The Persistence of Development Dynamics: Financial Frictions and Mobility Distortions*

Ei Yang†

Job Market Paper

March 13, 2016

Abstract

Successful economic reforms produce long-lasting transitional dynamics for developing countries. This paper analyzes how financial frictions and mobility distortions generate the persistence of post-reform development dynamics. I build a general equilibrium model with heterogeneous agents, occupation choice and rural-urban migration, and I calibrate it to China. The mobility distortion is an occupation distortion and it restricts a proportion of agents to the low-productive sector. A removal of distortions triggers the transition of the economy. The transitional path from calibration displays slow convergence. It shows persistent increases in output, productivity and urbanization, mimicking the patterns observed in data. The mobility distortion generates the slow convergence by creating more high-ability, but poor, agents. After the reform removes the distortions, it takes a long time for these agents to become entrepreneurs and to reach their efficient scales due to the financial friction. Compared with the literature that uses tax distortions, the economy with mobility distortions generates slower convergence. The model also generates the rural-urban migration, which contributes to the persistence, by providing high-ability workers and potential entrepreneurs after the reform.

*I am very grateful to Simon Gilchrist, Jianjun Miao and Robert G. King for their advice, guidance and endless encouragements. I also would like to thank Dilip Mookherjee, Alisdair McKay, Ying Lei, Weize Chen and various seminar participants for valuable comments and discussions.

†Department of Economics, Boston University, eiyang@bu.edu.

Keywords: Financial Frictions, Mobility Distortions, Development Dynamics, Urbanization

JEL Classification: E21 E44 O11 O16 O18 O25 O41 O47 O53

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).

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 Rebelo (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 Rebelo (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.

3 Motivation Evidence from China

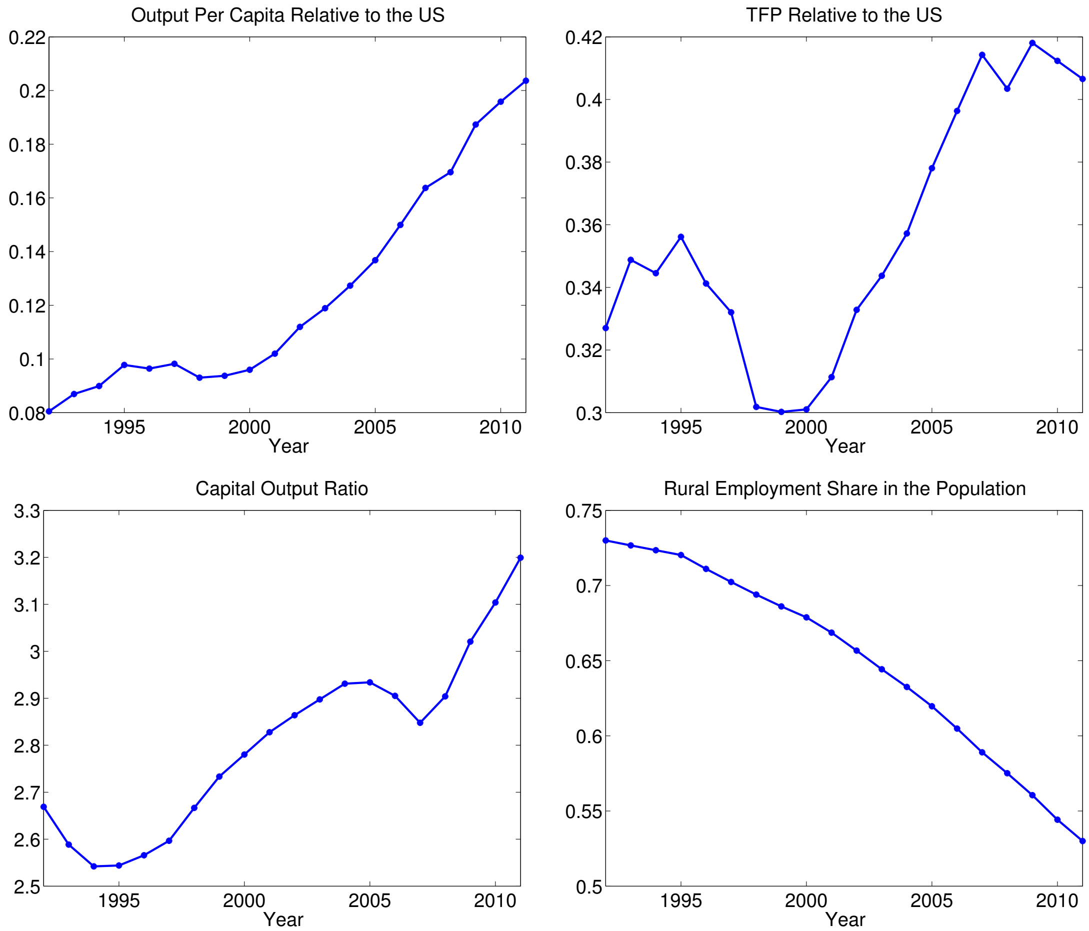
There are two major features of China: a long-lasting growth after reform and the Hukou policy.

