The Politico-Economic Dynamics of China's Growth

Yikai Wang*

October 2018

Abstract

China's rapid growth has been driven by policy reforms and political changes that significantly reduce market frictions. Policy reforms are determined by the government according to its own politico-economic considerations. This paper embeds these politico-economic considerations in a macro model of China to endogenously study government policies, market frictions, and growth. In the model, an elite runs the government and maximizes its own income, facing a political constraint: getting enough supporters. It provides high enough incomes for state workers to buy their support. It also controls capital allocations in the state and the private sectors to balance between keeping enough supporters and extracting more taxes from the private sector. These policies initially generate rapid growth accompanied by declining labor and capital market frictions but keep the frictions persistent in the long run, which is harmful to growth. The calibrated model can quantitatively account for salient aspects of China's recent development and provide predictions for future dynamics. (JEL: E22 E24 O41 O43 P16)

^{*}Wang: University of Oslo, Department of Economics, Box 1095 Blindern, 0317 Oslo, Norway. Email: yikai.wang@econ.uio.no. I am thankful to the editor and anonymous referees for helpful comments. I also thank Daron Acemoglu, Zheng Song, Kjetil Storesletten, Xiaodong Zhu, Fabrizio Zilibotti, and participants of many seminars and conferences for their suggestions. The research has received funding from the Swiss National Science Foundation (PBZHP1-138697), ERC grant (324085), and RGC Hong Kong (HKU 742211B). An earlier version of this paper was circulated under the title "Will China Escape the Middle-Income Trap? A Politico-Economic Theory of Growth and State Capitalism."

1 Introduction

China has been growing at a stellar rate for about 40 years. During this era, several policy reforms have been taken to reduce market frictions. The reduction of frictions has resulted in rapid economic growth. For example, in 1994, the government initiated the so-called "grasping the large and letting the small go" policies on state-owned enterprises (SOEs), and in late 1997, the 15th Congress of the Communist Party of China officially acknowledged the private sector as an important pillar of the economy, and these reforms led to a rapid state-to-private transition. Song et al. (2011) show that the labor and capital reallocations to the more efficient private sector have contributed significantly to the rapid growth since 1998. It is generally acknowledged that policy reforms have been crucial for the great transformation of China (see Chow (2015)) and that they have been determined by the government according to its own politico-economic considerations (see Shirk (1993) for the political logic of economic reforms). Understanding these considerations and policy reforms is crucial for understanding China's economy. This paper takes an important step toward this direction by incorporating into a macro model the objective function of the government and its political constraint: The government is run by the elite who maximizes its own income, subject to the constraint that enough workers support the government. Then the policy reforms and the dynamics of market frictions can be endogenously generated and studied. Providing the microfoundation for policies and frictions is useful for studying the dynamics of frictions and growth in the past, and necessary for predicting future frictions and growth. Otherwise, the government policies and the dynamics of frictions must be assumed exogenously, usually following past trends. This assumption can be largely counterfactual. For example, the recent trends of labor and capital wedges and allocations significantly deviate from the trends before 2008. Hsieh and Song (2015) document that after 2007, the capital wedge between state and private firms stopped declining. Storesletten and Zilibotti (2014) show that the reallocation of labor to private firms has stopped since 2008. Therefore, a model endogenously generating



Figure 1: Private employment share in manufacturing and growth of GDP per capita, PPP adjusted. Source: Song et al. (2011) and The World Bank.

these changes in frictions is necessary for predicting future dynamics.

How does the government set policies to satisfy the political constraint - getting enough supporters? First, it provides state sector workers high incomes, using both wages and transfers, in order to buy their support. This policy creates labor market distortions. For example, Ge and Yang (2014) find that China's state sector workers have been enjoying a wage premium between 20% and 30%, and this labor market wedge has been persistent and even increasing. State sector workers, not surprisingly, are found to be more supportive of the current regime and less supportive of democracy. Chen and Lu (2011) study survey data of Chinese individuals and show that state sector workers and the middle class are less supportive of democratic values, for example, multiparty competition, freedom of demonstrations, etc. Second, the government balances the capital allocations in state and private firms to maintain enough workers in the state sector. As shown in Storesletten and Zilibotti (2014) and reproduced here in Figure 1, China's private employment share in manufacturing stopped increasing when it reached about 60% in 2008, after ten years of rapid growth. In other words, state sector employment has stopped declining and has stayed around 40% since 2008. How does the state maintain this significant fraction of workers, while state firms are less productive than private firms (see Brandt and Zhu (2010))? It does so through the capital allocation in favor of state firms. State firms get more financial resources, for example, bank loans, so they can keep investing and

hiring a large fraction of workers. Brandt and Zhu (2010) document that the state sector's investment share stays around 60% though their employment share has been declining from 1998 to 2007.

In this paper, I embed the government's political constraint in a two-sector growth model to study how political considerations shape government policies and market frictions and how policies and frictions affect growth. In the model, a political elite runs the government, extracts surpluses from state firms, and taxes the private sector.¹ However, it faces a political constraint, that is, it must have support from a sufficient number of workers. The government uses the following policy tools to maintain support and to maximize the elite's income: in the labor market, setting wages for and transfers to state sector workers, and in the capital market, controlling capital allocations to state and private firms. First, the government sets state workers' incomes sufficiently high - as high as incomes they could expect in the democracy - so that state workers prefer to support the current regime instead of democracy. The government sets state workers' incomes by setting wages for and transfers to state workers, at the following costs: labor distortion from regulating wages and the direct cost of using transfers. If transfers from the government are large (given the targeted incomes of state workers, wage burdens on state firms are low), then state firms tend to hire more workers than efficiency would dictate. The redundant labor generates a labor wedge: The labor productivity in the state sector is lower than in the private sector. Second, the government can also control capital allocations in the state and private sectors to keep enough workers in the state sector. When the private sector is small and its capital level is low, maintaining enough political support is not costly, because the number of state workers is large and the expected wages and incomes in the democracy are low. When private sector capital grows, a trade-off for the government emerges: A larger private sector contributes more tax revenues, but also leads to a higher cost for maintaining the regime, because a growing private sector increases wages and competition

¹"The elite" and "the government" can be used interchangeably in this context, except later when I discuss democracy.

for workers. When private sector capital grows to a high level such that state employment declines to the critical level for sustaining the regime, the government must maintain a high enough investment in the state sector to hire enough workers there. If the private sector capital level grows too high, the government may even prefer to limit the size of the private sector, for example, by reducing loans to private firms.

Because the government changes capital market policies as private sector capital grows, the economy's growth pattern changes accordingly. It develops along a three-stage politicoeconomic transition. The first stage is rapid growth, during which the private sector grows rapidly, capital and labor are reallocated to the private sector, and the productivity gain from the reallocation contributes to the rapid growth. As privatization continues, the state employment share declines to the critical level, and then the economy enters the second stage, state capitalism. In this stage, the government over-invests in the state sector to keep its employment share sufficiently high. This stage features a halt to privatization and an increase of investment in the state sector. Though the relative size of the private sector does not change, its absolute level of capital still grows, while state sector capital grows proportionally. The economy still grows rapidly, partly because of the large investment. As private sector capital keeps growing, the economy enters the third stage, and there are two possible cases. The first is the *middle-income trap*. The government chooses to maintain the existing political system and the necessary frictions. It creates barriers to private sector capital growth, for example by tightening lending to private firms. Private sector capital growth slows and the economy stops growing before reaching the efficient output level. This case happens if the cost of sustaining the regime is low, e.g., if costs of transfers and investments are low. The other possible case is sustained growth. As private sector capital grows, it becomes too costly to continue investing in the less efficient state sector and therefore the government chooses to change the political system to democracy. Frictions in the labor and the capital market disappear, and the economy keeps growing until the output reaches the efficient level.

