoff policy, I have $h_0(z,t) = h(z,t)$ for $z \in (0, \bar{z}(t))$ and $h_1(z,t) = h(z,t)$ for $z \in [\bar{z}(t), +\infty)$.

[2] See for example Stokey (2009)

2.3.2 The Entrepreneur's Problem

I model the entrepreneur part closely following Moll (2014) for its tractability. All the individual variables are time varying in general. I drop out the time index when it is easy to distinguish. The entrepreneurs will be divided into lenders and borrowers. Those borrowers will be active in production.

The entrepreneurs are heterogeneous in their entrepreneurial productivity ϵ and asset holding in hand a. Entrepreneurs are risk averse and maximize their expected present discounted utility as[3]

$$E_0 \int_0^\infty e^{-\rho t} \log c(t) dt$$

The individual production function is assumed to be $y = f(\epsilon, k, l) = (\epsilon k)^\alpha l^{1-\alpha}$ where k denotes the capital and l denote effective labor units used in production. The urban labor market is competitive. Each entrepreneur is large enough such that the coming worker has the mean quality of all the workers in the urban labor market. Each entrepreneur faces financial constraint modeled as collateral constraint $k \leq \lambda a$ where $\lambda \in [1, \infty)$. When $\lambda = 1$, no one can borrow. The economy is in autarky. When $\lambda = \infty$, entrepreneur can borrow as much as she wants.

Assume there is a competitive capital rental market with a rental rate $R(t) = r(t) + \delta$. The individual asset holding evolves according to

$$\dot{a} = f(\epsilon, k, l) - wl - (r + \delta)k + ra - c$$

The production is intra-period. The problem is divided into a production problem and a consumption/saving problem. I can define the profit function $\Pi(a, \epsilon)$ from production:

$$\Pi(a, \epsilon) = \max_{k,l} \{ f(\epsilon, k, l) - wl - (r + \delta)k \ s.t \ k \leq \lambda a \}$$

It can be showed that the optimal choice for production is a corner solution $\underline{\epsilon}$, as factor demands and profits are linear in wealth a. The cutoff $\underline{\epsilon}$ is determined by $\underline{\epsilon}\pi = r + \delta$. Correspondingly, the

[3]Log utility can be extended to CRRA utility as Moll (2014) shows.

optimal capital holding policy is $k(a,\epsilon) = \begin{cases} \lambda a & \epsilon \geq \underline{\epsilon} \\ 0 & \epsilon < \underline{\epsilon} \end{cases}$, and labor demand function is $l(a,\epsilon) = \left(\frac{1-\alpha}{w}\right)^{\frac{1}{\alpha}} \epsilon k(a,\epsilon)$. The profit function then is summarized by $\Pi(a,\epsilon) = \max\{\epsilon\pi - r - \delta, 0\}\lambda a$ where $\pi = \alpha\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$.

Now turn to the consumption/saving problem. Given the log utility function, optimal savings policy is linear in wealth

$$\begin{aligned} \dot{a} &= s(\epsilon)a \\ c &= \rho a \end{aligned}$$

where $s(\epsilon) = \lambda \max\{\epsilon\pi - r - \delta, 0\} + r - \rho$ denotes the marginal propensity of saving.

Following Moll (2014), I assume the log of productivity follows another OU process: $d\log\epsilon = \frac{1}{\theta}(\mu - \log\epsilon)\,dt + \sigma\sqrt{\frac{1}{\theta}}dW$. By Ito's lemma, the productivity ϵ follows $d\epsilon = \frac{1}{\theta}\left(\frac{1}{2}\sigma^2 + \mu - \log\epsilon\right)\epsilon dt + \sigma\epsilon\sqrt{\frac{1}{\theta}}dW$. The nice feature of this setting is that the stationary distribution is invariant to the persistent parameter. The stationary distribution of ϵ is log-normal with mean zero and variance $\frac{\sigma^2}{2}$:$\log(\epsilon) \sim N\left(\mu, \frac{\sigma^2}{2}\right)$.

Let $g_t(a,\epsilon)$ denote the joint distribution and let $\varphi_t(a)$ and $\psi_t(\epsilon)$ denote the marginal distribution. These functions will be used to describe the equilibrium.

2.3.3 Equilibrium

The equilibrium is the combination of the time paths for prices $r(t)$ and $w(t)$ and corresponding individual quantities. Given these prices, the workers with rural Hukou maximize their utilities by optimally choosing the occupation; the entrepreneurs maximize their utilities by optimally making production and consumption/saving decisions, and all the markets are clear.

Capital market clearance requires:

$$\int k_t(a,\epsilon)dG_t(a,\epsilon) = \int a dG_t(a,\epsilon)$$

which says the capital demand from active entrepreneurs equals the total asset in the economy.

Labor market clearance equates the total effective labor demand from active entrepreneurs to the effective labor supply from migrant workers $\int zh_1(z,t)dz$ and corresponding urban workers $\bar{L}_c(t)$.

$$\int l_t(a,\epsilon)dG_t(a,\epsilon) = \int zh_1(z,t)dz + \bar{L}_c(t)$$

2.3.4 The aggregate characterization

In this subsection, I characterize the aggregate relationship between K_t, H_t, Y_t, the wealth share distribution $\omega(\epsilon,t)$, the distributions $h_1(z,t)$ and $h_0(z,t)$. The stationary equilibrium condition is derived in the appendix.

The joint distribution $G_t(a,\epsilon)$ has no stationary distribution. Instead, the wealth share distribution $\omega(\epsilon,t)$ do. The wealth share is defined as $\omega(\epsilon,t) = \frac{1}{K(t)}\int ag_t(a,\epsilon)da$, where $K(t) = \int ag_t(a,\epsilon)dad\epsilon$ is the total capital. I denote the total effective labor in city as $H(t) = \int l_t(a,e)dG_t(a,e) = \int zh_1(z,t)dz + \bar{L}_c$.

With market clear condition and optimal policy function, the aggregate variables satisfies the following equation system:

$$Y = \mathcal{E}K^{\alpha}H^{1-\alpha} + uL_r \tag{2.1}$$

$$\dot{K} = \alpha\mathcal{E}K^{\alpha}H^{1-\alpha} - (\rho+\delta)K \tag{2.2}$$

$$\mathcal{E} = \left(\frac{\int_{\underline{\epsilon}}^{\infty}\epsilon\omega(\epsilon)d\epsilon}{1-\Omega(\underline{\epsilon})}\right)^{\alpha} = \mathbb{E}_{\omega}[\epsilon|\epsilon \geq \underline{\epsilon}]^{\alpha} \tag{2.3}$$

$$\lambda(1-\Omega(\underline{\epsilon})) = 1 \tag{2.4}$$

Equation (2.1) describes that total output comes from urban sector $\mathcal{E}K^{\alpha}H^{1-\alpha}$ and rural sector uL_r. Equation (2.2) is aggregate capital evolution equation. Equation (2.3) describes how the urban TFP is determined. It is determined by the average productivity of active entrepreneurs, with the wealth share serves as a density. Equation (2.4) is derived from credit market clear condition and links the active cutoff for an entrepreneurs with financial friction and the wealth share distribution.

The factor prices are determined by

$$w(t) = (1-\alpha)\mathcal{E}K^{\alpha}H^{-\alpha} \tag{2.5}$$

$$r(t) = \frac{\underline{\epsilon}}{\mathbb{E}_{\omega}[\epsilon|\epsilon \geq \underline{\epsilon}]}(\rho+\delta)-\delta \tag{2.6}$$

Equation (2.5) says the effective wage is determined by the marginal product of effective urban labor supply. Equation (2.6) says the real interest rate is determined by the marginal product of the cutoff entrepreneur.

Following Moll [2014], given the specific diffusion process for entrepreneur productivity, the wealth shares can be proved to obey the following FP equation

$$\frac{d\omega(\epsilon,t)}{dt} = \left(s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)}\right)\omega(\epsilon,t) - \frac{d}{d\epsilon}\left[\frac{1}{\theta}\left(\frac{1}{2}\sigma^2 + \mu - \log\epsilon\right)\epsilon\omega(\epsilon,t)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\frac{1}{\theta}\sigma^2\epsilon^2\omega(\epsilon,t)\right]$$

2.3.5 Calibration Strategy

To provide a quantitative analysis of the joint effect of the frictions, I quantify the model with the data from China.

China has initiated a series of economy reforms since the 1980s. I treat the year 1986 as the beginning year of the economy. The economy is already on a transitional path in 1986. The degree of frictions is assumed to be constant over this 1986-2013 period. I assume that the steady state of urbanization rate is 80% under the frictions in the long run. Given this calibration strategy, I need to specify an initial condition for the economy and the degree of frictions for the economy since 1986. After calibration, I can answer the questions that what the economy will behave if frictions are eliminated or the economy structure is changed in 1986.

The initial condition is modeled by means of two distributions: the worker's ability distribution and the entrepreneur wealth share distribution. These two distributions are distorted from their stationary distribution.

My approach is same as Moll [2014] but different from Buera and Shin [2013], who model the economy transiting from one distorted steady state to another new efficient steady state after some once-for-all economic reform took at beginning. In Buera and Shin [2013], the initial source of distortion is modeled as exogenous idiosyncratic tax distortion, following Restuccia and Rogerson [2008], and the new steady state is calibrated from the US data.

I emphasize that my calibration strategy is equivalent to theirs in essential. An initial condition can be calibrated as the steady state under exogenous idiosyncratic taxes or can be specified directly.

2.3.6 Calibration of Parameters

The model is calibrated at year frequency and solved with finite difference method. Details and steps on the numerical method are in the appendix.

To do numerical analysis, I assume a reflective boundary for the worker ability process and entrepreneur productivity process. The maximum value for entrepreneur productivity ϵ is set to be the 95th percentile of its stationary log-normal distribution. And the maximum value for worker ability z is the 99th percentile of its stationary log-normal distribution.

Exogenous Parameters

Table 2.2 shows the exogenous given or calibrated parameters used in the model. They are borrowing from other literature or calibrated with the steady state condition. The model implies a relationship between total private debt D and the total capital K: $\frac{D}{K} = 1 - \frac{1}{\lambda}$ and $\frac{D}{Y} = \frac{D}{K}\frac{K}{Y} = \left(1 - \frac{1}{\lambda}\right)\frac{K}{Y}$ in steady state. I associate D with private credit by deposit money banks and other financial institutions to GDP from the 2013 financial data set constructed and maintained by Beck et al. [2009]. This definition for private credit follows Buera and Shin [2013]. In the data set, I can calculate that US $\lambda = 2.57$ and China $\lambda = 1.43$ in 2009.

Table 2.2: Exogenous Parameters

Entrepreneur Side			
Parameters	Meaning	Value	Source/Reason
α	capital share for urban production function	0.5	Bai, Hsieh and Qian 2006
ρ	discounted rate	0.05	risk-free rate 5%
δ	depreciation rate	0.10	$\frac{K}{Y_{city}} = 3.33$
θ	persistency parameter for entrepreneur productivity	6.1531	$e^{-\frac{1}{\theta}} = 0.85, \frac{\sigma^2}{2\theta} = 0.9684$
σ	diffusion parameter for entrepreneur productivity	1.3891	Asker, Collard-Wexler and Loecker 2014
μ	mean parameter for entrepreneur productivity	0	Normalization
λ	financial frictions	1.43	Calculated from Data of Beck, Demirguc-Kunt and Levine (2009)

Worker Side			
Parameters	Meaning	Value	Source/Reason
L_c	Urban Hukou labor share	24.73%	NBS 2000
μ_w	mean of working ability	0	normalization
$theta_w$	drift parameter	0.0619	$e^{-\frac{1}{\theta_w}} = 0.94$ Chen, Chen, Chen and Qiu (China
data 2008-2009)			
$sigma_w$	diffusion parameter	0.2268	$\frac{\sigma^2}{2\theta} = 0.22$ Chen, Chen, Chen and Qiu (China
data 2010)			
ξ	depressed migrant wage	0.66	Song (2013)

Calibrated Parameters

Given the exogenous parameters, including the friction parameters, I use steady state urbanization rate to calibrate the operational cost κ. The US urbanization rate is around 84% in the 2010s, but the US faces different degrees of frictions compared with China. Thus, the steady state urbanization rate is assumed to be 80% for China. I normalize the rural productivity u to be one, and set $\kappa = 0.615$ to achieve 80% urbanization rate in the long run in the model.

The initial distributions are characterized by the initial mean of log worker abilities distribution $\mu_w^{distorted}$, the initial mean of entrepreneur wealth share distribution $\mu^{distorted}$ and the initial capital stock K_0. These initial condition parameters are set to match initial capital-output ratio to be one and match the urbanization rate at 2013 and the ratio of urban-rural average income in 2013.