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). 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.

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

Figure 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.

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.

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.

4.1 Environment

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 $dz = \left[\psi\left(\mu - \log z\right) + \frac{\sigma^2}{2}\right] zdt + \sigma z dW$. 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.

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:

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

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.

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]

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

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.

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

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.

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.

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.

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

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

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.

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.

5.1 Calibration

5.1.1 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%.

Table 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%

5.1.2 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. Up to now, all parameters relevant to the terminal state are pinned down.

Table 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 Figure 3). Given these parameters, the initial equilibrium of the benchmark economy is pinned down.

Table 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

[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.

benchmark as possible. In this sense, the three economies have a similar distortion in terms of rural employment share (see Table 4).

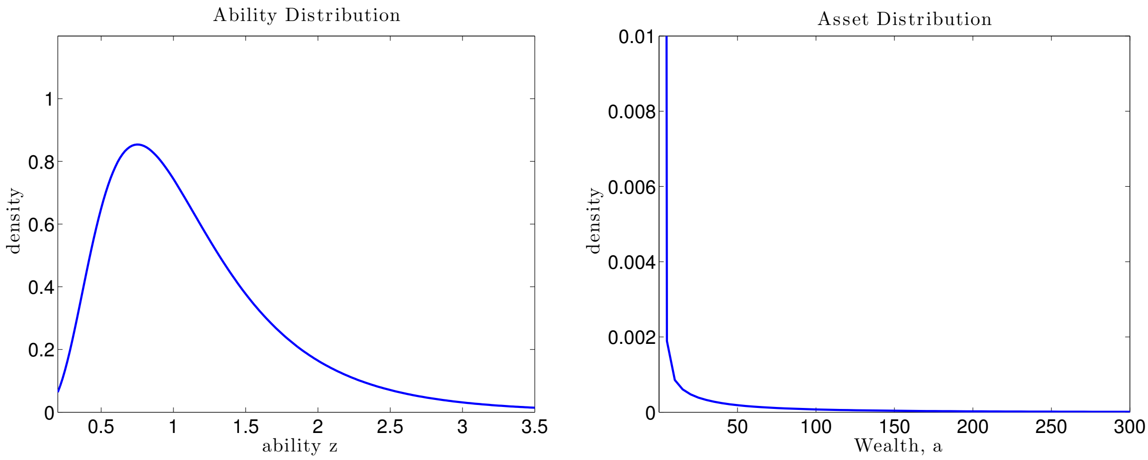
Table 4: Calibrated Parameters for Initial Conditions

	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%

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.