The model is calibrated to China's manufacturing sector and can account for salient aspects of China's recent growth experience, especially the following three: the state-toprivate transition, the dynamics of labor and capital market wedges between state and private firms, and output growth. The first stage - rapid growth - corresponds to the Chinese economy from the 1990s until 2007. During this period, the private sector employment share increases rapidly, the productivity of labor and capital in state firms relative to that in private firms increases, and output grows rapidly. See private sector employment growth and GDP per capita growth in Figure 1. As documented in Storesletten and Zilibotti (2014), the private employment share in the manufacturing sector increases from 15% in 1998 to 60% in 2007. The model in this paper can generate this rapid labor reallocation from state to private firms, with the help of the rapid increase of private sector capital, which benefits from the capital reallocation from state firms and also from the capital accumulation of private firms. Second, the productivity of labor and capital in state firms increases relative to that of private firms. Hsieh and Song (2015) document an increase of the labor productivity of state firms relative to that of private firms from 60% in 1998 to 75% in 2007, and an increase of relative capital productivity from 34% to 46%. Using the total factor productivity (TFP) growth of state firms relative to that of private firms implied by Hsieh and Song (2015) as the input, the model can successfully account for the dynamics of both labor and capital wedges. Third, the model generates a high output growth rate of above 10%, which is consistent with the high GDP per capita growth in China during this period.

The *state capitalism* stage started when the employment share of the state sector reached the critical level in 2008. The state employment share stays around 40% afterward. The model generates this stagnation of labor reallocation because the political constraint requires sufficiently many state sector workers. It also generates a continuous increase of the labor productivity of state firms relative to that of private firms, and a much slower increase of the capital productivity of state firms, consistent with the findings in Hsieh and Song (2015). Capital productivity grows slowly for two reasons: overinvestment in the state sector to maintain supporters and decline of transfers relative to incomes. As incomes increase, the cost of using transfers increases, and the relative size of transfers declines. The decline of transfers leads to increases of wages, the labor productivity, and the capital-labor ratio in the state sector, and consequently to slow growth of capital productivity. Finally, the growth rate decreases, but not too much, because of the large investment, consistent with the discussion in Zilibotti (2017). The output growth rate in the model is still above 5% in 2017 and stays above 4% until 2027.

The model also provides predictions for future dynamics of wedges and growth, and for potential development paths. The third stage in the calibrated economy is the *middleincome trap*, where labor and capital market frictions persist. These frictions harm economic growth, which slows down and stops before the income of the economy converges to the high level of developed countries. The government chooses to maintain these frictions in order to sustain the current regime. For example, it restricts lending to private firms to maintain the state employment share, resulting in persistent and even increasing capital market frictions. There is another possible development path: *sustained growth*. This path emerges if costs of investment and transfers are sufficiently large - much larger than in the benchmark calibration. Then the elite finds sustaining the regime too costly and chooses to democratize. The frictions are reduced and growth continues until income reaches the high level.

This paper contributes to three strands of literature. The first is on China's economic growth given labor and capital market frictions. This paper contributes by providing the microfoundation for capital and labor market frictions and by studying their endogenous evolution. Brandt and Zhu (2010) and Brandt et al. (2012) use a three-sector growth model with capital wedges between state and private firms to account for China's growth. They identify private sector TFP growth as the main contributor. Song et al. (2011) construct a two-sector growth model to study China's growth in the transition from the state to the

private sector, given exogenously the financial constraints on private firms. Cheremukhin et al. (2015) study China's structural changes and the dynamics of wedges in the labor and the capital markets and their contributions to growth from 1953. This paper offers explanations for the dynamics of labor and capital market wedges and frictions in the past, provides predictions for their future dynamics, and studies their implications for future growth. Second, this paper contributes to the study of the interaction between political development and economic development. One aspect of the interaction is that political institutions affect economic development. For example, Acemoglu (2008) studies the economic performance in oligarchic societies, in comparison with that in democratic societies. He shows that an oligarchic society may achieve higher efficiency at first because the elite protects its property better and invests more than in a democracy. However, in the long run, the elite blocks new entrepreneurs from entering, and economic growth becomes slower than in a democracy. This paper's analysis of growth in the oligarchy is in the spirit of Acemoglu (2008), but with an important difference regarding the implications for efficiency. In this paper, the higher short-run growth in the oligarchy is not due to higher efficiency but is precisely because of its inefficiency and distortions in the labor and the capital markets: The labor market friction leads to a low initial output such that growth in the beginning can be high when the friction declines; and overinvestment in the state sector in the state capitalism stage is inefficient but keeps the growth rate high. The mechanism generating slower long-run growth in the oligarchy is similar to that in Acemoglu (2008): In the long run, the elite creates barriers for entrepreneurs. The other aspect of the interaction, i.e., how economic development affects political development, is also studied in my paper. Economic progress may lead to political progress, e.g., democratization, but the latter happens only under the right conditions. The impact of economic development on political progress described in this paper is related to but different from the modernization theory started by Lipset (1959), which suggests that economic progress is sufficient for political progress such as democratization. The third strand is the literature on the middle-income

trap. There are many important factors determining the middle-income trap, such as the income threshold, labor costs, population dynamics, etc. This paper contributes to the literature by studying one of them - government policies. Why do some countries successfully adopt good policies and rapidly grow out of poverty but then suddenly fail to implement appropriate policies and to become rich countries? This paper offers a theoretical explanation: The interest of the elite aligns with growth in early stages of development but not in later stages. It also studies various government policies that could help middle-income countries to grow further and the conditions that induce governments to implement these policies.

The rest of the paper is organized as follows. Section 2 presents the model. In Section 3, the model is calibrated to the manufacturing sector in China from 1998 to present. The quantitative model accounts for the time series in China's recent development and offers predictions for future development. Extensions of the model, including the alternative development path and an calibration of China's urban economy since 1994, are also studied. Section 4 concludes. Proofs and more details of the model are in the Online Appendix.

2 The Model

This section presents a two-sector growth model, in which the dynamic general equilibrium is deterministic and agents have perfect foresight. The two sectors are the state (S) and the private (P) sector. There is a continuum of state firms and private firms. The firms are neoclassical: They produce the same final goods using capital and labor and maximize profits taking prices as given. There are three groups of infinitely many agents: the elite (e), private entrepreneurs (p), and workers (w). The population size of workers is normalized to 1, while the sizes of the elite and entrepreneurs are both infinitesimal - denoted as ε . The elite supplies capital to state firms, and entrepreneurs supply capital to private firms.

They receive incomes from capital returns, consume, and save. Workers supply labor and receive wages. A worker may work in an S firm or a P firm. We use a representative agent to describe the behavior of a group of agents, i.e., we use "the elite," "the entrepreneur," "the S firm," and "the P firm" to represent the corresponding groups of agents. We refer to workers as "S workers" and "P workers," respectively. ²

Political elements are incorporated into the growth model. The economy starts in an oligarchic regime, where the elite controls the government but faces a political constraint: It needs sufficiently many workers to support the regime. If the number of supporters is large enough, the regime survives; otherwise it collapses, and democratization occurs. Each worker decides whether to support the regime by comparing her income in the current regime with her expected income in the democracy. To maintain enough supporters, the government can strategically increase state workers' incomes using the following three policies: (1) setting state sector wages, (2) giving transfers to state workers, and (3) allocating capital in the state and the private sector.

In this section, I first describe the model setup in Subsection 2.1 to 2.4, and then solve for the model and characterize some analytical properties of the solution in subsections 2.5 to 2.7.

2.1 Preferences, Technology, and Markets

Agents live for infinite periods. The lifetime utility of the elite is the discounted sum of utility in all future periods:

$$U = \sum_{t=0}^{\infty} \beta^t \log\left(c_{et}\right)$$

The elite discounts the future at the rate β and the coefficient of relative risk aversion is 1. The representative entrepreneur has the same utility function as that of the elite and her consumption is denoted as c_{pt} . The entrepreneur supplies capital to the P sector

²Because the population sizes of state sector workers and private sector workers change over time, it is better to discuss them as groups instead of as two representative agents.

and receives capital income. The elite supplies capital to the S sector and receives capital income, and additionally, it has another income source in the oligarchy: the government's budget surplus, which will be discussed in detail when the government is described below. After the government redistributes using taxes and transfers, the elite and the entrepreneur consume and save for the next period.³ Savings can affect capital supplies and capital incomes in the next period. A worker supplies one unit of labor to a firm in each period. A worker is assumed to live hand to mouth and to be myopic: She consumes all her current-period income and cares only about current-period income, even when she makes political decisions.