Table 2.3: Calibration for Initial Condition

Value	Meaning	Target	Model
$\mu_w^{distorted} = -1.66$	initial mean of log worker ability	urban average income/ rural average income = 3.03 (NBS 2013)	3.07
$u^{distorted} = -1.70$	initial mean of wealth share	urbanization rate = 53.72% (NBS 2013)	53.74%
$K_0 = 0.88$	Initial capital	$\frac{K_0}{Y_0} = 1$	1.00

2.3.7 Quantitative Results from Benchmark Model

The benchmark model is a two-sector economy suffering from labor and financial market fractions. The long-run effect of frictions will be larger if we have a large degree of any friction. A larger labor market friction (high κ and low ξ) and larger financial friction (low λ) will lead to a larger loss in TFP, output and low level of urbanization rate. This paper focuses on the transitional dynamics.

Transitional Dynamics

Figure 2.1 shows initial distortion on the worker ability and entrepreneur wealth shares. Initially (the dotted red line), more capital is holding in the low productivity entrepreneurs' hand and most workers have low working abilities. In the long run (the blue line), the wealth will be concentrated in hands of more productive entrepreneurs and workers' ability improves as urbanization rate is high.

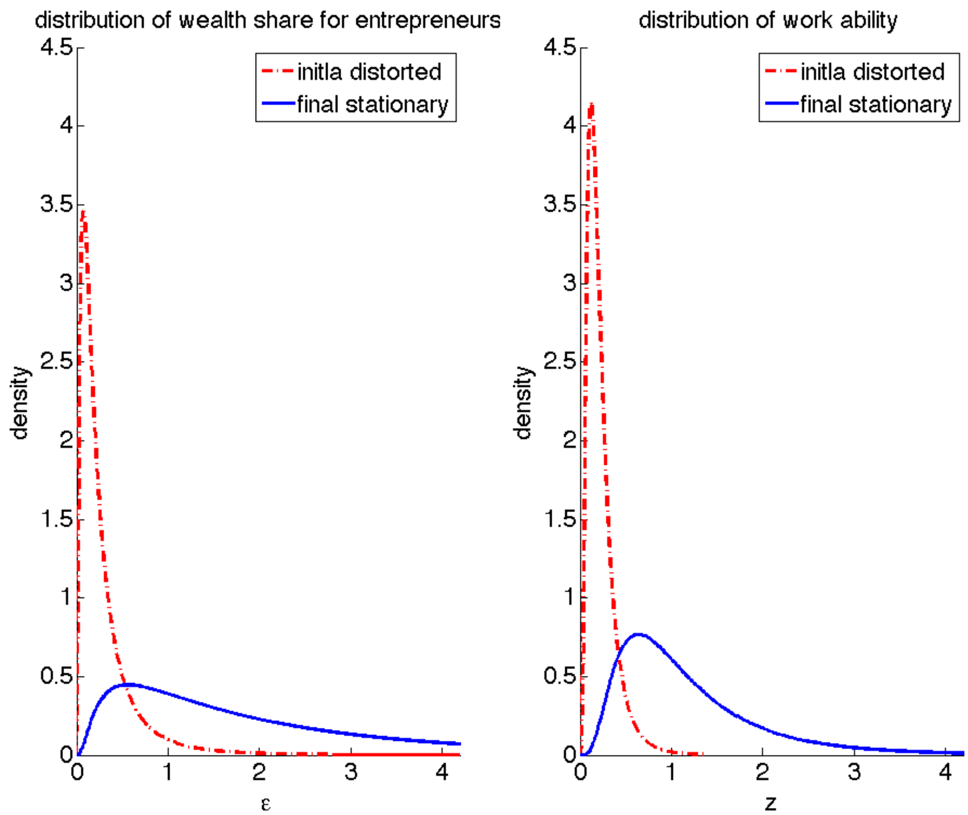


Figure 2.1: Initial Distortion

The model generates many features consistent with the facts observed in the data. First, the model generates an increasing and slowly urbanization process (see Figure 2.5). The speed is controlled by the labor market friction, given financial friction fixed. Second, the model produces an increasing urban TFP. Figure 2.3 shows the urban TPF dynamics and normalized it by the first-period value. As stated before, urban TFP gains from the reallocation of wealth in the hand of entrepreneurs. As more productive entrepreneurs produce more and less productive entrepreneurs quit the urban production, the TFP grows. The process takes time as the entrepreneur faces financial constraints.

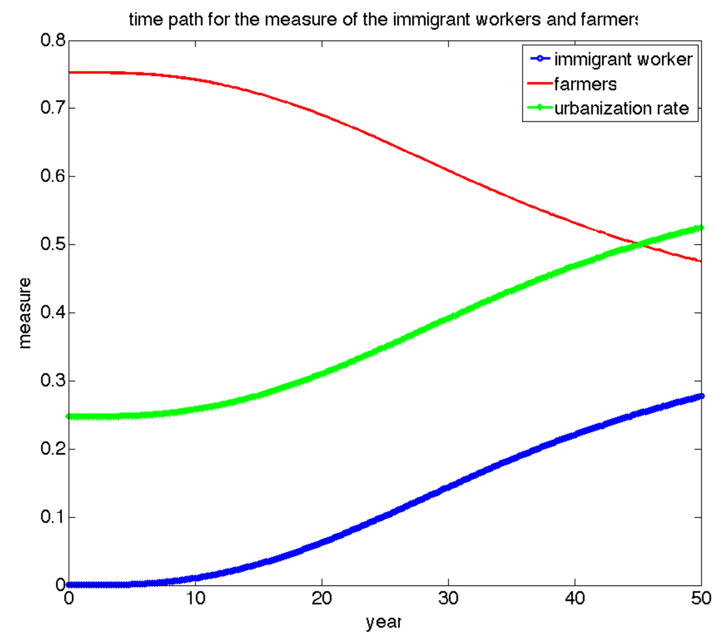


Figure 2.2: Urbanization

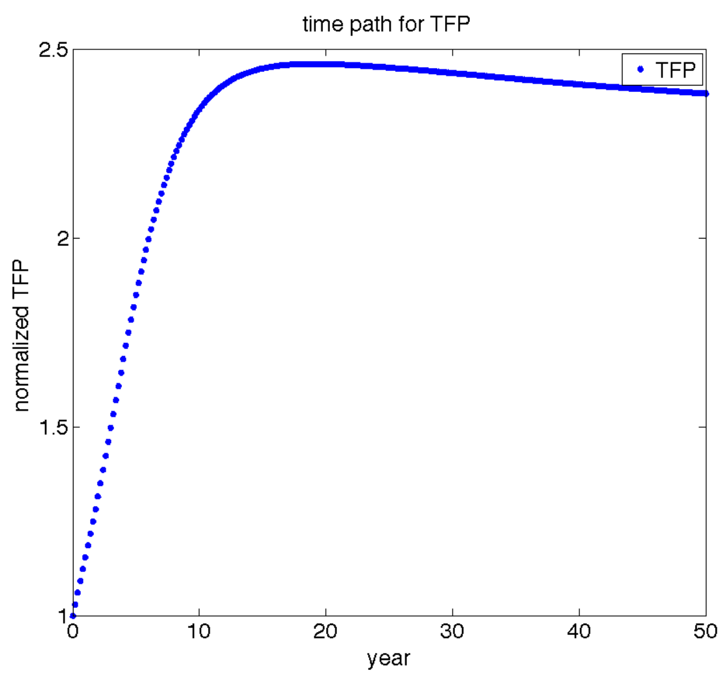


Figure 2.3: TFP

Other major macro variables shown in Figure 2.4 are consistent with the observed data. For example, the output per capita is increasing all the time, the investment-output ratio is increasing and high, the labor share is decreasing over time. The real interest rate increases at the beginning and drops afterward. It does not overshoot at beginning as the one in a neoclassical growth model studied by **??**, because the TFP moves increases slowly and initially both capital and effective labor used in the city are both small. So the capital-labor ratio is not small as in one sector model. The cutoff for active entrepreneur starts low. This is due to the effect of financial friction and initial capital misallocation. As time passes, the cutoff increases, and inefficient entrepreneurs quit. The cutoff for being an immigrant worker is dropping off all the time. As economy accumulates more capital and reallocates the capital in the hand of productive entrepreneurs, the wage rate is increasing, so it attracts more labor to the urban area.

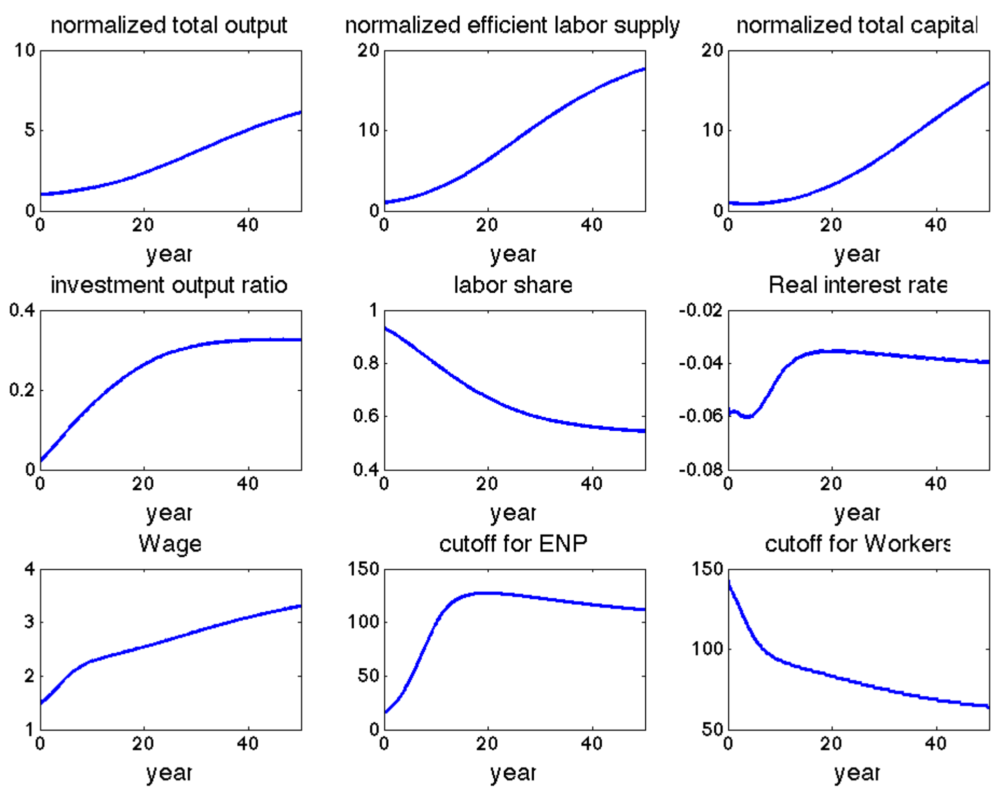


Figure 2.4: Dynamics for Other Variables

2.3.8 The Interaction between Frictions and Two Sector Structure

In this section, I compare different transitional paths in different counterfactual experiments. All experiments take the two-sector structure as fixed. Experiments comparing the economy with and without two-sector structure is in the appendix. Also, all experiments take the evolution of worker ability as exogenous. One extension of the endogenous evolution of worker ability is discussed in the extension part.

I first compare four economies: one without any friction, one with financial friction only, one with labor market friction only and one with both frictions. The four economies differ in friction but start from the same initial condition.

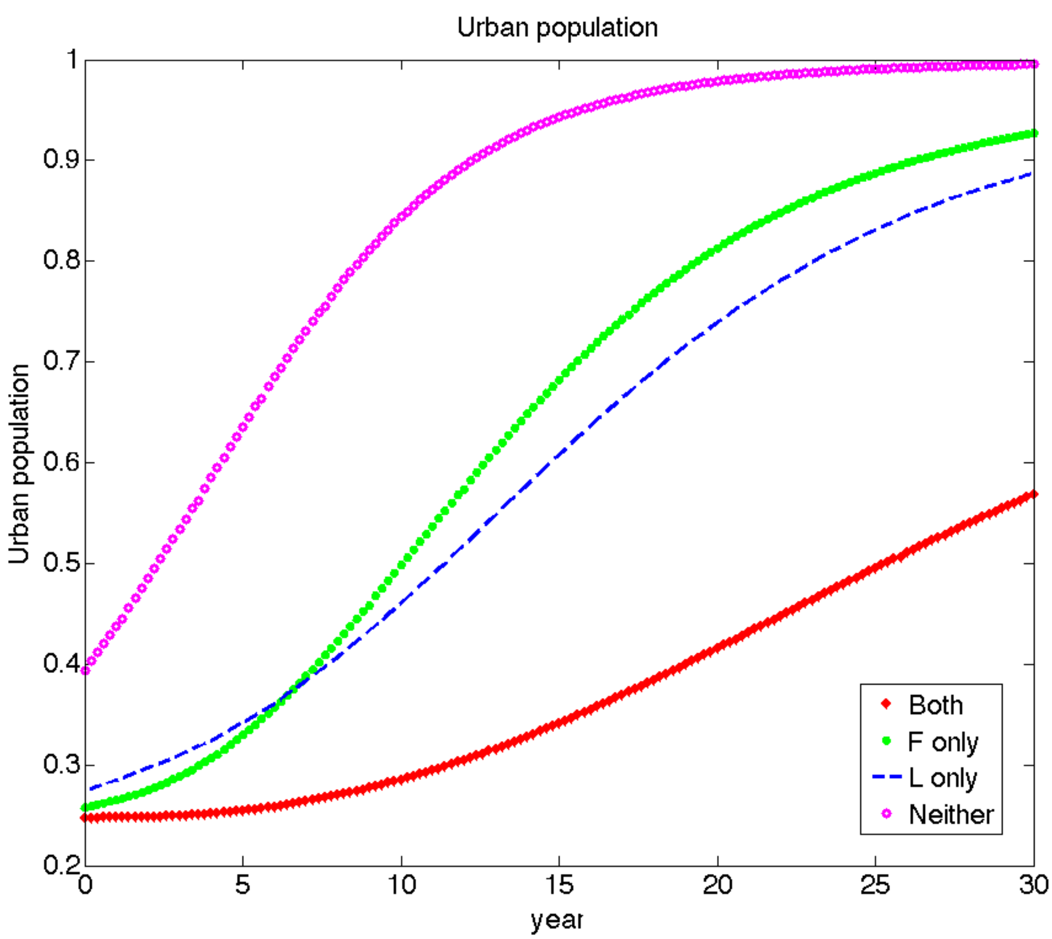


Figure 2.5: Urbanization Comparison

In Figure 2.5, it is clear to see that the financial friction has its maximum effect when there is labor market friction. Compare with the non-friction case, each friction has slowed down the urbanization speed. However, the joint effect of two frictions is much larger than the simple sum of separate effect. This nonlinearity of joint effect is showing up clearly from urbanization at 30 years. The non-friction economy has an almost 100% urbanization rate, while 92.6% for one with financial friction only, 88.57% for one with labor market only and 56.53% for one with both frictions.