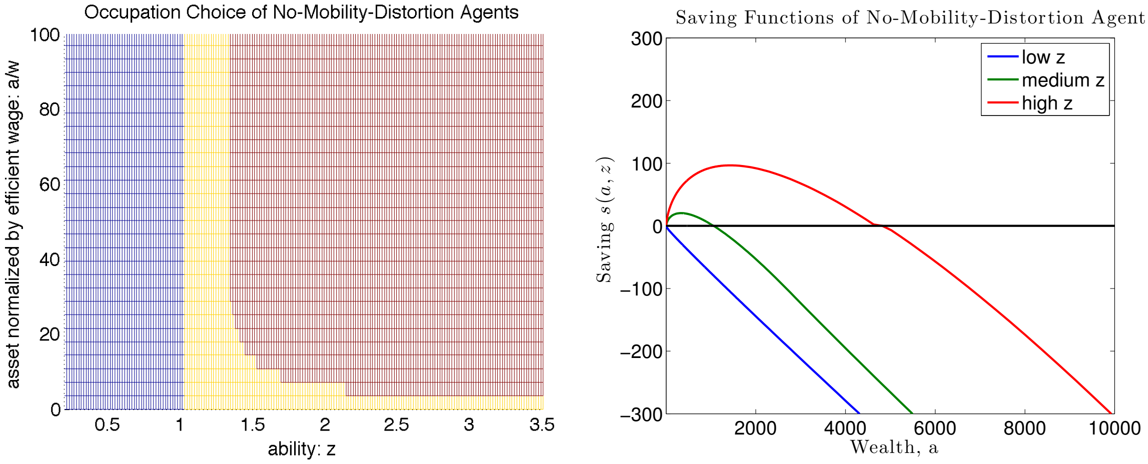
Figure 2: Ability Distributions and Asset Distribution



The diffusion process of the ability generates a stationary log-normal ability distribution (see Figure 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 3: Policy Functions for the Agents without Mobility Distortion



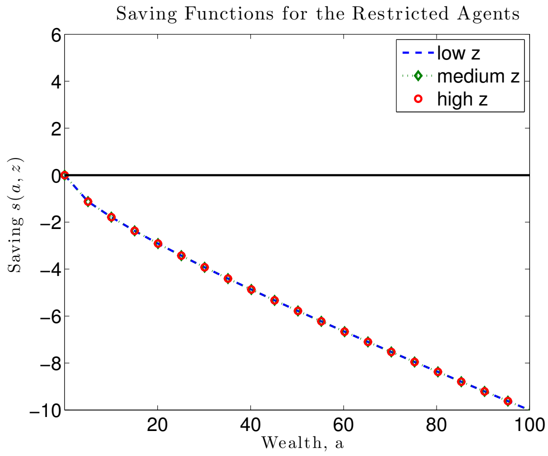
For the agents not restricted by mobility distortions, their occupation choices and saving functions are affected by financial frictions and the lump-sum tax (see Figure 3). In current calibration, if there are no financial frictions and distortions, the high-ability agents will choose to be entrepreneurs. 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 4).

Figure 4: Policy Functions for the Agents with Mobility Distortion



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.

5.3.1 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 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.