The S firm and the P firm are different in two aspects: access to the capital market and productivity. First, the S firm rents capital from the S-sector capital market, where capital is supplied by the elite, and the P firm rents capital from the P sector capital market, where capital is supplied by the entrepreneur. Both firms hire from the same pool of workers.⁴ Second, the S firm is less productive than the P firm. The technologies of the S and the P firms are described by the following Cobb–Douglas production functions:

$$Y_{St} = z_{St} K_{St}^{\alpha} L_{St}^{1-\alpha},$$

$$Y_{Pt} = z_{Pt} K_{Pt}^{\alpha} L_{Pt}^{1-\alpha},$$

where z_{jt} , K_{jt} , and L_{jt} , $j \in \{S, P\}$ stand for TFP, capital, and labor in sector j at time t, respectively. TFP sequences are dynamic and deterministic. Capital depreciates at the rate δ . Because firms are neoclassical, their profit maximization problems imply that wages

³Details of government redistribution and the capital market will be discussed in Subsection 2.3 below. ⁴In the equilibrium, all workers prefer to work in the S sector because S workers' incomes are higher, and then the S firm randomly draw a fraction of workers according to its labor demand.

and capital returns are equal to the marginal productivity:

$$w_{St} = (1-\alpha) z_{St} K_{St}^{\alpha} L_{St}^{-\alpha}, \qquad (1)$$

$$w_{Pt} = (1-\alpha) z_{Pt} K_{Pt}^{\alpha} L_{Pt}^{-\alpha}.$$
⁽²⁾

for wages, and

$$r_{St} = \alpha z_{St} K_{St}^{\alpha - 1} L_{St}^{-\alpha}, \tag{3}$$

$$r_{Pt} = \alpha z_{Pt} K_{Pt}^{\alpha - 1} L_{Pt}^{-\alpha}, \tag{4}$$

for gross returns of capital (without depreciation and taxes).

The labor market clearing condition is simply

$$L_{St} + L_{Pt} = 1. (5)$$

In the financial market, there is a representative competitive bank. It has access to the international bond market where the interest rate r is exogenously given. The gross interest rate is denoted by R = 1 + r. So the bank serves as an intermediary which allows domestic agents to save and borrow at this interest rate. The elite can borrow from the bank, and invest the bank loan and its own assets in the S firm. The elite has deep pockets, i.e., there is no constraint on how much it can borrow. The entrepreneur can also borrow from the bank, but she faces a borrowing constraint: the bank loan cannot exceed $\eta_t - 1$ fraction of her assets. In other words, the maximal capital-to-asset ratio in the P sector is η_t :

$$K_{Pt} \le \eta_t a_{pt},\tag{6}$$

where a_{pt} and K_{Pt} represent the entrepreneur's assets and the P firm's capital, respectively. η_t may vary over time, depending on the government's policy, which will be described below in subsection 2.3.

2.2 The Political Constraint

The economy starts in the oligarchy. The representative elite runs the government and sets policies to maximize her lifetime utility. However, she faces the following political constraint: She must obtain political support from sufficiently many workers. Otherwise, a revolution occurs, and the regime switches to democracy. Here we can take the incomes of the elite and workers in the democracy as given, and they are determined by the equilibrium in the democracy, which will be studied in Subsection 2.5. In every period, each worker in the S sector decides whether to support the oligarchic regime. If the number of supporters in the S sector is larger than a critical level \underline{L} , the regime survives in this period.⁵ Let us denote a worker's expected income in the democracy as y_{wt}^D . The expression of y_{wt}^D , depending on state and private sector capital, will be derived in Subsection 2.5. Now the political constraint is equivalent to the following two economic constraints. First, the high-income constraint, i.e., S workers' incomes are high enough:

$$y_{St} = w_{St} + T_t \ge y_{wt}^D,\tag{7}$$

where y_{St} stands for each S worker's income in period *t*, which includes the wage paid by the S firm w_{St} and transfers from the government T_t . Second, the large-state-employment constraint, i.e., the S sector's employment share is high enough:

$$L_{St} \ge \underline{L}.\tag{8}$$

⁵This implies that having a large enough number of supporters in the S sector is the sufficient and necessary condition for sustaining the oligarchy. This assumption is without loss of generality, because P workers' incomes are always lower, not only than S workers' incomes but also than workers' expected incomes in the democracy, and P workers always support democratization. This political process can be formalized as a global game according to Morris and Shin (2000). Details of the game are in the Online Appendix.

2.3 The Government

In the oligarchy, the elite controls the government and decides government policies.⁶ In the beginning of each period, the elite can choose to democratize voluntarily or to stay in the oligarchy. If the first is chosen, the regime switches to democracy forever. If the latter is chosen, then the elite, or equivalently the government, sets the following policies: the S sector wage, transfers to each S worker, S sector capital, and the credit constraint of the P firm.⁷ Moreover, the government taxes P sector workers and the entrepreneur at an exogenous rate $\tau > 0$.⁸ To simplify the expressions, I assume that the tax is imposed on the gross return on the entrepreneur's investment in the P firm, i.e., $r_{Pt}K_{Pt}$, so that taxes from P sector workers and the entrepreneur together equal τ fraction of the P sector output.

The cost of transfers to each S worker is a convex function of the transfers: $b_t T_t^2$, where T_t is the transfers to each S worker, and the cost parameter b_t can change over time. Given S workers' incomes and the S firm's capital, higher transfers imply lower wages paid by the S firm, higher labor demand, and lower labor productivity of the S firm. In the model, we interpret the variable T as cash transfers to state workers, while it can be more generally interpreted as all real-world policies similar to cash transfers, which increase political support of state workers and are financed by the government. These policies include non-cash benefits, subsidies and political propaganda, e.g., housing, education, and tax benefits for state workers, labor subsidies to state firms, which increase state workers'

⁶We can use "the elite" and "the government" interchangeably in this context.

⁷S sector capital is set optimally for the elite as a group. This model setup implies that the collective action problem for the elite is assumed away: every individual elite member supplies capital according to the decision made by the elite as a group, though she may have incentives to choose a different level of capital that she supplies.

⁸The tax rate can be endogenized. I also extend the model and allow the government to choose the tax rate of the entrepreneur given an upper bound $\bar{\tau}$, which is exogenously determined by the state capacity described in Besley and Persson (2009). Then the government can potentially choose $\tau < \bar{\tau}$. However, given the calibration in this paper and the initial assets of the entrepreneur, I find that the government always sets $\tau = \bar{\tau}$ along the transition. So in this paper, I do not need to consider the endogenous tax rate. The case $\tau < \bar{\tau}$ happens only in the event that the entrepreneur's asset is extremely small but this case is not relevant for the years studied in the model. Moreover, in reality, once a tax rate is set, it lasts for a relatively long time and it is not easy to adjust it annually. So, even in a year that the optimal tax rate can be lower, the government, knowing the optimal tax rate will be at $\bar{\tau}$ soon, should set the tax rate according to the optimal level in the medium and long run, instead of on an annual basis.

expected incomes in the oligarchy relative to their expected incomes in the democracy.⁹

The government can set S sector capital without any financial constraint, because it has deep pockets and can borrow as much as it wishes from the international financial market at the interest rate r. The government can also influence the P sector capital: It can set the P firm financial constraint η_t in a bounded region $[\underline{\eta}, \overline{\eta}]$. The government can choose η_t because it can create barriers to entrepreneurs' access to the financial market, or give administrative instructions to state banks regarding how much lending is allocated to entrepreneurs (see Brandt and Zhu (2000)). $\overline{\eta}$ represents the natural borrowing constraint on entrepreneurs if the government does not put additional restrictions on private sector borrowing.¹⁰ $\underline{\eta}$ represents the leverage of the private firm given the highest level of restriction that the government can set.¹¹ The setting on the financial market is similar to that used in Song et al. (2011), while an important difference is that capital in the state sector and credit constraints in the private sector are endogenously determined by the government in this paper. Therefore, this model can endogenously account for the past dynamics of capital market frictions and can be used to predict the future trend of capital market frictions.