Figure 2.6 provides a further comparison of urban TFP and total output between the economy with financial friction only and one with both frictions. This comparison can be thought as a policy change at the beginning of the period to remove the labor market friction and let the economy evolve.

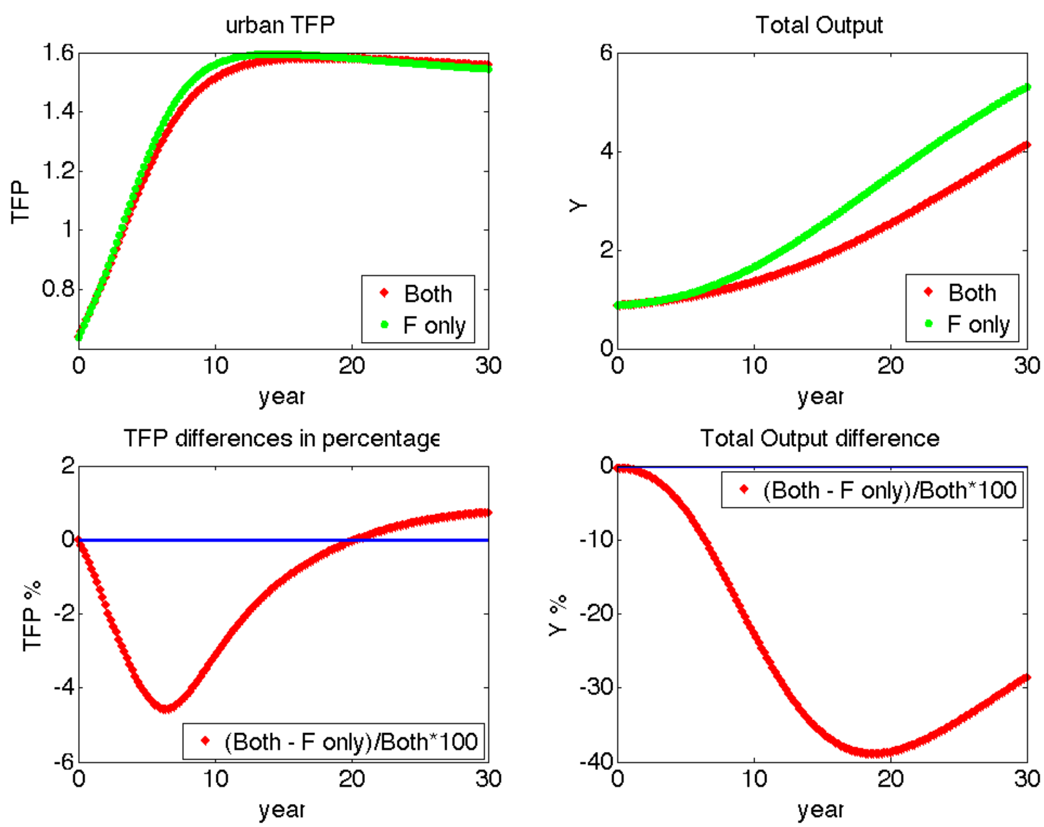


Figure 2.6: Urban TFP comparison

As the first column of figure shows, both economies have an increasing urban TFP over time as the capital misallocation is relieving endogenously. Use the benchmark economy as the base, the urban TFP difference drops to -4.6% at trough after 8 years. However, this small difference is strongly amplified by the labor market friction. In the second column of Figure 6, the output gap is keeping dropping beyond 8 years and touch the trough at -39% after 19 years. The urban sector uses the immigrant worker to produce. When the urbanization rate is lower in the economy with both friction, the output will be much less. Labor market friction slows down the wealth reallocation and the urbanization, by doing so, it amplifies the effect of financial friction strongly.

2.3.9 An Extension of the Model

In the benchmark calibration, the evolution of worker ability is set as exogenous. In this section, I discuss how endogenous evolution of worker ability can be a further channel to amplify the misallocation over time.

When urbanization happens, there are two ways increasing efficient labor supply in the urban sector. If the worker ability distribution is fixed, any new migrant worker increases the effective labor supply. This is the allocation channel. If the amount of migrant worker is fixed, improvement in the worker ability distribution will increase the effective labor supply. This is the human capital upgrading channel. The human capital upgrading channel will play an endogenous and larger role if itself is linked with the level of urbanization rate. Intuitively, slow urbanization leads to lower human capital upgrading, and the latter will slow down the urbanization further given the frictions in the economy.

It is reasonable to think that the speed of human capital upgrading will depend on how many workers are engaged in the urban production. When more workers are in the city, the worker ability will have a higher mean.

This can be modeled as an endogenous shift in the mean of the long run stationary distribution of the log of worker ability. Suppose the shift relationship is linear in initial measure : $\tilde{\mu}_w = \mu_w^{distortion} L_r + \mu_w(1 - L_r)$. As more workers are in the urban sector, the mean $\mu(z)$ is closer to μ_w by assigning more weights on it. The FP equation for worker ability distribution is

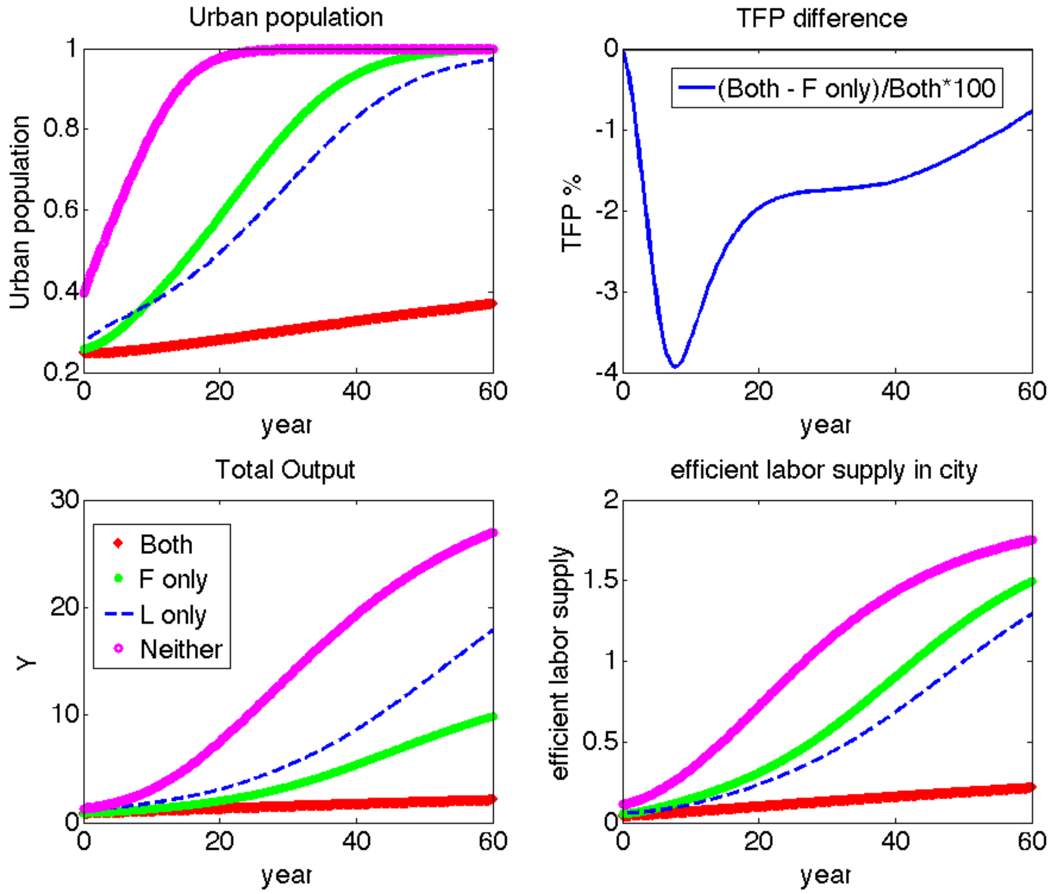


Figure 2.7: Comparison with endogenous work ability

$$\partial_t h(z,t) = -\partial_z \left[\tilde{\mu}(z)h(z,t)\right] + \frac{1}{2}\partial_{zz}\left[\sigma^2(z)h(z,t)\right]$$

with modified $\tilde{\mu}(z) = z\theta_w(\tilde{\mu}_w - \log z) + \frac{1}{2}z\sigma_w^2$.

Figure 2.7 presents a simple numerical simulation from the benchmark model. The experiments compares four economies with different frictions but with human upgrading channel specified as endogenous. The dependence of human capital upgrading on urbanization clearly exaggerates the effects of frictions. This can be seen from the response of urbanization rate, TFP, efficient labor supply and output.

2.4 Conclusion

To understand the development dynamics for countries like China and India, I argue that, in addition to the financial market friction, the labor market friction dimension and two sector economy structure are important.

Financial market friction restricts high ability entrepreneurs to expand and lower the TFP level from capital using. This paper argues that the two sector economy setting and labor market friction affect how fast the capital reallocates. Lower labor supply pushes up wages and depress profits. Thus, the TFP gain from capital reallocation is lower with labor friction and two-sector economy. The counterfactual experiment shows that capital reallocation is relatively fast, and the urban TFP loss is small. However, labor market friction can amplify these differences and leads to a large total output loss.

An economy with low urbanization rate suffers from a high inequality in consumption. Thus, with welfare in mind, when thinking about policy issues in financial friction, it will be better to analyze an economy with labor market friction and calculate how much of the gain in development can be translated into the welfare of the majority.

The future work is to analyze within the framework that how the unemployment in the urban area can be an endogenous source for government policy in urban labor market and how does this affect the whole economy.

2.5 Appendix

2.5.1 Stationary Equilibrium

Market clear conditions:

$$\begin{aligned} \int k(a,\epsilon)dG(a,\epsilon) &= \int a dG(a,\epsilon) \\ \int l(a,\epsilon)dG(a,\epsilon) &= \int_{z_L}^{\infty} e^{z}h(z)dz + \bar{L}_c = H \end{aligned}$$

The aggregate variables obey the following equations:

$$Y = \mathcal{E}K^{\alpha}H^{1-\alpha} + uL_r$$

$$\alpha\mathcal{E}K^{\alpha-1}H^{1-\alpha} = \rho + \delta$$

$$\mathcal{E} = \left(\frac{\int_{\underline{\epsilon}}^{\infty} \epsilon\omega(\epsilon)d\epsilon}{1-\Omega(\underline{\epsilon})}\right)^{\alpha} = \mathbb{E}_{\omega}[\epsilon|\epsilon \geq \underline{\epsilon}]^{\alpha}$$

$$\lambda(1-\Omega(\underline{\epsilon})) = 1$$

The factor prices are

$$\begin{aligned} w &= (1-\alpha)\mathcal{E}K^{\alpha}H^{-\alpha} \\ r &= \frac{\underline{\epsilon}}{\mathbb{E}_{\omega}[\epsilon|\epsilon \geq \underline{\epsilon}]}(\rho+\delta) - \delta \end{aligned}$$

The stationary wealth shares obey

$$0 = s(\epsilon)\omega(\epsilon) - \frac{d}{d\epsilon}\left[\frac{1}{\theta}\left(\frac{1}{2}\sigma^2 + \mu - \log\epsilon\right)\epsilon\omega(\epsilon)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\frac{1}{\theta}\sigma^2\epsilon^2\omega(\epsilon)\right]$$

The two boundary condition for the 2nd order ODE are

$$\begin{aligned} \omega(0) &= 0 \\ \int \omega(\epsilon)d\epsilon &= 1 \end{aligned}$$

2.5.2 Numerical Method

I solve the system with finite difference methods. First, I summarize how to solve the dynamic system.