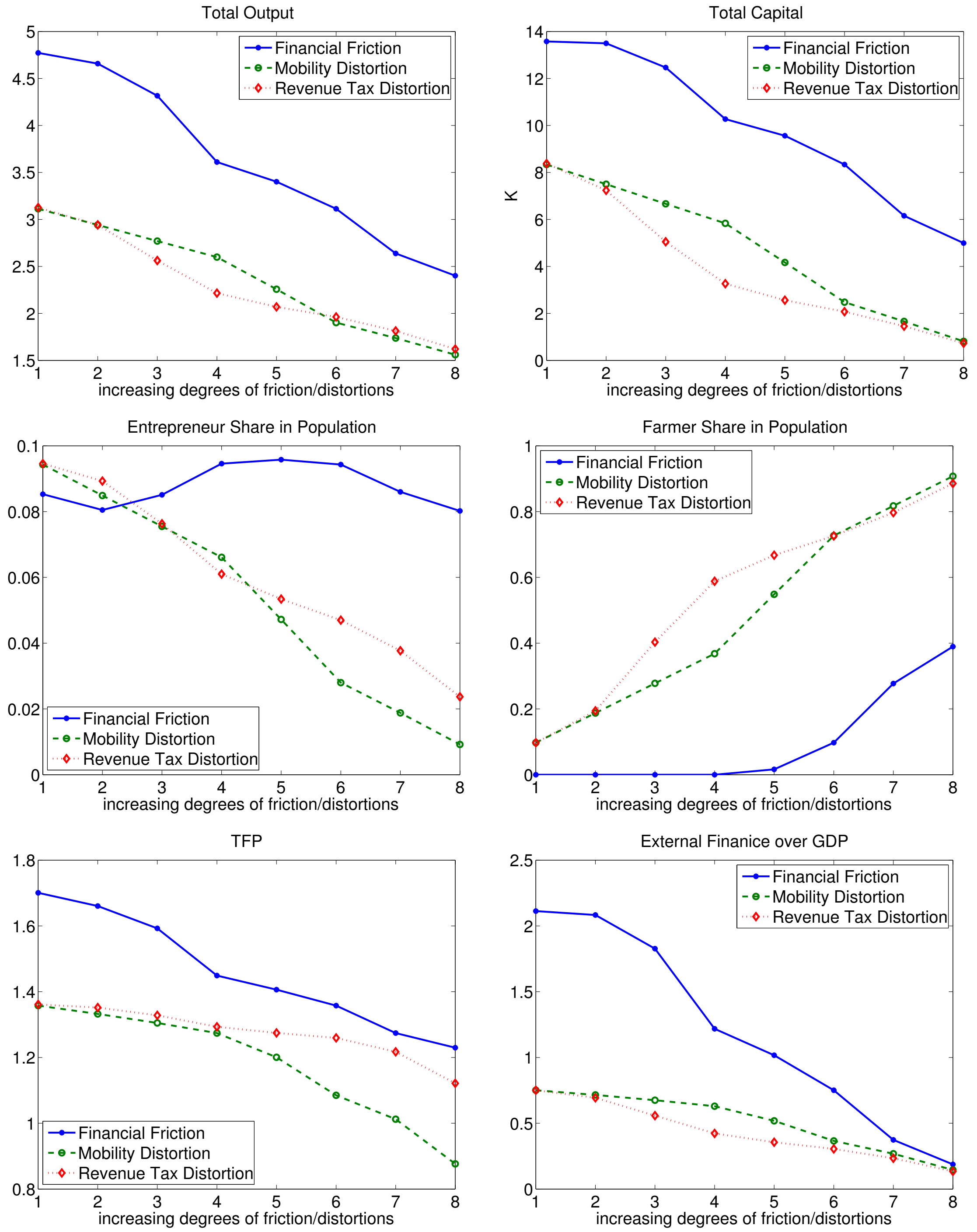
Financial frictions lead to a low level of entrepreneurship. The steep drop in the external finance-to-GDP ratio reflects the channel financial friction affecting the economy. Given the collateral constraint, some of the talented, but poor, agents cannot be entrepreneurs, and some are operating on a 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.

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.

5.3.2 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

Figure 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 $q_3 = 0, 0.01, 0.03, 0.05, 0.06, 0.068, 0.08, 0.1$.

revenue tax distortions can change outcomes of the stationary equilibrium. See Figure 5.

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.

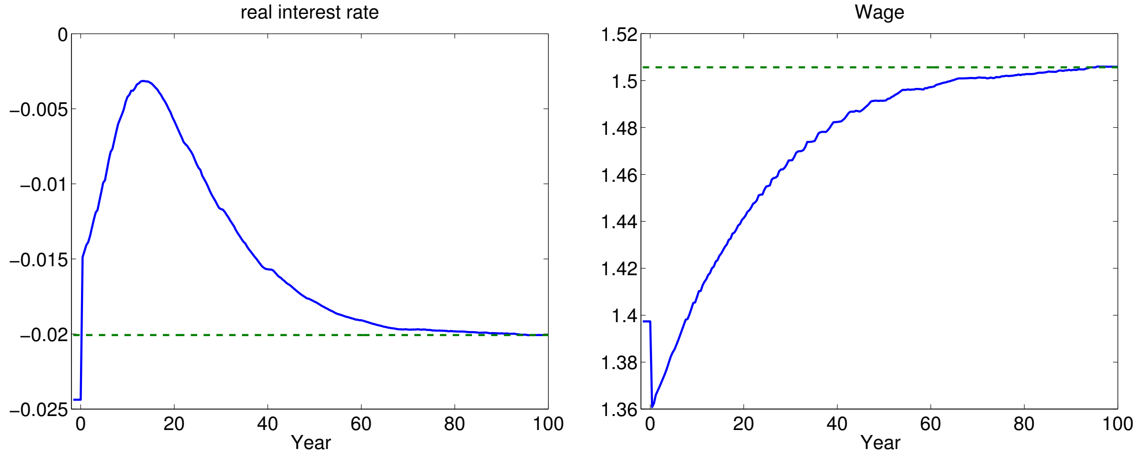
5.4 Transition Dynamics

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.

5.4.1 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 re-

Figure 6: Price Evolutions in the Benchmark



Note: The dotted line denotes the level of variable in the new stationary equilibrium.

alize 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 6).