The revenues in the government budget include the profit from the S firm and taxes from the P sector, and the expenditure is the cost of transfers to S workers. The budget surplus in every period is assumed to be claimed by the elite; so the government's budget is assumed to be balanced in every period.¹² The government's budget constraint can be

⁹For example, if the government gives state workers better access to public schools, or starts a political campaign trying to convince state workers that the current regime is very efficient or that democratization leads to large risks and costs, state workers will be more likely to prefer to stay in the current regime. These policies work in the same way as cash transfers. Moreover, if some policies and factors change the workers' utility without affecting their incomes in the oligarchy and in the democracy, e.g., non-monetary benefits of democracy, then such policies and factors can also be captured by variable T in the model. See more details about the general definition of transfers in the Online Appendix.

¹⁰In this case, there is still a natural limit on the borrowing of private entrepreneurs, due to the following moral hazard problem: If a private entrepreneur gets a loan that is too large to be supported by her assets, she may choose to default on the loan and leave with the money.

¹¹For example, if the most stringent policy the government can set is to let the bank lend nothing to the entrepreneur, then the entrepreneur has to finance her investment using her own savings. This implies $\eta = 1$.

 $[\]frac{1}{12}$ We can also assume that the government's budget does not need to be balanced in every period, and

written as

$$\pi_{St} + \tau Y_{Pt} = b_t T_t^2 L_{St} + y_{et}, \qquad (9)$$

where π_{St} is the profit of the S firm and y_{et} is the elite's non-asset income which equals the whole government surplus. The government's budget constraint can be used to derive the expression for y_{et} .

$$y_{et} = \pi_{St} + \tau Y_{Pt} - b_t T_t^2 L_{St}$$

= $\alpha z_{St} K_{St}^{\alpha} L_{St}^{1-\alpha} - r K_{St} + \tau z_{Pt} K_{Pt}^{\alpha} L_{St}^{1-\alpha} - b_t T_t^2 L_{St}.$

In later analysis, we will substituted this expression into the elite's problem, and then the government's budget constraint will be automatically satisfied.

2.4 Timing of Events

The events in each period are the following:

1. In the beginning of period *t*, the elite decides whether to voluntarily democratize. If it decides to, the political system switches to democracy forever; if not, the following events occur.

- 2. Capital allocation: The elite sets K_{St} and η_t , and then the entrepreneur chooses K_{Pt} .
- 3. The elite sets the S sector's wage w_{St} and transfers T_t .

4. The S firm and the P firm hire workers. Workers are randomly selected into the S firm according to its labor demand.

5. S workers decide whether to support the regime. If there are not enough supporters, the political system switches to democracy. If there are enough supporters, the oligarchy

the results are the same. For example, if the government transfers more to the elite than the government budget surplus in a certain period, the the gap is financed by government debt and will be paid back in the future. Consequently, the government's budget surpluses will be smaller in the future, given the No-Ponzi condition. If the interest rate for the government debt equals the interest rate for the elite's savings and loans, then the government's debt is essentially the elite's debt, and what matters is simply the elite's assets minus the government debt, which is equal to the elite's assets when we assume the government budget is balanced in every period.

survives.

6. Firms produce. Labor and capital incomes are distributed. Taxes are collected and transfers are made.

7. The elite and the entrepreneur consume and save. The economy enters the next period.

Notice that the game in the oligarchy may end in either step 1 or step 5 if democratization happens. If so, the exit payoffs of the elite and the workers are determined by the equilibrium in the democracy, discussed below.

2.5 The Democracy

In the oligarchy, workers decide whether to support democratization according to their expected incomes in the democracy, and the elite also decides whether to voluntarily democratize according to their incomes in the democracy. So the outcomes in the democracy are useful for pinning down the equilibrium in the oligarchy. Here I summarize the model setup and some main results in the democracy, especially the outcomes relevant for the equilibrium in the oligarchy, i.e., (1) workers' incomes y_{wt}^D ; and (2) the elite's non-asset income y_{et}^D . The other equilibrium outcomes and conditions are in the Online Appendix.

The major difference is that in the democracy, the government is run by the representative worker.¹³ The government taxes the elite and the the entrepreneur at the exogenous rate $\tau > 0$ and then transfers the tax revenue to all workers. The economy is simply a competitive equilibrium given taxes. The economy is similar to the economy in the competitive equilibrium described by Song et al. (2011).

The competitive labor market implies that wages in the S and the P sectors are the

¹³Because the workers' dominant population size guarantees that they win the majority of votes. In this sense, some partial democracies where the voting systems exist but are controlled by the elite do not correspond to the democracy but rather to the oligarchy in this model.

same, determined by the marginal productivity of the labor.

$$w_t^D = (1 - \alpha) z_{St} (K_{St})^{\alpha} (L_{St}^D)^{-\alpha} = (1 - \alpha) z_{Pt} (K_{Pt})^{\alpha} (L_{Pt}^D)^{-\alpha}$$

= $(1 - \alpha) (z_{St}^{\frac{1}{\alpha}} K_{St} + z_{Pt}^{\frac{1}{\alpha}} K_{Pt})^{\alpha}$, (10)

where L_{St}^D and L_{Pt}^D represent the labor allocation in the democracy and $L_{St}^D + L_{Pt}^D = 1$. The representative worker's income includes the wage and transfers:

$$y_{wt}^{D} = w_{t}^{D} + \tau \left(\alpha z_{St} K_{St}^{\alpha} \left(L_{St}^{D} \right)^{1-\alpha} + \alpha z_{Pt} K_{Pt}^{\alpha} \left(L_{Pt}^{D} \right)^{1-\alpha} \right)$$
$$= \left(1 + \frac{\tau \alpha}{1-\alpha} \right) w_{t}^{D}$$
$$= \left(1 - \alpha + \tau \alpha \right) \left(z_{St}^{\frac{1}{\alpha}} K_{St} + z_{Pt}^{\frac{1}{\alpha}} K_{Pt} \right)^{\alpha}.$$
(11)

The capital market in the democracy is also competitive. Each elite member supplies capital to the S sector, taking prices as given. The major difference from the oligarchy is that in the democracy S sector capital supply is no longer collectedly decided by the elite as a group who internalize the impact of capital supply on prices. The competitive capital market implies that the elite's marginal return from investing in the S firm equals the interest rate r charged by the bank, if the S sector exists. If the S sector does not exist because the P sector capital is large enough and the return to S sector capital is lower than r, then the elite's income simply equals the return from its assets at the rate r, implying that the elite's non-asset income is 0. Basically, the elite "retires" after democratization:

$$y_{et}^D = 0, (12)$$

where y_{et}^D is the elite's non-asset income in the democracy.

The entrepreneur faces the borrowing constraint; so if the marginal rate of return on her investment in the P firm is higher than r, she borrows given the constraint $K_{Pt} \le \eta_t a_{pt}$. The government prefers a higher η_t , which implies a higher K_{Pt} and a higher w_t . So it always sets η to its upper bound:

$$\eta_t^D = \bar{\eta}.\tag{13}$$

The dynamic equilibrium in the democracy is similar to that in Song et al. (2011). If the S sector exists, then the P sector's capital return equals the world interest rate; otherwise it can be lower:

$$r_{St}^{D} \begin{cases} = r & \text{if } L_{St}^{D} > 0, \\ \leq r & \text{if } L_{St}^{D} \le 0. \end{cases}$$

$$(14)$$

If the S sector exists, then the rate of return determines the capital-labor ratio in the S sector as in equation 3, and the latter determines the wage as in equation 1. The wage in the P sector is the same as that in the S sector, and therefore, the P sector's wage w_{Pt} and rate of return to capital r_{Pt} are determined by equations 2 and 4, respectively. If r_{Pt} is high enough, the entrepreneur's assets and the P sector capital grow over time. Eventually, all workers move to the P sector and the S sector disappears. In this case, the P sector labor equals one, and then the P sector's capital-labor ratio determines the wage and the return to capital. The main results in the democracy are summarized in the proposition below.