There are 12 unknowns in the dynamic system:

$\{Y(t), \mathcal{E}(t), K(t), H(t), \omega(\epsilon, t), h(z, t), \underline{\epsilon}(t), \bar{z}(t), w(t), r(t), \pi(t), s(\epsilon, t)\}$. I solve this system by the following steps.

Given the wealth share distribution $\omega(\epsilon, t)$, I find out the cutoff $\underline{\epsilon}$, using credit market clearance condition:

$$\lambda(1 - \Omega(\underline{\epsilon})) = 1 \tag{2.7}$$

This provides efficiency factor from entrepreneurs

$$\mathcal{E} = \left(\frac{\int_{\underline{\epsilon}}^{\infty} \epsilon\omega(\epsilon)d\epsilon}{1 - \Omega(\underline{\epsilon})} \right)^{\alpha} = \mathbb{E}_{\omega}[\epsilon|\epsilon \geq \underline{\epsilon}]^{\alpha} \tag{2.8}$$

and the factor prices

$$r = \frac{\underline{\epsilon}}{\mathbb{E}_{\omega}[\epsilon|\epsilon \geq \underline{\epsilon}]}(\rho + \delta) - \delta \tag{2.9}$$

Given $K(t)$, $\mathcal{E}(t)$ and $h(z, t)$, I can derive $\bar{z}$ from the labor market clear condition, the cutoff policy function:

$$\begin{aligned} u &= \xi w \bar{z} - \kappa \\ w &= (1 - \alpha)\mathcal{E}K^{\alpha}H^{-\alpha} \end{aligned}$$

Then I have $\frac{u+\kappa}{\xi} = (1 - \alpha)\mathcal{E}K^{\alpha}H^{-\alpha}\bar{z}$, that is

$$(u + \kappa)\left(\int_{\bar{z}}^{\infty} zh_1(z)dz + \bar{L}_c\right)^{\alpha} = \xi(1 - \alpha)\mathcal{E}K^{\alpha}\bar{z} \tag{2.10}$$

The LHS is decreasing in $\bar{z}$, starting from $\frac{u+\kappa}{\xi}\left(\int_0^{\infty} zh_1(z)dz + \bar{L}_c\right)^{\alpha}$ to $\frac{u+\kappa}{\xi}\bar{L}_c^{\alpha}$. The RHS is increasing in $\bar{z}$, starting from 0 to ∞. A unique solution for $\bar{z}$ is guaranteed for reasonable parameters.

Next step, I calculate $w(t)$ and $H(t)$ by

$$w(t) = (1 - \alpha)\mathcal{E}K^{\alpha}H^{-\alpha} \tag{2.11}$$

$$H = \int_{\bar{z}}^{\infty} zh_1(z)dz + \bar{L}_c = L_1 \int_{\bar{z}}^{\infty} z\frac{h_1(z)}{\int_{\bar{z}}^{\infty} h_1(z)dz}dz + \mathcal{Z}_c L_c = \mathcal{Z}_1 L_1 + \mathcal{Z}_c L_c \tag{2.12}$$

Combined with $H(t)$ and $K(t)$, I have output

$$Y = \mathcal{E}K^{\alpha}H^{1-\alpha} + \bar{L}_c \tag{2.13}$$

Also, I can derive the dynamics of capital evolution

$$\dot{K} = \alpha\mathcal{E}K^{\alpha}H^{1-\alpha} - (\rho + \delta)K \tag{2.14}$$

The Dynamics for the Distributions $\{\omega(\epsilon, t), h(z, t)\}$**:**

The dynamics for the wealth share distribution are described by the FP equation. In general

$$\frac{d\omega(\epsilon,t)}{dt} = \left(s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)}\right)\omega(\epsilon,t) - \frac{d}{d\epsilon}\left[\mu(\epsilon,t)\omega(\epsilon,t)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2(\epsilon)\omega(\epsilon,t)\right]$$

With the specific diffusion process $d\epsilon = \frac{1}{\theta}\left(\frac{1}{2}\sigma^2\epsilon - \epsilon\log\epsilon\right)dt + \sigma\epsilon\sqrt{\frac{1}{\theta}}dW$, I have the wealth share equation:

$$\frac{d\omega(\epsilon,t)}{dt} = \left(s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)}\right)\omega(\epsilon,t) - \frac{d}{d\epsilon}\left[\frac{1}{\theta}\left(\frac{1}{2}\sigma^2\epsilon - \epsilon\log\epsilon\right)\omega(\epsilon,t)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\frac{1}{\theta}\sigma^2\epsilon^2\omega(\epsilon,t)\right] \tag{2.15}$$

where $\omega(\epsilon) = \frac{1}{K}\int a g_t(a,\epsilon)da$; $K = \int a g(a,\epsilon)da d\epsilon$.

The elements in FP equation are

$$s(\epsilon,t) = \lambda\max\{\epsilon\pi - r - \delta, 0\} + r - \rho \tag{2.16}$$

$$\pi(t) = \alpha\left(\frac{1-\alpha}{w(t)}\right)^{\frac{1-\alpha}{\alpha}} \tag{2.17}$$

There are two boundary conditions for the 2nd order PDE.

$$\omega(0,t) \quad = \quad 0$$

$$\int \omega(\epsilon, t) d\epsilon = 1$$

The dynamics for the work ability distribution:

$$\frac{d}{dt} h(z,t) = -\frac{d}{dz}\left[\mu(z) h(z,t)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2(z) h(z,t)\right] \tag{2.18}$$

where $dz = \left(z\theta_w(\mu_w - \log z) + \frac{1}{2} z\sigma_w^2\right) dt + \sigma_w z dW = \mu(z)dt + \sigma(z)dW$

There are two boundary conditions for the 2nd order PDE.

$$h(0,t) = 0$$
$$\int h(z,t) dz = 1$$

The above equations consist the system. Now I turn to discrete version of the dynamic equations and solve them numerically.

Capital equation can be written as:

$$\dot{K} = \alpha \mathcal{E} K^{\alpha} H^{1-\alpha} - (\rho + \delta) K$$

$$K_{n+1} = dt\left(\alpha \mathcal{E} K_n^{\alpha} H^{1-\alpha} - (\rho + \delta) K_n\right) + K_n$$

the density function of $\omega(\epsilon, t)$

FP equation:

$$\frac{d\omega(\epsilon,t)}{dt} = \left(s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)}\right)\omega(\epsilon,t) - \frac{d}{d\epsilon}\left[\frac{1}{\theta}\left(\frac{1}{2}\sigma^2\epsilon - \epsilon\log\epsilon\right)\omega(\epsilon,t)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\frac{1}{\theta}\sigma^2\epsilon^2\omega(\epsilon,t)\right]$$

Put into another way

$$\frac{d\omega(\epsilon,t)}{dt} = \left(s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)}\right)\omega(\epsilon,t) - \frac{d}{d\epsilon}\left[\mu(\epsilon)\omega(\epsilon,t)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2(\epsilon)\omega(\epsilon,t)\right]$$

where $\mu(\epsilon) = \frac{1}{\theta}\left(\frac{1}{2}\sigma^2\epsilon - \epsilon\log\epsilon\right)$ and $\sigma^2(\epsilon) = \frac{1}{\theta}\sigma^2\epsilon^2$.

Simplify it into a differential equation of ω

$$\frac{d\omega(\epsilon,t)}{dt} = a(\epsilon,t)\omega(\epsilon,t) + b(\epsilon,t)\frac{d\omega(\epsilon,t)}{d\epsilon} + c(\epsilon,t)\frac{d^2\omega(\epsilon,t)}{dz^2}$$

where

$$\begin{aligned}
a(\epsilon,t) &= s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)} - \mu'(\epsilon) + \frac{1}{2}(\sigma^2)''(\epsilon) \\
b(\epsilon) &= -\mu(\epsilon) + (\sigma^2)'(\epsilon) \\
c(\epsilon) &= \frac{1}{2}\sigma^2(\epsilon)
\end{aligned}$$

in general.

And specifically, I have

$$\begin{aligned}
a(\epsilon,t) &= s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)} - \frac{1}{\theta}\left(\frac{1}{2}\sigma^2 - \log\epsilon - 1\right) + \frac{1}{\theta}\sigma^2 \\
&= s(\epsilon,t) - \frac{\dot{K}(t)}{K(t)} + \frac{1}{\theta}\left(\frac{1}{2}\sigma^2 + \log\epsilon + 1\right) \\
b(\epsilon) &= -\frac{1}{\theta}\left(\frac{1}{2}\sigma^2\epsilon - \epsilon\log\epsilon\right) + \frac{2}{\theta}\sigma^2\epsilon \\
c(\epsilon) &= \frac{1}{2\theta}\sigma^2\epsilon^2
\end{aligned}$$

Using the evolution of $\dot{K}$, I have

$$a(\epsilon,t) = s(\epsilon,t) - \alpha\mathcal{E}K^{\alpha-1}H^{1-\alpha} + (\rho+\delta) - \frac{1}{\theta}\left(\frac{1}{2}\sigma^2 - \log\epsilon - 1\right) + \frac{1}{\theta}\sigma^2$$

$$
\begin{aligned}
b(\epsilon) &= -\frac{1}{\theta}\left(\frac{1}{2}\sigma^2\epsilon - \epsilon\log\epsilon\right) + \frac{2}{\theta}\sigma^2\epsilon \\
c(\epsilon) &= \frac{1}{2\theta}\sigma^2\epsilon^2
\end{aligned}
$$

- Use a forward-difference approximation in the time dimension
- Use a central-difference approximation in the space dimension
- and write the right hand side in an implicit scheme (evaluated at the $n+1$ time level)

I have $\frac{d\omega(\epsilon,t)}{dt} = a(\epsilon,t)\omega(\epsilon,t) + b(\epsilon,t)\frac{d\omega(\epsilon,t)}{d\epsilon} + c(\epsilon,t)\frac{d^2\omega(\epsilon,t)}{dz^2}$ as

$$
\frac{\omega_i^{n+1} - \omega_i^n}{\Delta t} = a_i^n\omega_i^{n+1} + b_i\frac{\omega_{i+1}^{n+1} - \omega_{i-1}^{n+1}}{2\Delta\epsilon} + c_i\frac{\omega_{i+1}^{n+1} - 2\omega_i^{n+1} + \omega_{i-1}^{n+1}}{(\Delta\epsilon)^2}
$$

Then I have a system in $\{\omega_i^{n+1}, \omega_i^n\}$:

$$
x_i\omega_{i-1}^{n+1} + y_i\omega_i^{n+1} + z_i\omega_{i+1}^{n+1} = \omega_i^n
$$

where

$$
\begin{aligned}
x_i &= b_i\frac{\Delta t}{2\Delta\epsilon} - c_i\frac{\Delta t}{(\Delta\epsilon)^2} \\
y_i &= 1 - a_i\Delta t + c_i\frac{2\Delta t}{(\Delta\epsilon)^2} \\
z_i &= -b_i\frac{\Delta t}{2\Delta\epsilon} - c_i\frac{\Delta t}{(\Delta\epsilon)^2}
\end{aligned}
$$

This system holds for $i = 2, ..., I-1$. I have $I-2$ equations in $\omega_2^{n+1}, ..., \omega_I^{n+1}$.

The boundary conditions are

$$
\begin{aligned}
\omega_1^{n+1} &= 0 \\
\Delta\epsilon\sum_{j=1}^{I}\omega_j^{n+1} &= 1
\end{aligned}
$$

which give 2 additional equations.

For a given $\{\omega_i^n\}_{i=2}^{I-1}$, I have a system of $I-1$ equations in $I-1$ unknowns $\{\omega_i^{n+1}\}_{i=2}^{I}$. I can summarize this in a matrix form

$$\begin{bmatrix} y_2 & z_2 & 0 & 0 & \dots & 0 \\ x_3 & y_3 & z_3 & 0 & \dots & 0 \\ 0 & x_4 & y_4 & z_4 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & x_{I-1} & y_{I-1} & z_{I-1} \\ \Delta\epsilon & \Delta\epsilon & \Delta\epsilon & \Delta\epsilon & \Delta\epsilon & \Delta\epsilon \end{bmatrix} \begin{bmatrix} \omega_2^{n+1} \\ \omega_3^{n+1} \\ \vdots \\ \vdots \\ \omega_{I-1}^{n+1} \\ \omega_I^{n+1} \end{bmatrix} = \begin{bmatrix} \omega_2^{n} \\ \omega_3^{n} \\ \vdots \\ \vdots \\ \omega_{I-1}^{n} \\ 1 \end{bmatrix}$$

the density function of $h(z,t)$

Specifically, I have $dz = \left(z\theta_w(\mu_w - \log z) + \frac{1}{2}z\sigma_w^2\right)dt + \sigma_w z dW = \mu(z)dt + \sigma(z)dW$, then the FP equation is

$$\frac{d}{dt}h(z,t) = -\frac{d}{dz}\left[\mu(z)h(z,t)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2(z)h(z,t)\right]$$

where $\mu(z) = z\theta_w(\mu_w - \log z) + \frac{1}{2}z\sigma_w^2$ and $\sigma^2(z) = \sigma_w^2 z^2$.