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 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.

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 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.

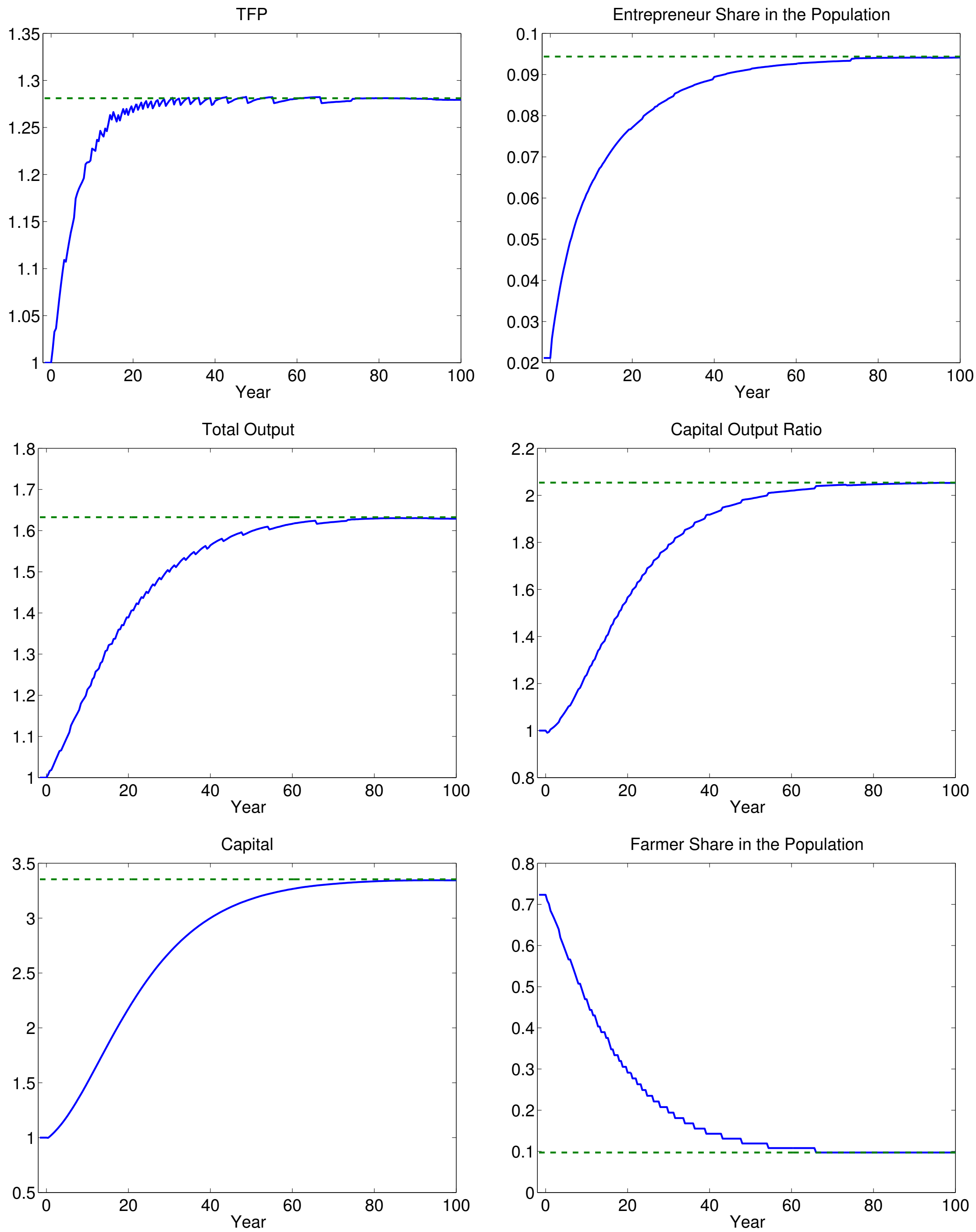
The rural-urban migration in this model contains high-ability workers and potential entrepreneurs. This feature does not depend on the mobility distortion. As long as there are heterogeneities in ability and occupation choice, this channel exits. The mobility distortions only strengthen this channel 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.