Proposition 1. In the democracy, the representative worker's income equals the wage determined by the competitive equilibrium given S and P sector capital, plus transfers which equal $\tau \alpha / (1 - \alpha)$ fraction of the wage.

The elite "retires": it receives only the return on its assets but zero non-asset income, i.e., $y_{et}^D = 0.$

The equilibrium is a competitive equilibrium which determines K_{St} and w_{St} . Moreover, $\eta_t^D = \bar{\eta}$.

2.6 The Equilibrium in the Oligarchy

Given the model setup discussed above, I now describe the problems of the two major agents (the entrepreneur and the elite), define the equilibrium, and solve for it. First, I formalize the entrepreneur's problem. She maximizes her lifetime utility by choosing the P sector capital supply subject to the borrowing constraint and the deterministic sequences of consumption and savings, taking the return to P sector capital r_{Pt} as given:

$$\max_{\{K_{Pt},a_{pt},c_{pt}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c_{pt}$$
(15)

s.t.
$$c_{pt} = Ra_{pt} + ((1 - \tau) r_{Pt} - \delta - r) K_{Pt} - a_{p,t+1}.$$
 (16)

The No-Ponzi condition is imposed. Each entrepreneur takes r_{Pt} as given because each entrepreneur's choice cannot affect aggregate variables and the political system.

The elite chooses the political system, government policies in the oligarchy, consumption, and savings to maximize its lifetime utility. We can think of the elite's problem as consisting of three steps, (1) choosing the sequence of political systems - $M_t \in \{O, D\}$; (2) setting government policies in the oligarchy, subject to the political constraints; and (3) choosing consumption and savings throughout the lifetime. First, because once democratization happens, democracy is consolidated and the elite can never return to power, the choice on the whole sequence of the political systems is simply to choose the period T^D to democratize, where T^D can also be $+\infty$, representing that the elite chooses to stay in oligarchy forever. We can write $T^D \in \mathbb{Z}_{\geq 0} = \mathbb{Z}_{\geq 0} \cup \{+\infty\}$. Second, in the oligarchy, i.e., for each $t < T^D$, the elite chooses government policies $\{K_{St}, w_{St}, T_t, \eta_t\}$, given the political constraint. The elite takes into account that these government policies affect the aggregate prices and the entrepreneur's choices on assets and capital. After democratization happens, i.e., for each $t \geq T^D$, the elite "retires" and basically receives only the return on its assets, as I discussed in Proposition 1. The third set of choice variables for the elite is very standard: it chooses the consumption and saving subject to the budget constraint. To sum up, we can form the elite's problem as:

$$\max_{T^{D}, \{K_{St}, w_{St}, T_{t}, \eta_{t}\}_{t=0}^{T^{D}-1}, \{c_{et}, a_{et}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} \log c_{et}$$
(17)

subject to the following sets of constraints: (1) the budget constraint

$$c_{et} = Ra_{et} + y_{et} - a_{e,t+1}, (18)$$

where

$$y_{et} = \begin{cases} \alpha z_{St} K_{St}^{\alpha} L_{St}^{1-\alpha} - rK_{St} + \tau z_{Pt} K_{Pt}^{\alpha} L_{St}^{1-\alpha} - b_t T_t^2 L_{St} & \text{if } t < T^D \\ 0 & \text{if } t \ge T^D, \end{cases}$$
(19)

(2) the political support constraint for all $t < T^D$, or equivalently, two economic constraints described by equation 8 and equation

$$w_{St} + T_t \ge (1 - \alpha + \tau \alpha) \left(z_{St}^{\frac{1}{\alpha}} K_{St} + z_{Pt}^{\frac{1}{\alpha}} K_{Pt} \right)^{\alpha}, \qquad (20)$$

which is obtained by substituting the expression of workers' incomes in the democracy (equation 11) into the high-income constraint in the oligarchy (equation 7), (3) that firms optimize and prices are functions of aggregate capital and labor, described in equations 1 to 4, (4) that the entrepreneur optimizes: the sequence of allocations $\{K_{Pt}, a_{pt}\}_{t=0}^{\infty}$ is the solution of the entrepreneur's problem given the sequence $\{r_{Pt}\}_{t=0}^{\infty}$, as described by equations 15, 16 and 6, and (5) that the labor market clears. Moreover, the No-Ponzi condition is imposed to the elite's problem.

It is helpful for understand the elite's problem to compare it with a problem we are familiar with: the household's problem with the endogenous retirement decision. The elite chooses to "retire," or more precisely, to democratize, in period T^D . Before "retirement," it chooses some "efforts" - government policies - to obtain income y_{et} , in addition to the return from its assets, while after that, it lives on the return on its assets but has no

additional income: $y_{et} = 0$ for $t \ge T^D$. Of course, there are also important differences, including the infinite horizon, the political constraint, and the fact that the elite takes into account how its choices affect aggregate variables and prices. The fact that $y_{et} = 0$ after democratization implies that other equilibrium outcomes after T^D do not matter for the elite's problem.

Notice that the equilibrium in the oligarchy is not a competitive equilibrium, because the elite can directly choose aggregate variables and prices, e.g., K_{St} and w_{St} . So the equilibrium is characterized as the solution of a planner's problem, where the planner is the elite, who chooses a sequence of policies, prices, and allocations to maximize its own utility, subject to the budget constraint, the political constraint, and the incentive constraints that other agents - firms and the entrepreneur - optimize. To focus on the equilibrium in the oligarchy, we can take the equilibrium outcomes in the democracy, for example, the elite's income y_{et}^D , as exogenously given. Moreover, firm productivities, the cost parameter of transfers, and other static parameters such as the tax rate, the minimal number of supporters in the oligarchy, and other standard parameters are taken as exogenously given. These exogenous sequences are dynamic, but deterministic. So there are no uncertainties in the perfect foresight equilibrium . In other words, taking $\{z_{St}, z_{Pt}, b_t, y_{et}^D\}_{t=0}^{\infty}$ and other static parameters exogenous, we can formally define the dynamic politico-economic equilibrium starting in the oligarchy as follows.

Definition 1. A sequence of policies, prices, and allocations

 $\{M_t, \eta_t, T_t, w_{St}, w_{Pt}, r_{St}, r_{Pt}, K_{St}, K_{Pt}, a_{et}, a_{pt}, c_{et}, c_{pt}\}_{t=0}^{\infty}$ is a dynamic politico-economic equilibrium starting in the oligarchy if

- 1. The sequence of the political systems $\{M_t\}_{t=0}^{\infty}$ can be partitioned into two subsequences by an integer T^D : $M_{t < T^D} = O$ and $M_{t \ge T^D} = D$.
- 2. Firms maximize their profits taking prices as given, implying that prices satisfy equations 1 to 4.
- 3. The entrepreneur chooses $\{K_{Pt}, a_{pt}, c_{pt}\}_{t=0}^{\infty}$ to maximize her lifetime utility described

by equation 15, subjects to constraints described by equations 6 and 16.

- 4. The elite chooses T^{D} , $\{K_{St}, w_{St}, T_{t}, \eta_{t}\}_{t=0}^{T^{D}-1}$, and $\{c_{et}, a_{et}\}_{t=0}^{\infty}$ to maximize its lifetime utility, subject to its budget constraint described by 18 and 19, the politico-economic constraints 7 and 8, and the incentive constraints that firms and the entrepreneur optimize.
- 5. The market clears, as described by equation 5.
- 6. Aggregate variables after democratization are given by the competitive equilibrium in the democracy. In particular, $\{y_{et}, \eta_t, r_{St}\}_{t=T^D}^{\infty}$ are determined by equations 12, 13, and 14.

Notice that to focus on the equilibrium in the oligarchy, I make two simplifications in the above definition. First, in item 6, I describe only the equilibrium conditions in the democracy that are different from those in the oligarchy, but not the same ones. For example, the firms' maximization conditions and the entrepreneur's maximization problem are the same in both regimes, thought the prices are different. Second, some intermediate variables are omitted in the definition of the equilibrium because they can be expressed by the variables. For example, workers' income in democracy y_{wt}^D is expressed as a function of capital, as in equation 11, and I substitute this expression into the elite's problem, as in equation 17. Then we don't have to include y_{wt}^D into the definition of the equilibrium. Similarly, the government budget constraint is not in the definition of the equilibrium because it is used to obtain the expression for y_{et} , so it is automatically satisfied.