Simplify it into a differential equation of h

$$\frac{dh(z,t)}{dt} = a(z,t)h(z,t) + b(z,t)\frac{dh(z,t)}{dz} + c(z,t)\frac{d^2h(z,t)}{dz^2}$$

I have

$$\begin{aligned} a_h(z) &= \theta_w + \frac{1}{2}\sigma_w^2 - \theta_w(\mu_w - \log z) \\ b_h(z) &= \theta_w - \frac{1}{2}\sigma_w^2 - \theta_w(\mu_w - \log z) + 2\sigma_w^2 z \\ c_h(z) &= \frac{1}{2}\sigma^2 z^2 \end{aligned}$$

- Use a forward-difference approximation in the time dimension
- Use a central-difference approximation in the space dimension

- and write the right hand side in an implicit scheme (evaluated at the $n+1$ time level)

I have $\frac{dh(z,t)}{dt} = a(z,t)h(z,t) + b(z,t)\frac{dh(z,t)}{dz} + c(z,t)\frac{d^2h(z,t)}{dz^2}$ as

$$\frac{h_i^{n+1} - h_i^n}{\Delta t} = a_{hi}^n h_i^{n+1} + b_{hi}\frac{h_{i+1}^{n+1} - h_{i-1}^{n+1}}{2\Delta z} + c_{hi}\frac{h_{i+1}^{n+1} - 2h_i^{n+1} + h_{i-1}^{n+1}}{(\Delta z)^2}$$

Then I have a system in $\{h_i^{n+1}, h_i^n\}$:

$$x_{hi}h_{i-1}^{n+1} + y_{hi}h_i^{n+1} + z_{hi}h_{i+1}^{n+1} = \omega_i^n$$

where

$$\begin{aligned} x_{hi} &= b_{hi}\frac{\Delta t}{2\Delta\epsilon} - c_{hi}\frac{\Delta t}{(\Delta\epsilon)^2} \\ y_{hi} &= 1 - a_{hi}\Delta t + c_{hi}\frac{2\Delta t}{(\Delta\epsilon)^2} \\ z_{hi} &= -b_{hi}\frac{\Delta t}{2\Delta\epsilon} - c_{hi}\frac{\Delta t}{(\Delta\epsilon)^2} \end{aligned}$$

This holds for $i = 2, ..., I-1$. I have $I-2$ equations in $h_2^{n+1}, ..., h_I^{n+1}$.

The boundary conditions are

$$\begin{aligned} h_1^{n+1} &= 0 \\ \Delta z\sum_{j=1}^{I} h_j^{n+1} &= 1 \end{aligned}$$

which give 2 additional equations.

For a given $\{h_i^n\}_{i=2}^{I-1}$, I have a system of $I-1$ equations in $I-1$ unknowns $\{h_i^{n+1}\}_{i=2}^{I}$. I can summarize this in a matrix form and solve it quickly in Matlab.

$$\begin{bmatrix} y_2 & z_2 & 0 & 0 & \dots & 0 \\ x_3 & y_3 & z_3 & 0 & \dots & 0 \\ 0 & x_4 & y_4 & z_4 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & x_{I-1} & y_{I-1} & z_{I-1} \\ \Delta z & \Delta z & \Delta z & \Delta z & \Delta z & \Delta z \end{bmatrix} \begin{bmatrix} h_2^{n+1} \\ h_3^{n+1} \\ \vdots \\ \vdots \\ h_{I-1}^{n+1} \\ h_I^{n+1} \end{bmatrix} = \begin{bmatrix} h_2^n \\ h_3^n \\ \vdots \\ \vdots \\ h_{I-1}^n \\ 1 \end{bmatrix}$$

Chapter 3

Intergenerational Income Mobility and Income Inequality in China

The intergenerational income mobility decreases and the Gini coefficient of income increases to high level since China began economic reform in 1978. I propose a theoretical overlapping generation model with missing capital markets, increasing the return to human capital and increasing education cost to explain these facts. After the economic reform happens, all levels of wages go up and all families accumulate and update human capital. However, the increasing education cost and credit constraint prevent the children from rural families from accumulating human capital quickly. The urban families accumulate human capital faster than the rural families. These predictions from the model are verified in the census data. Whether this process continues or not depends on the subsidy of education. Government education policy can improve the allocation of education in the economy.

3.1 Introduction

Over the last 30 years, China has made impressive progress in economic development. The GDP per capita grows from $220.68 in 1980 to $2891.09 in 2010, measured in constant 2005 US dollar (World Bank). On average, it is 8.38% over the last 31 years. The remarkable constant growth combined with the size of the economy has put China the second largest economy in the world since 2009. However, this growth has also accompanied with increase income inequality and decreasing intergenerational income mobility.

Both facts pose concerns about the quality of China's economic success and its future. The income inequality itself indicate different welfare between different groups. The benefits of economic growth enter more into the pockets of the high-income family. What's more, when the economic mobility decreases, the poor income family in the economy will have a higher probability to stay poor, while the rich income family stays where they are. The economy is heading to a less mobile economy, which can mean a potential waste of human capital and political instability.

This paper tries to answer what caused this pattern in China. I build a simple overlapping generation model with education choice and missing capital market to explain why this pattern happens in China. I propose two elements that are keys to understanding this pattern and supported by the data. The first one is the increasing return to education and the second one is the increasing education cost.

The increasing return to education is supported by empirical works and can be modeled as skill-biased technology. It is triggered by the economic reform in China and become the source for increasing income inequality.

Even though the increasing return to the human capital increase the incentive to accumulate the human capital for children from both poor and rich families, the increasing education cost, combined with the missing capital market, prevents the poor families from taking more high education than the rich families. It may be too expensive or not affordable for the children from the poor families to take higher education. Therefore, the intergenerational income mobility will decrease in the long run.

The education system is taken as exogenous in this paper. The increasing in the education cost can be thought as a response to an increasingly high demand for education and a slow supply response controlled by the government. The results from this paper open the door to discuss the government policy in response to the inequality and mobility in the future.

Consist with the logic proposed in the paper, I investigate the census data for China by looking at the human capital patterns over time and regions. The data show that all families accumulate and update human capital for the last 30 years. However, when divided into groups, the urban families, relative rich ones, take higher education than the rural families, relative poor ones, as a whole.

This paper is the first one to explain the dynamic facts of increasing inequality and decreasing intergenerational mobility in China.

3.2 Stylized Fact for China

The first set of empirical facts consists of the income inequality and intergenerational income mobility.

The income inequality in China is increasing over the last 30 years. The income inequality is usually measured by the Gini coefficient of income. A study by Ravallion and Chen [2007] calculate the numbers are 0.309 in 1981 and 0.447 in 2001. The National Bureau of Statistics of China (NBS) also announces the Gini coefficients are 0.454 in 2001 and 0.473 in 2013. Although it is high relative to other countries, the NBS numbers are relative constant over years. This creates a lot doubts on the accuracy of these numbers. For example, another study by Xie and Zhou [2014] claims the income inequality is from 0.53-0.55 in 2013, much higher than the official numbers. This high level of income inequality is also echoed by Li [2013], who get the Gini coefficient of 0.61 in 2010 from the China Household Finance Survey (CHFS).

The intergenerational income mobility is decreasing over the same period. One way to measure the income mobility is to use the intergenerational elasticiy. This measure shows how closely the children's income depends on their parents' income. Controled various other elements, Deng et al. [2013] shows that the intergenerational elasticiy of son/father pair increases from 0.47 in 1995 to 0.53 in 2002 in the urban area. Fan et al. [2015] claims the intergenerational income elasticiy increases from 0.315 to .442 between cohorts born before and after 1970. If we use the rank-rank elasticity estimate, which is more reliable compared to the income elasticity estimate, Fan et al. [2015] shows it increases from 0.273 to 0.347 for these two cohorts.

Another way to measure the intergenerational income mobility is to use the transition matrix. Table 3.1 shows the parent/child quintile transition matrix for the two cohorts studied in Fan et al. [2015]. The population is divided into 5 groups equally from low to high according to their income. The (i, j) element in the matrix shows the probability that the children move to group j when the parents' income is in group i. A comparison between two cohorts income transition matrix indicates that it is harder for the children from the low-income groups (1 and 2) to move up while the children from the high-income groups (3, 4 and 5) have a higher chance to stay in the high-

Table 3.1: Income Transition Matrix

Panel A. Early cohort 1956-1970, Income						
		Children Quintile				
		1	2	3	4	5
	1	0.392	0.1935	0.2126	0.168	0.0317
Parent	2	0.24	0.3145	0.2441	0.128	0.0714
Quintile	3	0.184	0.25	0.1969	0.256	0.1111
	4	0.152	0.1694	0.1811	0.232	0.2698
	5	0.032	0.0726	0.1654	0.216	0.5159

Panel B. Late cohort 1971-1985, Income						
		Children Quintile				
		1	2	3	4	5
	1	0.4512	0.2303	0.1697	0.1091	0.0364
Parent	2	0.3171	0.303	0.2424	0.0727	0.0667
Quintile	3	0.1524	0.2545	0.2485	0.2303	0.1152
	4	0.0671	0.1636	0.1697	0.2727	0.3273
	5	0.0122	0.0485	0.1697	0.3152	0.4545

income groups. For example, the probability of staying in the lowest income group when come from this group increases from 0.39 to 0.45. The probability to stay in at least group 4 for the children come from group 4 increases from 0.5 to 0.63.

The second set of facts relates to the two assumptions used in this paper. As clearly documented in Li et al. [2012], the Mincer-style rate of return to education urban China increases from 2.3 percent in 1988 to 9 percent in 2000 and 9.5 percent in 2009. At the same period, the premia of college education relative to high school education increased from 7.4 percent in 1988 to 49.2 percent in 2009, which is a dramatic increase. Both indicates the increasing return to the education. As for the education cost, the ratio of all the tuition over government education expenditure increased from 5% in 1991 to around 35% in 2007 (NBS 2013) and the tuition over GDP increased from 0.1% in 1991 to around 0.8% in 2007. These tuitions are the costs for formal schools and they are only a part of the total schooling cost a family faced. For example, students from the rural area may need to pay for more living expensive and housing when entering a high school in the city. At the same time, the children from urban families can choose additional education investment by taking extra tutoring.

The third set of facts involves the human capital patterns over last 30 years from the census data of China. I first check the education distribution of the population at 1982, 1990, 2000 and 2010 (See Figure 3.1). The education is grouped into 5 levels: no schooling, primary, secondary, high school and the college degree and above. A trend of more human capital is shown up. In 1982, most population are concentrated in low education, like no schooling or primary schooling, which are around 40% and 35% respectively. There is less than 10% of population has degrees over high school. Over years, the population gradually accumulate more human capital. The distribution shifts its weight to the high levels of education. In 2010, the majority of the population around 43% have secondary schooling, followed by primary schooling and high schooling. Around 10% of the population has the college degree and above in 2010, while this number is less than 1% in 1982.

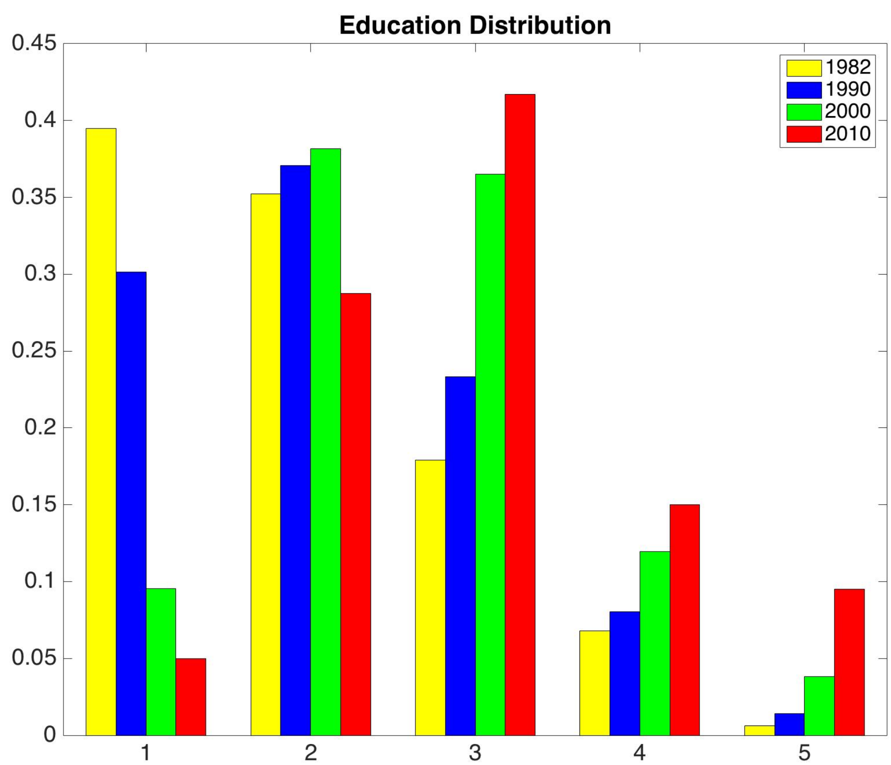


Figure 3.1: Education Distribution over Years

When the data are divided into regions, another pattern shows up: The urban population the accumulate human capital much faster than the rural population. Given the availability of region data, I compare the data from 2000 and 2010 (see Figure 3.2). At the aggregate level, the trend is same as before: the distribution of education is moving towards higher education levels. However, when to compare the rural group with the city group, two facts stand out. First, the rural group concentrated in the lower education levels of primary and secondary schooling, and the city group centered around the secondary schooling with some tail in the higher education levels. Second, ten years later, the rural population still has a lot of sharing in primary and secondary schooling, only a very few share in the high schooling and above. At the same time, the city population quickly update the human capital and around 20% of them get the college education and above.