5.4.2 Comparisons with Two Alternative Economies

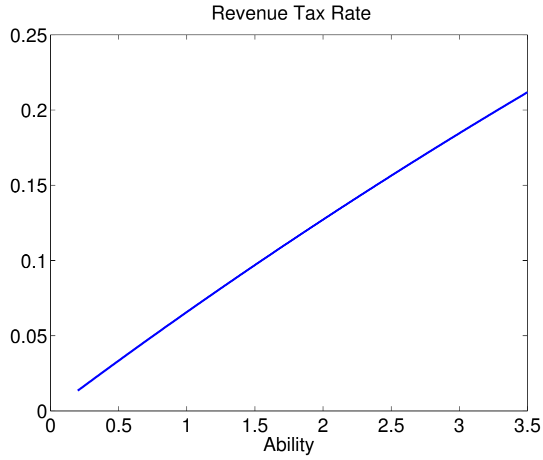
The mobility distortions and the rural-urban migration generate more persistence in the transitional dynamics compared with the revenue tax distortion. This subsection explains

Figure 7: Transitional Dynamics in the Benchmark



Note 1: The dotted line denotes the level of the variable in the new stationary equilibrium.
Note 2: The total output, TFP, capital, and capital output ratio are normalized by their pre-reform level.

Figure 8: Revenue Tax Distortions



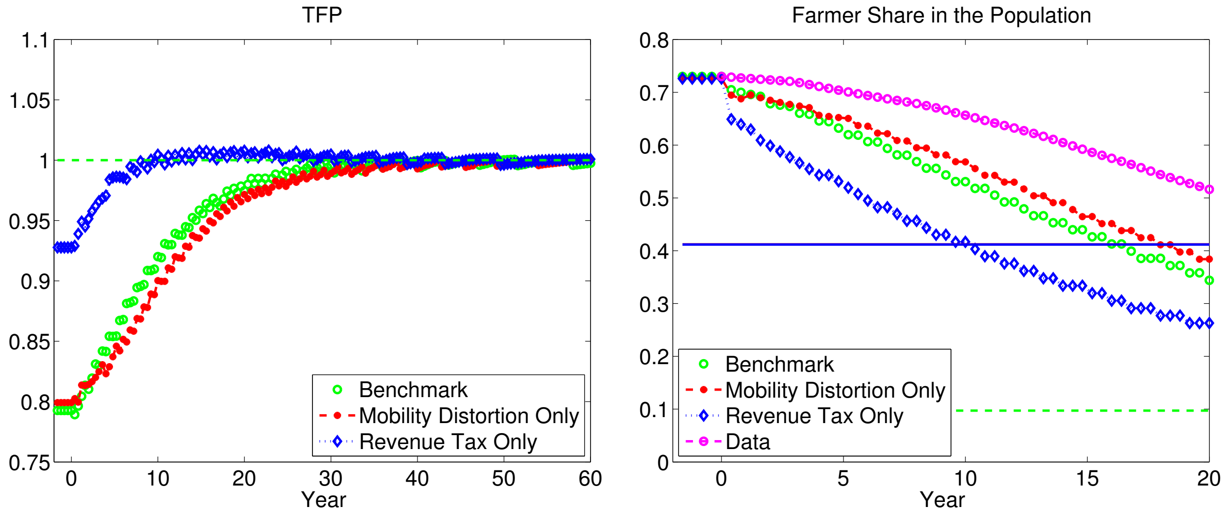
the mechanism and shows the quantitative differences through the comparisons between three economies.

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 is in Table 4 in Section 5.1.2.

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 8).