The reason that we can simplify the definition and focus on the equilibrium in the oligarchy is that except y_{et}^D , other variables in the democracy do not matter for the equilibrium in the oligarchy, and y_{et}^D has a very simple expression. For example, $\{\eta_t, r_{Pt}\}_{t=T^D}^{\infty}$, are useful for defining the entrepreneur's problem, but they do not really matter for the entrepreneur's choices in the oligarchy, because as we will see later in the solution of the entrepreneur's problem, the entrepreneur simply saves a constant fraction of her total

resources, independent of future rate of return on her savings.

2.7 Analytical Properties of the Equilibrium

Before solving for the equilibrium using numerical methods, we can first analytically characterize two important properties of the equilibrium on the allocations and the wedges in the labor market and in the capital market.

Let us first focus on how S sector labor is determined, given S sector and P sector capital. The S sector wage determines the marginal productivity of labor in the S firm, and thereby the labor demand of the S firm conditional on capital, as captured by equation 1. Moreover, given the constraint that S workers' incomes need to be high enough, described by equation 7, S workers' wages are then determined by their expected incomes in the democracy and transfers, if the constraint is binding. In this case, we can express S sector labor as

$$L_{St} = ((1 - \alpha) z_{St})^{\frac{1}{\alpha}} w_{St}^{-\frac{1}{\alpha}} K_{St}$$

= $((1 - \alpha) z_{St})^{\frac{1}{\alpha}} (y_{wt}^D - T_t)^{-\frac{1}{\alpha}} K_{St}.$ (21)

This equation shows that a higher level of transfers leads to a lower wage, and then higher labor demand of the S firm. Intuitively, when S workers are directly paid with large transfers, the labor cost for the S firm is low; so, the S firm is willing to hire many workers, thereby generating redundant labor and low labor productivity in the S sector. This result can be formally stated as the following:

Proposition 2. If $T_t \ge \frac{\tau \alpha}{1-\alpha} w_{St}$, then $L_{St} \ge L_{St}^D$, and $\frac{Y_{St}}{L_{St}} \le \frac{Y_{Pt}}{L_{Pt}}$.

If transfers are larger than a certain fraction of the wage, then there is redundant labor in the S sector. The size of S sector labor is larger than the efficient size L_{St}^D given the S sector and the P sector capital, and labor productivity in the S sector is lower than the counterpart in the P sector.

The efficient labor allocation given the capital allocation is determined by the competi-

tive equilibrium and is the same as in the democracy, denoted as $L_{St}^D = z_{St}^{1/\alpha} K_{St} / (z_{St}^{1/\alpha} K_{St} + z_{Pt}^{1/\alpha} K_{Pt})$. The proof of the proposition is in the Online Appendix.¹⁴ This result states that if the equilibrium transfers are large enough, the labor wedge - the lower labor productivity in the S sector than that in the P sector - exists. The government chooses transfers T_t given the cost parameter b_t . If b_t is low, transfers are large, and S sector labor productivity is low. Then as we will see in the calibration of the model, we can choose b_t to match the labor wedge.

 L_{St} also depends on K_{St} , as we see from equation 21. A larger K_{St} implies a larger L_{St} , and makes it easier to satisfy the large S sector labor constraint represented by equation 8. Conditional on T_t , a large enough K_S is necessary for keeping L_{St} greater than \underline{L} . This result can be formally stated as follows.

Proposition 3. Given any $\mu > 0$, if $T_t \le \mu w_{St}$, then the two constraints $L_{St} \ge \underline{L}$ and $T_t + w_{St} \ge y_{wt}^D$ together imply that $K_{St} \ge \kappa K_{Pt}$, where $\kappa = \frac{\left(1 + \frac{\tau \alpha}{1-\alpha}\right)^{\frac{1}{\alpha}}L}{\left(1 + \mu\right)^{\frac{1}{\alpha}} - \left(1 + \frac{\tau \alpha}{1-\alpha}\right)^{\frac{1}{\alpha}}L} \left(\frac{z_{Pt}}{z_{St}}\right)^{\frac{1}{\alpha}}$.

If transfers are bounded above, then large enough state sector capital relative to private sector capital is necessary for sustaining the oligarchy.

The detail of the proof is in the Online Appendix. The intuition is the following: If transfers as a percentage of wages are below a certain upper bound, to provide high enough incomes for S workers, S sector wages cannot be too low. Given S sector wages and the corresponding marginal productivity of labor, keeping enough workers in the S sector implies investing enough capital in the S sector, as we see from equation 1. This minimal S sector capital requirement may result in over-investment in the S sector, and a capital wedge between the S and the P firms. Notice that this result is quite different from the results implied by a competitive equilibrium. In the competitive equilibrium, more P sector capital implies less S sector capital, as described by Song et al. (2011),

¹⁴The reason for obtaining this simple expression for the lower bound of T_t is that a worker's expected income in the democracy is also a function of the S sector and the P sector capital. It includes the wage in democracy and the tax revenue, which equals $\frac{\tau \alpha}{1-\alpha}$ fraction of the wage. So if the size of transfers in the oligarchy is larger than that in the democracy, the S sector wage is lower than the efficient wage, resulting in the redundant labor and lower labor productivity.

because with more capital, the P firm hires more workers, and then the S firm hires fewer workers and needs less capital. In this model, differently, when P sector capital grows, the political constraint implies that S sector capital grows proportionally to guarantee enough well-paid workers in the S sector. Moreover, over-investment in the S sector can result in a capital market wedge: lower S sector capital productivity than the counterpart in the P sector.

The above two propositions together imply that if transfers are bounded above and below, then the political constraint generates labor and capital wedges in the S and the P sectors. As we will see in the quantitative exercise studied below, this is indeed the case: Equilibrium transfers vary in a bounded region, and there are labor and capital wedges.

3 Quantitative Analysis

In this section I calibrate the model to the Chinese economy and show that the quantitative model can deliver a successful account of China's growth experience since the 1990s, including the private sector growth, the labor and capital market frictions, and the economic growth. More specifically, as the benchmark, I calibrate the model to the manufacturing sector in China since 1998, and I also calibrate the model to urban China since 1994, including both the manufacturing and the service sectors in an extension. The main reason for using the manufacturing sector as the benchmark is the data availability: Some necessary statistics and datasets for the calibration, e.g., labor and capital productivity gaps between state firms and private firms, are available in the manufacturing sector and are accurately estimated and well-understood in the literature, while they are often not available in the service sector. To study the service sector and the urban economy, it is useful to first calibrate the model to the manufacturing sector and back out those crucial moments and parameters, e.g., the labor market wedge and the cost of transfers, and then under the assumption that these moments and parameters can be reasonably

generalized to the urban economy, the model can be used to study the urban economy.¹⁵ The calibration to the urban economy will be discussed in Subsection 5.5.

The model does not apply to rural China, because the government uses different strategies there: Instead of creating a large state sector and getting support from state workers by providing them high incomes, the government use other policies such as propaganda to maintain political support from the rural population as a group, because of different politico-economic conditions in the rural area. For example, low education of the rural population makes political propaganda very effective in getting support from the rural people and low in cost. Therefore, there is no need to maintain a large fraction of rural workers in the state sector at a high cost. ¹⁶

The exogenous inputs of the model include the respective TFP growth in the S and the P sectors, the upper bound of the borrowing constraint, and the world interest rate. Other major parameters are endogenously backed out in the calibration: For example, the discount factor of the entrepreneur is chosen to match the growth of the private employment share, and the cost of transfers to match the labor productivity of state firms relative to that of private firms. The calibrated model generates untargeted moments, including the output growth rate since 1998 and invest share of the state sector, which can be used to check the validity of the model. Moreover, the model provides predictions of future growth, market frictions and development paths.

¹⁵For example, the NBS data of manufacturing firms allow Hsieh and Song (2015) to back out the labor and capital market wedges in the state and the private sectors, while this dataset is not available for the service sector.