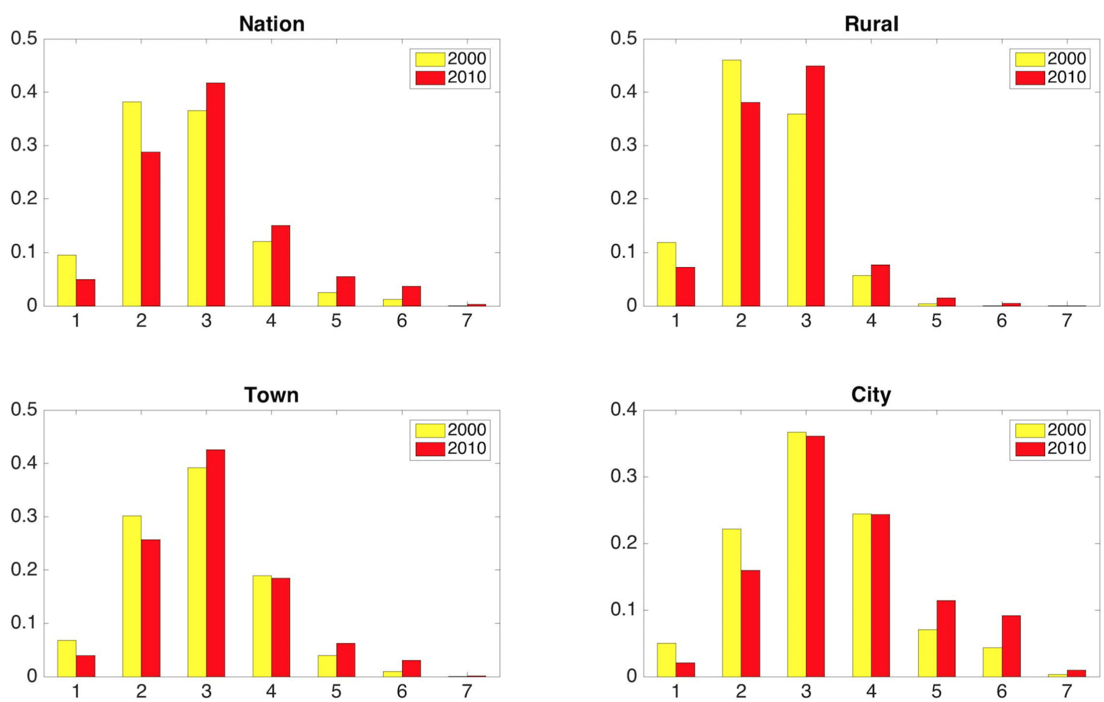


Figure 3.2: Educaiton Distribution over Regions

3.3 Related Literature

This paper has a close link with 3 strands of literature.

The first is the theoretic papers on the intergenerational mobility. The literature starts with **?** and Loury [1981], where human capital is introduced to explain the pattern. Later work of Becker and Tomes [1986], Solon [2004] and Becker et al. [2015] add credit constraint and another persistent element in the theory model to explain the patterns of intergenerational mobility. For more works, good reference are Piketty [2000] and Black and Devereux [2011]

The second one is the empirical paper working on the intergenerational mobility and inequality in China. These research include Gong et al. [2012], Deng et al. [2013], **?** and Fan et al. [2015]. They all try to estimate the intergenerational mobility and inequality in China over time. The estimation method and data source are different but the results all point to a decreasing mobility.

The last one is related to the effects of credit constraint and income distribution on the development dynamics. This line of research concerns how the incompleteness of credit market or capital market affects the dynamics of development. The income distribution matters as it affects the choices of individual given the credit constraint, so it has the huge effect on the transitional dynamics and a permanent effect on the economy. The related literature include Galor and Zeira [1993], Aghion and Bolton [1992], Banerjee and Newman [1993], Aghion and Bolton [1997] and Maoz and Moav [1999].

3.4 Model

3.4.1 Environment

The model is a two-period overlapping generation model with missing capital market and education choices.

The population is constant over time and normalized to one. Each family is of measure zero. Each family consists one adult and one child. An agent has two periods: young and old.

Agent i, who is born at time t, first draws the innate ability (cognitive ability) θ^i_{cog} from a normal distribution $\theta^i_{cog} \sim N(\mu, \sigma)$. The cognitive ability distribution is common for new born child.

Agent i inherits two things from his parent. One is the bequest x^i_{1t}, and the other is non-cognitive ability θ^i_{ncog}. The bequest x^i_{1t} is in monetary term. The non-cognitive ability is determined by the parents' education level. It is assumed that parents with high education level will provide a high level of non-cognitive ability. This can be seen as a shortcut for modeling the genetic heritage and difference in the family early education.

The schooling utility a^i for agent i combines both effects from cognitive ability and non-cognitive ability in an additive way, $a^i = \theta_{cog} + \theta_{ncog}$. Given the distribution assumption of θ_{cog}, we have $a^i \sim N(\mu + \theta_{ncog}, \sigma)$.

Each agent maximizes his lifetime utility. The utility comes from consumption streams, schooling and the utility of his child. The utility from his child is modeled as a warm glow, which is standard in the literature.

Assume the utility takes the following form, consisting all the elements mentioned above, $U(C^i_{1,t}, C^i_{2,t}, s^i_{1,t}, x^i_{2,t}; a^i, x^i_{1,t}) = \log C^i_{1,t} + a^i T^i + \beta \log C^i_{2,t} + \alpha\beta \log x^i_{2,t}$. The state variables for agent i is the ability a^i and bequest $x^i_{1,t}$. The consumption choices are $C^i_{1,t}$ and $C^2_{2,t}$, and the bequest choice $x^i_{2,t}$. The educaiton choice $s^i_{1,t}$ denotes what level of education agent i is choosing. More details on the educaiton choice will be followed.

The schooling utility is modeled as $a^i T^i$, where T^i is the schooling year for the agent i. Thus, when a^i is negative, a larger T^i denotes a larger disutility for the agent i.

There is only one final good and it is the e in this economy. The production is a constant return to scale function with constant elasticity of substitution among five types of labor inputs:

$$Y_t = A_t \left[\sum_{j=1}^{5} z_{j,t} L^{\rho}_{j,t} \right]^{\frac{1}{\rho}}$$

The elasticity of substitution between different types of labors are $\frac{1}{1-\rho}$. The parameters A_t and $z_{j,t}$'s are exogenous productivity terms.

The good market and labor market are both competitive. The final good producers take the wages as given and choose the labor inputs to maximize the profit:

$$\max_{\{L_{j,t}\}_{j=1}^{5}} \left\{ Y_t - \sum_{j=1}^{5} w_{j,t} L_{j,t} \right\}$$

The first order condition gives the wage equation $\frac{\partial Y_t}{\partial L_{j,t}} = w_{j,t}$ and for each level of education $j \in \{1, ..., 5\}$,

$$w_{j,t} = A_t \left[\sum_{j=1}^{6} z_{j,t} L_{j,t}^{\rho} \right]^{\frac{1}{\rho}-1} z_{j,t} L_{j,t}^{\rho-1}$$

The relative wage is determined by the supply of labor force and the ratio of specific technology $z_{j,t}$'s:

$$\frac{w_{j,t}}{w_{k,t}} = \frac{z_{j,t} L_{j,t}^{\rho-1}}{z_{k,t} L_{k,t}^{\rho-1}}$$

3.4.2 Individual Problem

When agent i is young, she chooses how to divide his bequest $x_{1,t}^i$ into consumption C_{1t}^i and education investment ξ_t^i. The education investment involves choices among 5 levels of education, $s^i \in \{1, ..., 5\}$.

When making consumption and education investment, the financial market is shut down. The only resource is the bequest from parent. This assumption is strong but it captures the difficulty to borrow against the future income of children. It provides a lower bound in the model. The model result can be seen as a worst case scenario. Also, this assumption justifies the government education policy. But at the same time the effect from government policy will be at the upper bound given this assumption.

When agent i is old, she becomes a parent. Her education level and the income y_{t+1}^i realize, and she chooses how much to consume $C_{2,t}^i$ and to leave as bequest for his child $x_{2,t}^i$.

The problem of agent i can be summariezd as following:

$$\max_{s^i, e^i, C_{1,t}^i, C_{2,t}^i, x_{2,t}^i} \mathbb{E} \left\{ \log C_{1,t}^i + a^i T^i + \beta \log C_{2,t}^i + \alpha\beta \log x_{2,t}^i \right\}$$

subjected to

$$\begin{aligned} C^i_{1,t} + \xi^i_t &= x^i_{1,t} \\ C^i_{2,t} + x^i_{2,t} &= y^i_{t+1} \\ C^i_{1,t}, C^i_{2,t}, \xi^i_t, x^i_{2,t} &\geq 0 \end{aligned}$$

Education Choice

The education system consists 5 levels of education, denoted as s_j, $j = 1, ..., 5$. The education levels represent from low to high: no schooling, primary schooling, junior schooling, high schooling and college education and above.

Each level of education consists an entry cost $\eta_{j,t} = \gamma_j + \kappa_j w_{j,t}$, where $\gamma_j, \kappa_j \geq 0$. The entry cost is assumed proportional to the current wages of the same education level. This assumption is flexible to capture the cost of education at different levels. As the skill-biased technology pushes up the return to high education, the demand for these will rise. Compared with the slow adjustment in the supply of these education levels, as a result, the costs of obtaining these education level are increasing.

For the education choice $s_i \in \{1, ..., 5\}$, the education outcome $\hat{s}^i$ will be the same as the choice s_i. This is different from the choice of the college education. Even agent i is targeting the college education, the outcome is uncertain. There is a probability $p(e)$ that agent will success obtaining the college education. This probability also depends on the additional education input e. It is assumed that the more input is, the larger the probability $p(e)$ is. And the increase in probability is decreasing over e. In terms of the derivatives, it is assumed that $p(e) \geq 0$, $p'(e) > 0$ and $p''(e) < 0$. No matter what the result is, the additional education cost is paid.

For the functional form, the probability is assumed as

$$p_5(e) = \min\{b_0 + b_5 \log(1 + e), 1\}$$

where b_0 denotes the basic probability to enter college and above if agent i choose to take college education and above. The coefficent b_5 controls the elasticity of additional education input.

Thus, the realization of education $\hat{s}^i$ has two outcomes when $s^i = 5$:

$$\hat{s}^i = \begin{cases} s_5 & with\ p_5(e) \\ s_4 & with\ 1 - p_5(e) \end{cases}$$

3.4.3 Individual Optimal Choices

The individual problem is solved backward. I first characterize the optimal choice in the old period, given the realization of education level and income, then I characterize the choices in the first period.

Given the education choice s^i for the first period, after the realization of $\hat{s} = j$, the income for agent i at the second period is $y^i_{j,t+1}$. The second period choices are consumption $C^i_{2,t}$ and beqeust $x^i_{2,t}$. The solution is derived by the first order condition. The bequest is $x^i_{2,t} = \frac{\alpha}{1+\alpha} y^i_{j,t+1}$.

We can use in direct utiltiy function to summarize the optimal choices in the second period. The indirect utility function $W(y^i_{j,t+1})$ for the second period is

$$W(y^i_{j,t+1}) = (1 + \alpha) log \left[\frac{(1 - \tau_{t+1}) y^i_{j,t+1}}{2} \right] + m$$

where $m = \log\left(\frac{1}{1+\alpha}\right) + 2\log\left(\frac{\alpha}{1+\alpha}\right)$.

Now turn to the first period, agent i will choose the best education choice and consumption given her state $(a^i, x^i_{1,t})$. We can also use the value function to denote the expected utility from different education choices.