Figure 9: Comparisons between Economies

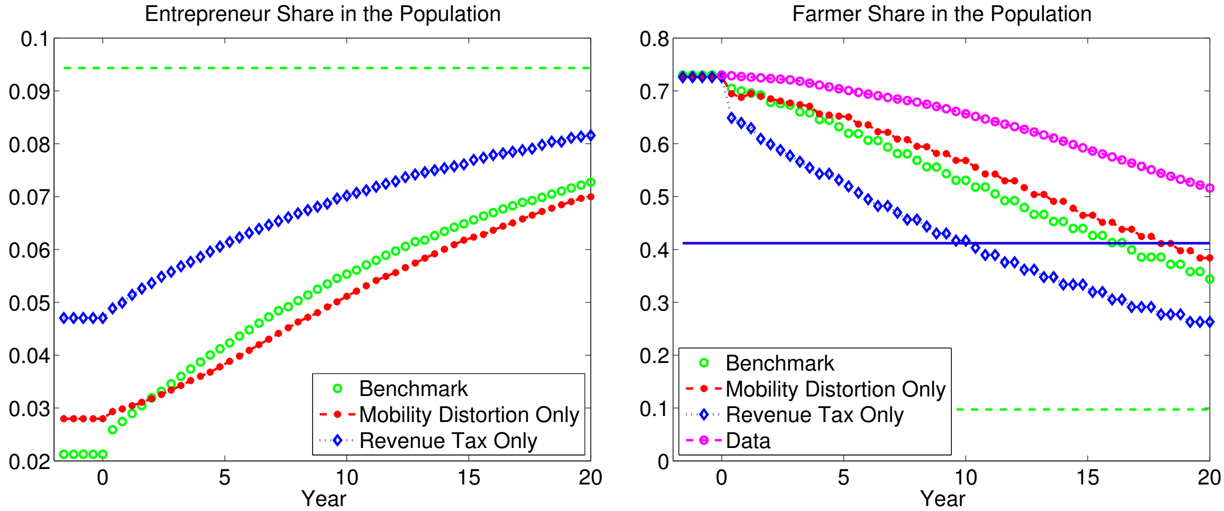


Note: The solid lines are the middle points of the initial level and the terminal level.

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.

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.

Figure 10: Comparisons between Economies



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

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 10). 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.

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.

References

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

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

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

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

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

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

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

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

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

Chang, C., K. Chen, D. F. Waggoner, and T. Zha (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., A. Demirguc-Kunt, E. Feyen, and R. Levine (2012). Benchmarking Financial Systems Around the World. Policy Research Working Paper Series 6175, The World Bank.

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

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

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

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

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

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

Hsieh, C.-T. and P. J. Klenow (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., A. Imrohoroglu, and K. Chen (2006). The Japanese Saving Rate. *American Economic Review 96*(5), 1850–1858.

King, R. G. and S. T. Rebelo (1993). Transitional Dynamics and Economic Growth in the Neoclassical Model. *American Economic Review 83*(4), 908–31.

Levine, R. (2005). Finance and Growth: Theory and Evidence. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1 of *Handbook of Economic Growth*, Chapter 12, pp. 865–934. Elsevier.

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

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

Midrigan, V. and D. Y. Xu (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.

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

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

Song, Z., K. Storesletten, and F. Zilibotti (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.

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.

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

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

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.

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.

8.1.1 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)

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

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.

$$\begin{aligned}
\partial_{a,B} v_{i,j} &= \frac{v_{i,j} - v_{i-1,j}}{\Delta a} \\
\partial_{a,F} v_{i,j} &= \frac{v_{i+1,j} - v_{i,j}}{\Delta a}
\end{aligned}$$

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

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

Substitute the definition in and simplify, we have

$$\begin{aligned}
\frac{v^{n+1}_{i,j} - v^n_{i,j}}{\Delta} + \rho v^{n+1}_{i,j} &= u(c^n_{i,j}) + x_{i,j}v^{n+1}_{i-1,j} + y_{i,j}v^{n+1}_{i,j} + z_{i,j}v^{n+1}_{i+1,j} + \chi_j v^{n+1}_{i,j-1} + \zeta_i v^{n+1}_{i,j+1} \\
x_{i,j} &= -\frac{(s^n_{i,j,B})^-}{\Delta a} \\
y_{i,j} &= -\frac{(s^n_{i,j,F})^+}{\Delta a} + \frac{(s^n_{i,j,B})^-}{\Delta a} - \frac{\sigma^2_j}{(\Delta z)^2} - \frac{\mu_j}{\Delta z} \\
z_{i,j} &= \frac{(s^n_{i,j,F})^+}{\Delta a} \\
\chi_j &= \frac{\sigma^2_j}{2(\Delta z)^2} \\
\zeta_j &= \frac{\mu_j}{\Delta z} + \frac{\sigma^2_j}{2(\Delta z)^2}
\end{aligned}$$

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

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

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

8.1.2 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.

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 $t=0,...,T$.

8.2.1 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.

8.2.2 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:

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