¹⁶Through the lens of this model, the high efficiency of propaganda in the rural area can be interpreted as the very low cost of *T*: the government can easily keep the whole rural population's expected gains from democratization low. More generally speaking, propaganda and political control are very effective and result in a low expected return from democratization and a high cost of democratization, because of low education, low population density, and the weak civil society in the rural area. Though the rural population are quite poor and should expect higher incomes in the democracy, the demand for democracy in the rural population is quite low, or at least, there are not many strong challenges to the regime.

Para.	Value	Description	Source
α	0.5	Capital share	Bai et al. (2006)
δ	0.1	Depreciation rate	Song et al. (2011)
r	0.05	World interest rate	Song et al. (2011)
$\bar{\eta}$	1.24	Max. leverage of the P firm	Li et al. (2008)
η	1	Min. leverage of the P firm	
$\overline{g_{z_P}}$	3%	Growth of the P firm TFP	Brandt and Zhu (2010)
z_S/z_P	0.45, 0.58	TFP ratio between the S and P firm	Hsieh and Song (2015)

 Table 1: Exogenous Parameters

3.1 Calibration

The model is calibrated to the Chinese manufacturing sector since 1998, following Song et al. (2011). Table 1 summarizes the parameters that are obtained exogenously from the literature. The production function is Cobb–Douglas with the capital share $\alpha = 0.5$ (Bai et al. (2006)). Following Song et al. (2011), the annual depreciation rate is set to $\delta = 0.1$ and the world interest rate r = 5%.¹⁷ The leverage of private firms in China in 2002 is 24%, according to Li et al. (2008). We set $\bar{\eta}$ such that the leverage in 2002 generated by the model is consistent with the empirical finding. In the calibrated model, before 2008, the economy is in the rapid growth stage and it is always the case that $\eta_t = \bar{\eta}$. So setting $\bar{\eta} = 1.24$ gives the correct leverage in 2002. ¹⁸ $\underline{\eta}$ is set to 1, implying that in the extreme case, the government can order banks to lend only to state firms but not to any private firms.

The sequences of TFP levels of the S firm and the P firm are also taken exogenously from the literature. The annual growth rate of the P firm TFP between 1998 and 2007 is set to 3% - the average private sector TFP growth during that period estimated by Brandt

¹⁷The world interest rate in Song et al. (2011) corresponds to the rate of return of capital in financially integrated firms when the iceberg cost is zero. Their Figure 8 shows that as the iceberg cost declines, the rate of return converges to about 5%.

¹⁸This result - $\eta_t = \bar{\eta}$ in 2002 - is also robust to different parameters, as long as we target finishing state to private labor reallocation in 2008. Theoretically, it is possible that $\bar{\eta} > 1.24$ and in 2002 the government optimally chooses $\eta_t = 1.24 < \bar{\eta}$, but it does not happen in the calibrated economy because as long as the economy is still in the rapid growth stage in 2002, the government does not want to create additional barriers to slow down P firm growth.



Figure 2: Sequences of Parameters

and Zhu (2010). This TFP growth rate is assumed to continue until 2030, to gradually decline to 2% in 2040, and to stay at 2% after that, as shown in panel 1 of Figure 2. The TFP levels of the S firm relative to the P firm in 1998 and 2007 are backed out using the labor and capital productivity gaps between the state and the private firms estimated by Hsieh and Song (2015). They report in their Table 5 that state firms' labor productivity is 60% of private firms' in 1998 and 75% in 2007. They also report the counterparts for capital productivity: 35% in 1998 and 47% in 2007. In the neoclassical framework, the Cobb–Douglas production function implies that the TFP gap can be backed out by combining the labor and capital productivity gaps, as follows:

$$\frac{z_{St}}{z_{Pt}} = \left(\frac{Y_{St}/L_{St}}{Y_{Pt}/L_{Pt}}\right)^{1-\alpha} \left(\frac{Y_{St}/K_{St}}{Y_{Pt}/K_{Pt}}\right)^{\alpha},$$

where Y_{St}/L_{St} stands for the labor productivity in the S sector, etc. This equation implies that S firm TFP is 45% of P firm TFP in 1998 and 58% in 2007.¹⁹ The TFP gaps between the S and P sectors in 1998 and 2007, discussed above, imply that the S sector TFP grew at 4% between 1998 and 2007. Afterward, the S firm TFP growth rate is assumed to linearly decline until it converges to the P firm TFP growth rate in 2017, implying that in 2017 the S firm TFP is 68% of the P firm TFP and the gap stays at this level afterward, as shown in panel 2 of Figure 2.

¹⁹These numbers are broadly consistent with the estimation by Brandt and Zhu (2010), who also use the neoclassical framework - 45% and 50%, while lower than the TFP gaps estimated using the monopolistic competition framework in Hsieh and Song (2015) - 55% and 75%.

Para.	Value	Description	Target
β	0.825	Discount factor of the entrepreneur	$L_{P,08} - L_{P,98} = 0.45$
\underline{L}	0.385	Min. state employment share	$mean(L_{P,08\sim12}) = 0.615$
τ	0.2	Tax rate	$y_{S}/y_{P} = 1.2$
b_{98}, b_{07}	0.28,0.64	Cost Parameter of transfers	$MPL_{S} / MPL_{P} = 0.6, 0.75$

 Table 2: Endogenous Parameters

The rest of the parameters are endogenously determined in the model to match the corresponding moments in the data, as summarized in Table 2. Storesletten and Zilibotti (2014) document that the private employment share increases from 15% in 1998 to above 60% in 2008 and then stays around 61.5% afterward. The growth speed of the private employment share is matched by setting the discount factor of the entrepreneur to $\beta = 0.825$. The final stable level of the private employment share pins down the minimal state employment share needed for sustaining the regime as $\underline{L} = 38.5\%$. The private employment share in 1998 pins down the initial assets of the entrepreneur. The tax rate is set to $\tau = 20\%$, generating an average of 20% income premium for state workers between 1998 and 2007, which is consistent with Ge and Yang (2014). The cost parameter of transfers - b_t - is chosen to match the labor market wedge: A higher cost parameter implies less transfers, lower state sector wages, and lower state labor productivity. As I discussed above, Hsieh and Song (2015) document that state sector's labor productivity is 60% of the private sector's counterpart in 1998, and 75% in 2007. Setting b to 0.28 in 1998 and 0.64 in 2007 matches the corresponding labor wedges. After 2007, b is assumed to keep growing at the same rate as before - 0.04 per year - until 2040, when all exogenous parameters, including the TFP growth rates, reach their steady-state levels. Panel 3 of Figure 2 shows the sequence of the cost parameter of transfers b. The calibration gives an increasing sequence of b because the labor productivity of state firms increases fast relative to that of private firms, as documented by Hsieh and Song (2015). If instead, b decreases and becomes smaller than in the benchmark, then transfers will be larger and will increase faster than in the benchmark, and consequently, state firms will hire more redundant

labor and the increase of state firms' labor productivity will be slower than the calibration targets. It is reasonable to expect that in reality, the cost of transfers increases over time. The cost of raise funding for cash transfers may not increase after the 1994 SOE reform, as Brandt and Zhu (2000) point out, but transfers in the general sense, including non-cash benefits, subsidies, and political propaganda, become less efficient in increasing workers' real incomes and therefore become more costly over time.²⁰

3.2 The Solution of the Model

First, we solve the representative entrepreneur's problem, and then we can solve for the elite's problem, who takes the optimal choices of the entrepreneur as a given. The solution of the entrepreneur's problem is very simple and is the same as in Moll (2014): Given that the entrepreneur's income is proportional to her assets and the log utility, the entrepreneur saves β fractions of her resources in hand, including assets and income, to the next period and consumes the rest, independent of the rate of return on her assets in each period.