For the education choice $s^i \in \{1, 2, 3, 4\}$, the income for the second period is for sure. The first question for agent i is whether the education is affordable or not, that is, $x^i_{1,t} \geq \gamma_j + \kappa_j w_{j,t}$. The bequest from parent determines the education level can be chosen for the child. This is the credit constraint emphasized in the literature.

When the education is affordable, agent i compare the expected utility from choosing the education level j. The expected utility is

$$V^j(x^i_{1,t}, a^i_t) = \log\left(x^i_{1,t} - \eta_{j,t}\right) + a^i T_j + \beta W(w_{j,t+1})$$

The value functions for $j \in \{1, .., 4\}$ shows that the value function is linear in a^i and the wage $w_{j,t+1}$ work as a shifter for the value function. Higher wage in the future provides more incentive to invest in that education level.

For a given bequest $x^i_{1,t}$, any comparison between V^j's, $j \leq 4$, leads to a cutoff for the ability a_{jk}, such that for any $a \geq a_{jk}$, $V^j \geq V^k$:

$$a_{jk} = -\frac{1}{T_j - T_k}\left\{\log\left(\frac{x^i_{1,t} - \gamma_j - \kappa_j w_{j,t}}{x^i_{1,t} - \gamma_k - \kappa_k w_{k,t}}\right) + (1+\alpha)\beta log\left[\frac{w_{j,t+1}}{w_{k,t+1}}\right]\right\}$$

This equation shows different incentives in chosing education j over k. Suppose $T_j > T_k$ and $a_{jk} < 0$. Then the ability disutility comes from education j is larger than education k. The larger is $T_j - T_k$, the larger is a_{jk} and less probability to choose education j. This is the ability incentive. At the same time, when the wage gap $\frac{w_{j,t+1}}{w_{k,t+1}}$ is larger, the smaller is a_{jk} and larger the probability to chooose education j.

For the education choice $s^i = 5$, the income for the second period is random. Agent i need to choose the optimal additional education input e.

$$\begin{aligned} V^j(x^i_{1,t}, a^i_t) &= \max_e \left\{p_j(e)\left[\log\left(x^i_{1,t} - \eta_{j,t} - e^i\right) + a^i T_j + \beta W(y^i_{j,t+1})\right]\right. \\ &\quad \left. + \left(1 - p_j(e^i)\right)\left[\log\left(x^i_{1,t} - \eta_{4,t} - e^i\right) + a^i T_4 + \beta W(y^i_{4,t+1})\right]\right\} \end{aligned}$$

The first order condition provides the optimal additional education input:

$$\begin{aligned} & p'_j(e)\left[\log\left(\frac{x^i_{1,t} - \eta_{j,t} - e^i}{x^i_{1,t} - \eta_{4,t} - e^i}\right) + a^i(T_j - T_4) + W(y_{j,t+1}) - W(y_{4,t+1})\right] \\ = & \quad \frac{p_j(e)}{x^i_{1,t} - \eta_{j,t} - e^i} + \frac{1 - p_j(e^i)}{x^i_{1,t} - \eta_{4,t} - e^i} \end{aligned}$$

The above equation describes the marginal gain from additional education investment equals the marginal loss. It defines the policy function $e^*_j(x^i_{1,t}, a^i)$. Given the choice of additional education investment, then we can get the value function of $V^5(x^i_{1,t}, a^i_t)$.

3.4.4 The Equilibrium

The equilibrium is a standard competitive equilibrium. Given the prices $\{w_{j,t}\}_{t=0}^{\infty}$, each agent optimally chooses the consumption, education level and bequest. At the same time, the labor markets and the good markets need are at equilibrium. The supplies of labor with different education levels are equal to the demands of each education level.

Let $m_{kj,t}$ denote the share of children from family education level k moving to the education j. The evolution for the labor supplies are

$$\sum_{k=1}^{5} m_{kj,t} L_{k,t} = L_{j,t+1}$$

for $j = 1, 2, 3, 4, 5$ and any $t \geq 0$. The labor markets also requires the aggregate labor supply is constant over time: $\sum_{j=1}^{5} L_{j,t} = 1$.

3.5 Numerical Example

The numerical example is designed to illustrate the model's power in generating the empirical patterns we observe for China. Given the data available at this time, the numerical example is not a calibration exercise. A careful quantitative exercise is left for future research.

The idea of numerical example is to study two stationary equilibria of the model. One represents the economy of China in 1982 and the other in 2010. The differences between these two stationary equilibria are lying in the technology levels of z_j's and the education cost parameters γ_j's. The differences in z_j's represent the skill-biased technology in the model assumption. The differences in the education cost parameters γ_j's represent the change in the education system between these two years.

3.5.1 Parameters

I first choose some parameters from literature, and these parameters are kept constant over two years. Then I choose values for the left ones to match the labor distribution in 1982 and in 2010 respectively.

Parameters Constant between Years

The discounted rate β are chosen by $\left(\frac{1}{1.04}\right)^{30} = 0.555$. The preference coefficient α are chosen such that agents leave 40% of income in the second period to their children. That is, $\frac{\alpha}{1+\alpha} = 0.4$, then $\alpha = 2/3$.

The innate ability distribution is assumed as $\theta_{cog} \sim N(-0.05, 1)$. And the non-cognitive abilities are assigned numbers of (-0.1375, -0.0750, -0.0375, -0.0125, -0.0075, 0.0025) for the educations from low to high. These numbers reflect the advantages and disadvantages of being born into different education level families. They can be interpreted as parents' genetic impact or early education impact.

The schooling years are chosen as following, $T_1 = 3$, $T_2 = 6$, $T_3 = 9$, $T_4 = 12$ and $T_5 = 16$ to represent education levels of no schooling, primary schooling, secondary schooling, high schooling and college degree and above. The education cost proportional coefficients are $\kappa_1 = \kappa_2 = \kappa_3 = 0$,$\kappa_4 = 0.16$ and $\kappa_5 = 0.06$.

For the production function, I choose $\rho = 1 - 1/1.64$ such that the elasticity of substitution is 1.64.

Parameters Changed between Years

The Table 3.2 show the different values used in matching the labor distribution in 1982 and 2010.

The numerical values for z_j's display the skill-biased technology assumption. Technology improvements are uneven among different levels of education. It decreases for education level one, and it increases for other education levels. The increase in A_0 also increases all the types of wages in the economy.

Table 3.2: Parameters Changed Between Years

	z_1	z_2	z_3	z_4	z_5	A_0
Equilibrium 1982	0.2	0.74	0.7	0.36	0.18	1
Equilibrium 2010	0.02	2.7	13.75	9.075	7.5	1.5
	γ_1	γ_2	γ_3	γ_4	γ_5	b_5
Equilibrium 1982	0	0	0.2	0.5	-0.5	0.3
Equilibrium 2010	0	0	0.2	100	50	0.11

The changes in the fixed part of education cost γ_j's capture part of the uneven increase in the education cost. The other part is captured in the increasing wages. Notably, the increase in the fixed education cost in high education and college education and above are huge. These reflect the high demand for these education levels.

The last change in the elasticity of additional education input is a decrease in b_5 from 0.3 to 0.11. This means the marginal probability of entering college and above from additional education input is lower. This captures the competition of entering college is increasing over time.

3.5.2 Numerical Results

Given these values for the model, I could match the labor distribution in 1982 and 2010 quite well. As we can see in the Figure 3.3, the model matches the most of the levels, except education level 4, which is the high schooling. The patterns from these two years are consistent with the theory. Initially, most population are concentrated in the low education levels. The distribution is skew to the left in 1982. In 2010, the distribution moves its center to the middle. The share of secondary schooling is the largest and there is a significant amount of agents obtain the high education.

After the aggregate labor distribution is matched, we now turn to the income transition matrix. The results are shown in the Figure 3.4.

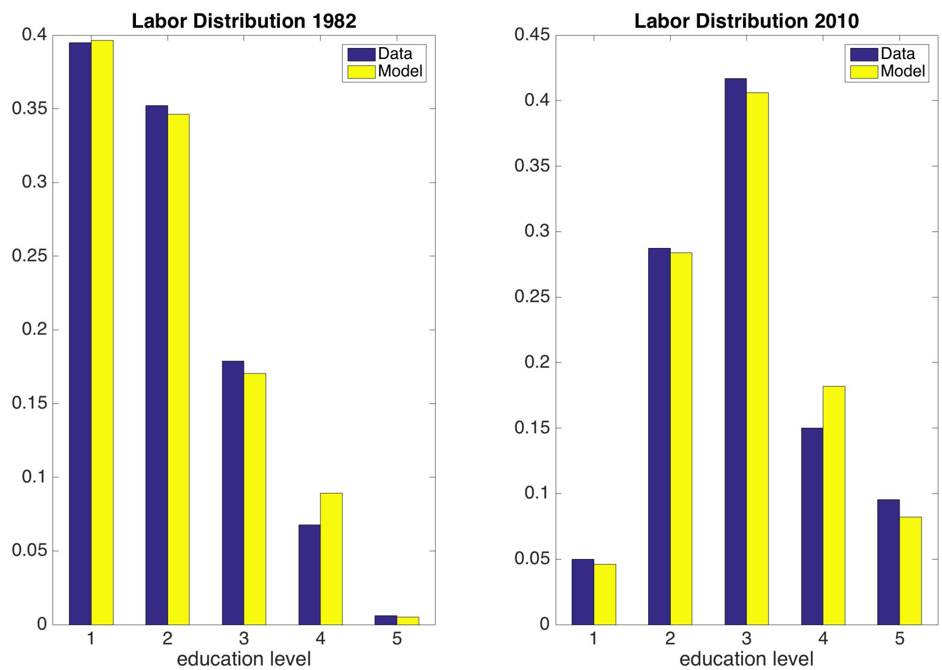


Figure 3.3: Labor Distribution: Model vs Data

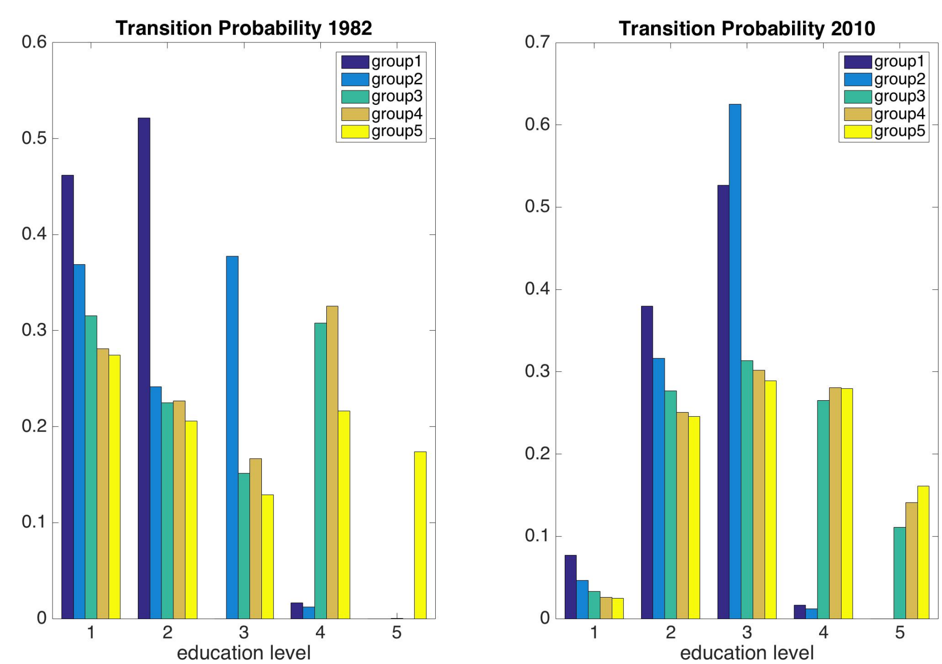


Figure 3.4: Model Generated Transition Matrix

The five groups indicate the parent education level. Each bar shows the probability of the child to be in that education group. In this model, education group corresponds to income group. It is also a two-period overlapping generation model, the income transition matrix corresponds to the transition of parent/child income group. So the transition probability is the intergenerational income transition probability.

The first result is the distributions of probability for group 1 and group 2 between the two years. These two groups are relatively poor families in the model and can be thought as the rural families in the data. For them, the probabilities of staying in education level 1, 2 and 3 are very high in 1982. It is also true that their probabilities of entering education level 3 increase a lot in 2010. At the same time, in both years of 1982 and 2010, the children of group 1 and group 2 has little probabilities of entering education level above 4. This is exactly what the theory predicts: the relatively poor cannot afford or will not choose the high education, even though their wages are increasing over time. These facts are consistent with the data we observe for China, that is the second graph in Figure 3.4.

The second result is the distributions of probability for group 3, group 4 and group5 between the two years. These three groups are relatively rich families in the model, and they can be thought as families in the urban area, corresponding county, and city in the data. In 1982, children from these groups have most probabilities distributed among education levels from 1 to 4 with a U-shape. In 2010, the probability distribution changes into an inverse U-shape, showing more probabilities in the education levels from 2 to 4, and at the same time, there is a significant chance, from 10% to 18%, entering the highest education level. These facts are also consistent with the data we observe for China, that is the fourth graph in Figure 3.4.

Combined these two results together, the reasons for the decreasing in mobility are clear. Even though all families accumulate more human capital, the quantity and quality are different. Children from the rural families are relatively poor. They will not choose or can not afford the high education because of the increasing education cost. Children from urban families are much lucky, their parents' income can afford the increasing education cost, as a result, they become the majority source for the higher education.

The results indicate the misallocation and the waste of human capital. Some children from relatively poor families are smart and they do not obtain the higher education. Some children from relatively rich families obtain the high education because they can invest more. However, if their credit constraints are away, the entry of children from poor families will make the competition for high education entry more severe. The probability of obtaining the high education is not affected by the demand in this model, but it can be modeled as an increase in the fixed part of the education cost. And the increased supply of high education agents will depress the wage rate a little. Therefore, increasing cost and decreasing wage benefit will prevent some relative low ability children from the rich families from obtaining the high education. This is an improvement in the welfare. In theory, it is also good to compare current results with a planner solution.

3.6 Conclusion

This paper is the first one to investigate why China has decreasing intergenerational income mobility and increasing income inequality in the last 30 years along with the rapid growth. It proposes that increasing return to human capital and increasing education cost are the reason behind these facts. The theory is tested in a numerical experiment based on an overlapping generation model with missing capital markets. The results are confirmed that these two elements can generate the patterns we observe for China. Also, the results are consistent with the education distribution both in national level and in region level observed in the last 30 years.

This paper is the first step to thinking deeply about the policy response. In the future work, a full quantitative overlapping generations model is required to evaluate any education policy. As pointed in the paper, there is a misallocation in the current, caused by the credit constraint. However, different from Restuccia and Urrutia [2004] and Yang and Qiu [2016] who claim the subsidy for early education is more important, subsidy of education may have more weight on the high education in the face of skill-biased technology. An expansion of college loan may increase the efficiency more than a subsidy for early education for current China. And these subsidies are important to predict whether the patterns of mobility and inequality will continue or not. All these questions are left for future research.

3.7 Appendix: Numerical Solution Method

The state variables are $L_{j,t}$ and the prices are $w_{j,t}$, $j \in \{1, ..., 5\}$.

Given the current wage prices, we know the bequest for children from family with occupation j at time t is

$$x_{2,t-1}^{ij} = x_{1,t}^{ij} = \frac{\alpha}{1+\alpha} w_{j,t}$$

. As the bequest is linked with the parents income and the bequest is same for the children who's parents are in the same occupation, I use $x_{1,t}^{ij}$ to denote the bequest for children from the occupation j family. The heratige distribution is $\left\{\frac{\alpha}{1+\alpha} w_{j,t}\right\}_{j=1}^{5}$.

Given the bequest distribution, the education costs, and an expected wage next period, we can find out the best occupation for different abilities. Then we can solve out the labor forces next period. The key is to find out the state evolution equation.

At the steady state, the wages and the labor forces are constant over time. This simplifies the dynamic system into a static one. We can drop out the index t for the dynamic system.

1. Guess an initial distribution of labors L_j, $j \in \{1, ..., 5\}$.

2. solve the value function $V^j(x_1^k, a)$ for $k \times j = 6 \times 6$.

   (a) Given the state L_j's, we can calculate the wages w_j, $j \in \{1, ..., 5\}$.

   (b) given the wages, we know the bequests x_1^k's.

   (c) then calculate value function for each group k: $V^j(x_1^k, a)$ for $j \in \{1, ..., 4\}$.

   (d) use grid approximation on (a, e) for V^5 and find out the optimal additional education input e, and the value function $V^5(x_1^k, a)$

   (e) calculate the transitional probability and labor supply

3. Calculate the labor supply $L_j^{'}$'s, update the state variables L_j's; The update parameter is 0.9.

4. when the state variables are close to each other $\sqrt{\sum_{j=1}^{5}(L_j - L_j^{'})^2} < \epsilon$, stop updating.

Bibliography

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2015). Heterogeneous Agent Models in Continuous Time. *Working Paper*.

Aghion, P. and Bolton, P. (1992). Distribution and Growth in Models of Imperfect Capital Markets. *European Economic Review, Papers and Proceedings of the 6th Annual Congress of the European Economic Association*, 36(2/3):603–611.

Aghion, P. and Bolton, P. (1997). A Theory of Trickle-Down Growth and Development. *lead article in the Review of Economic Studies*, 64(2):151–172.

Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics*, 109(3):659–684.

Bai, C.-E., Hsieh, C.-T., and Qian, Y. (2006). The Return to Capital in China. *Brookings Papers on Economic Activity*, 37(2):61–102.

Banerjee, A. V. and Duflo, E. (2005). Growth Theory through the Lens of Development Economics. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, chapter 7, pages 473–552. Elsevier.

Banerjee, A. V. and Newman, A. F. (1993). Occupational Choice and the Process of Development. *Journal of Political Economy*, 101(2):274–98.

Barles, G. and Souganidis, P. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis*.

Beck, T., Demirguc-Kunt, A., and Levine, R. (2009). Financial institutions and markets across countries and over time - data and analysis. Policy Research Working Paper Series 4943, The World Bank.

Becker, G., Kominers, S. D., Murphy, K. M., and Spenkuch, J. L. (2015). A Theory of Intergenerational Mobility. MPRA Paper 66334, University Library of Munich, Germany.

Becker, G. S. and Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4(3):S1–39.

Bewley, T. (1986). *Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers*. Amsterdam:North Holland.

Black, S. E. and Devereux, P. J. (2011). *Recent Developments in Intergenerational Mobility*, volume 4 of *Handbook of Labor Economics*, chapter 16, pages 1487–1541. Elsevier.

Buera, F. J. and Shin, Y. (2013). Financial Frictions and the Persistence of History: A Quantitative Exploration. *Journal of Political Economy*, 121(2):221 – 272.

Cagetti, M. and Nardi, M. D. (2006). Entrepreneurship, Frictions, and Wealth. *Journal of Political Economy*, 114(5):835–870.

Chang, C., Chen, K., Waggoner, D. F., and Zha, T. (2015). Trends and Cycles in China's Macroeconomy. In *NBER Macroeconomics Annual 2015, Volume 30*, NBER Chapters. National Bureau of Economic Research, Inc.

Cihak, M., Demirguc-Kunt, A., Feyen, E., and Levine, R. (2012). Benchmarking Financial Systems Around the World. Policy Research Working Paper Series 6175, The World Bank.

Deng, Q., Gustafsson, B., and Li, S. (2013). Intergenerational Income Persistence in Urban China. *Review of Income and Wealth*, 59(3):416–436.

Dollar, D. and Jones, B. F. (2013). China: An Institutional View of an Unusual Macroeconomy. NBER Working Papers 19662, National Bureau of Economic Research, Inc.

Fan, Y., Yi, J., and Zhang, J. (2015). The Great Gatsby Curve in China: Cross-Sectional Inequality and Intergenerational Mobility.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, 105(10):3150–82.

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

Gilchrist, S., Sim, J. W., and Zakrajsek, E. (2013). Misallocation and Financial Market Frictions: Some Direct Evidence from the Dispersion in Borrowing Costs. *Review of Economic Dynamics*, 16(1):159–176.

Gong, H., Leigh, A., and Meng, X. (2012). Intergenerational income mobility in urban china. *Review of Income and Wealth*, 58(3):481–503.

Harris, J. R. and Todaro, M. P. (1970). Migration, Unemployment & Development: A Two-Sector Analysis. *American Economic Review*, 60(1):126–42.

Hayashi, F. and Prescott, E. C. (2008). The Depressing Effect of Agricultural Institutions on the Prewar Japanese Economy. *Journal of Political Economy*, 116(4):573–632.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2013). The Allocation of Talent and U.S. Economic Growth. NBER Working Papers 18693, National Bureau of Economic Research, Inc.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Huggett, M. (1993). The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. *Journal of Economic Dynamics and Control*, 17(5-6):953–969.

Imrohoroglu, S., Imrohoroglu, A., and Chen, K. (2006). The Japanese Saving Rate. *American Economic Review*, 96(5):1850–1858.

Jovanovic, B. (2014). Misallocation and Growth. *American Economic Review*, 104(4):1149–71.

King, R. G. and Levine, R. (1993). Finance and Growth: Schumpeter Might Be Right. *The Quarterly Journal of Economics*, 108(3):717–37.

Levine, R. (2005). Finance and Growth: Theory and Evidence. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, chapter 12, pages 865–934. Elsevier.

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, 22(2):139–191.

Li, G. (2013). Income Inequality and Consumption in China.

Li, H., Li, L., Wu, B., and Xiong, Y. (2012). The End of Cheap Chinese Labor. *Journal of Economic Perspectives*, 26(4):57–74.

Loury, G. C. (1981). Intergenerational Transfers and the Distribution of Earnings. *Econometrica*, 49(4):843–67.

Lucas, R. E. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics*, 9(2):508–523.

Lucas, R. E. (2004). Life Earnings and Rural-Urban Migration. *Journal of Political Economy*, 112(S1):S29–S59.

Maoz, Y. D. and Moav, O. (1999). Intergenerational Mobility and the Process of Development. *Economic Journal*, 109(458):677–97.

Meng, X. (2012). Labor Market Outcomes and Reforms in China. *Journal of Economic Perspectives*, 26(4):75–102.

Midrigan, V. and Xu, D. Y. (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review*, 104(2):422–58.

Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *American Economic Review*, 104(10):3186–3221.

Piketty, T. (2000). Theories of persistent inequality and intergenerational mobility. In Atkinson, A. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 1 of *Handbook of Income Distribution*, chapter 8, pages 429–476. Elsevier.

Rajan, R. G. and Zingales, L. (1998). Financial Dependence and Growth. *American Economic Review*, 88(3):559–86.

Ravallion, M. and Chen, S. (2007). China's (uneven) Progress Against Poverty. *Journal of Development Economics*, 82(1):1–42.

Restuccia, D. and Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics*, 11(4):707–720.

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

Solon, G. (2004). *A Model of Intergenerational Mobility Variation over Time and Place*, chapter 2, pages 38–47. Cambridge University Press: Cambridge.

Song, Z., Storesletten, K., and Zilibotti, F. (2011). Growing Like China. *American Economic Review*, 101(1):196–233.

Stokey, N. L. (2008). *The Economics of Inaction: Stochastic Control Models with Fixed Costs*. Princeton University Press.

Todaro, M. P. (1969). A Model for Labor Migration and Urban Unemployment in Less Developed Countries. *American Economic Review*, 59(1):138–48.

Wolff, E. N. (2012). The Asset Price Meltdown and the Wealth of the Middle Class. NBER Working Papers 18559, National Bureau of Economic Research, Inc.

Xie, Y. and Zhou, X. (2014). Income Inequality in Today's China. *Proceedings of the National Academy of Sciences of the United States of America*, 111:6928–2933.

Yang, J. and Qiu, M. (2016). The Impact of Education on Income Inequality and Intergenerational Mobility. *China Economic Review*, 37(C):110–125.

Zhu, X. (2012). Understanding China's Growth: Past, Present, and Future. *Journal of Economic Perspectives*, 26(4):103–24.

Curriculum Vitae

Education

Ph.D., Economics, Boston University, Boston MA, May 2016 (expected)
Dissertation Title: *Essays on Financial Frictions, Misallocation and Development Dynamics*
Dissertation Committee: Simon Gilchrist, Jianjun Miao, Robert G. King

M.A., Economics, New York University, New York NY, 2010
B.A. (*Honours*), Economics and Mathematics, Renmin University of China, Beijing, China, 2008

Fields of Interest

Macroeconomics, Computational Economics, Development Economics

Fellowships and Awards

Teaching Fellowship, Boston University, 2011-2015
Outstanding Graduate of Renmin University of China, 2008
Excellent Student Award, Renmin University of China, 2004-2008
The Credit Scholarship of Hong Kong, Renmin University of China, 2007

Working Papers

"The Persistence of Development Dynamics: Financial Frictions and Mobility Distortions" (Job Market Paper).
"The Welfare Analysis of Depressed Migrant Wage in China: A Dynamic View"
"Intergeneration Mobility and Inequality in China"

Work in Progress

"Constrained Efficiency in an Incomplete Economy with Endogenous Job Risk"

Teaching Experience

Teaching Fellow, Introductory Macroeconomics, Department of Economics, Boston University, Spring 2015
Teaching Fellow, Macro Theory II (first year Ph.D core), Department of Economics, Boston University, Spring 2013, Spring 2014
Teaching Fellow, Macro Theory I (first year Ph.D core), Department of Economics, Boston University, Fall 2012, Fall 2013
Teaching Assistant, Intermediate Macroeconomics, Department of Economics, Boston University, Fall 2011, Spring 2012

Computer Skills: MATLAB, R, STATA, Microsoft Office, LaTeX

Citizenship: China

Contact: 270 Bay State Road, Boston MA 02215 USA

Email: eiyang@bu.edu