Given the entrepreneur's reactions, we can solve the elite's problem. The elite acts like a social planner but cares only about its own utility; so after solving for this problem, the equilibrium is essentially solved. The elite's problem and the equilibrium in the oligarchy cannot be solved sequentially but have to be solved recursively. The reason is that the elite's decision on whether to democratize is discrete and also depends on the lifetime

²⁰The general interpretation of the variable *T* in the model includes also non-cash benefits, subsidies, and political propaganda, as discussed in Subsection 2.3 and in the Online Appendix. First, subsidies and in-kind wages enjoyed by state workers become less efficient and less popular, as the market develops and supplies more variety of goods and services. For example, clothes, education, and housing provided by the government were very valuable for workers 20 years ago, because these goods and services supplied by the market are usually similar and limited. It is quite different now: Market supplies are abundant, high quality and with more varieties, so subsidies and in-kind wages become less efficient in increasing state workers' real incomes. For example, a state worker may not like the clothes or the apartment provided by the government because she prefers some other types supplied by the market, so she does not find the government subsidies valuable. Brandt and Zhu (2000) document that those non-cash benefits become less important and less popular over time. Second, political propaganda may also become less efficient over time. As people become richer and more educated, their utility in the democracy may increase, and it is also harder to change their expected incomes and utility. Changes in efficiency of these policies similar to transfers can be modeled as the increasing cost parameter of transfers *b*.

utility in the oligarchy; so it cannot be characterized by a first-order condition (FOC).²¹ The optimal choices cannot be characterized by sequences of equations, which allow the problem to be solved sequentially. Instead, I write down the elite's problem recursively and solve it backward. First, I form the elite's problem at the steady state recursively and solve for it using value function iterations. The value function of the elite in the oligarchy gives us the lifetime utility of the elite if it chooses oligarchy, so we can compare the lifetime utility of the elite in the oligarchy with the counterpart in the democracy, to obtain the elite's political choice. Second, knowing the elite's value functions and policy functions at the steady state, say, in period t_{ss} , I can solve for the elite's problem in period t_{ss} – 1. The elite's value functions and policy functions in period t_{ss} – 1 are different from those in the steady state, because the parameters are different. For example, if the S firm TFP is lower than the steady-state level, even if the elite faces the same state variable, i.e., the entrepreneur's assets, the elite may choose a higher S sector capital level in period t_{ss} – 1 than in the steady state. Finally, we iterate this process backward and obtain the elite's value functions and policy functions in periods $t_{ss} - 2$, $t_{ss} - 3$, ..., 1, 0. Now we know the elite's policy functions in period 0, or equivalently, in year 1998, that is to say, we know what the elite would choose given any level of the entrepreneur's assets.²² Because a higher level of the entrepreneur's assets implies a larger private employment share, the initial private employment share from the data helps us to pin down the right initial level of the entrepreneur's assets in the model. The equilibrium starting from that level is the one that represents Chinese economy and the one we focus on. Details of the algorithm

²¹The elite compares the lifetime utility if it chooses to sustain the oligarchy to the lifetime utility in the democracy, and to characterize this decision, we need to know the lifetime utility if it chooses oligarchy. In a standard competitive equilibrium, e.g., the equilibrium in the democracy, choice variables are continuous and optimal choices in each period can be characterized by first-order conditions depending on variables in that period and/or in the next period, e.g., the Euler equation. The optimal choices in a competitive equilibrium can be transformed into sequences of equations, whose solution gives the equilibrium path, while this approach does not work in this problem.

²²There are two state variables in this model, the elite's assets and the entrepreneur's assets. The elite's asset level affects only its consumption and saving, but not other equilibrium variables, e.g., P sector labor, because the elite has deep pockets: no matter how low or high its own asset level is, it can always implement the capital and labor market policies that maximize its lifetime income. So the real relevant state variable for the equilibrium is only the entrepreneur's assets. More details are in the Online Appendix.



Figure 3: The Solution of the Equilibrium at the Steady State

for solving the equilibrium are in the Online Appendix.

In the following, the solution of the calibrated model is explained and the intuition and the logic are discussed. In the steady state, or more precisely, the balanced growth path, many variables, including assets, capital, incomes, and transfers, grow at the same rate as the TFP does - 2%. So here I normalize these variables using the TFP growth rate such that the normalized ones converge to their steady-state values in the long run. Using the normalized variables, Figure 3 shows the solution of the elite's problem at the steady state, i.e., how policies and equilibrium variables depend on the state variable - the entrepreneur's asset a_p . A noticeable and important feature is that the relations between the equilibrium outcomes, for example, the government policies, and the state variable, i.e., the entrepreneur's assets, are not monotonic, which is different from the feature in a standard competitive equilibrium. How these variables co-move with the entrepreneur's assets depends on whether the entrepreneur's assets are small or large. If the entrepreneur's assets are small, corresponding to $a_p < 0.8$ in Figure 3, increasing the entrepreneur's assets leads to increasing P sector capital and decreasing S sector capital and labor. This pattern - the negative relation of the S sector and the P sector sizes - is similar to the pattern in the competitive equilibrium in the democracy. The reason is that when the P sector is small, the political constraint is not binding, and the economy behaves similarly to a competitive equilibrium: As the entrepreneur's assets grow, the P sector grows and employs more workers and the S sector shrinks.

However, if the entrepreneur's asset is large enough, i.e., $a_p \ge 0.8$, the equilibrium pattern differs from that of the previous case. S sector capital and labor no longer decrease as the entrepreneur's assets increase, because now the political constraint becomes binding and relevant. S sector labor stays at the critical level <u>L</u> to keep enough workers in the S sector. Moreover, S workers' incomes and wages have to be high enough; so keeping at least <u>L</u> workers in the S sector implies that S sector capital needs to be large enough compared to P sector capital, as stated in Proposition 3.

Moreover, if the entrepreneur's assets are too large, i.e., $a_p > 1.2$, noticeably, the P sector leverage decreases with the entrepreneur's assets. The elite does not want the P sector capital to be larger anymore and wants to reduce the P sector size, meaning that though a larger P sector contributes more tax revenues, it also makes maintaining enough workers in the S sector more difficult and now the cost of P sector growth exceeds the benefit for the elite. The elite's interests are no longer in line with growth. The elite reduces the lending to the P sector and slows down growth of the P sector and the aggregate economy. It wants to keep the size of the P sector and also the size of the economy at a median level, corresponding to $a_p = 1.2$ in this example. As we can see from panel 3 and panel 4 of Figure 3, the lending to the P sector drops after the entrepreneur's assets reach 1.2, and the elite's lifetime income is maximized at that point. However, there is a limit on how much the elite can slow down P sector growth, i.e., $\eta_t \ge \underline{\eta}$; so after the leverage drops to the minimal level, growth of the entrepreneur's assets can still lead to growth of P sector capital, though not of P sector labor anymore.

Transfers and wages for state workers generally follow the pattern of private sector capital, as shown in panel 5: As private sector capital becomes larger, the government

increases both wages and transfers for state workers, to maintain their incomes high enough. Finally, the entrepreneur's assets tomorrow are a concave function of the assets today, as shown in panel 6. This function intersects with the 45 degree line, and the entrepreneur's assets converge.

Above is the solution at the steady state. In fact, in other periods, given different parameters, the policy functions are qualitatively similar. There are some quantitative differences. One difference is that in the first few periods, the cost parameter *b* is lower than in later periods. So transfers are cheaper and larger, the S sector wage is lower, and labor productivity gaps between the S sector and the P sector are larger. In later periods, costs of transfers become larger, and wages and the labor productivity in the S sector also become higher.

3.3 The Dynamics of The Economy

The initial level of the entrepreneur's assets can be set such that the initial P sector employment share equals the 1998 private sector employment share in the data - 15% (see Storesletten and Zilibotti (2014)). Then policy functions and equilibrium conditions solved above give us sequences of variables from 1998. Now we can see how well the dynamics generated by the model match with the facts it was calibrated to and the historical data it was not calibrated to.

Targeted Moments Figures 4 shows the targeted moments and the corresponding sequences generated by the model from 1998 to 2017. The dashed lines show the data, while the solid lines are the time series generated by the model economy. The targeted moments - the state employment share from 1998 to 2012, labor and capital productivity gaps between state firms and private firms in 1998 and in 2007 - are matched well.²³

²³The dashed line in the first panel of Figure 4 represents the state employment share in each year between 1998 and 2012, documented in Storesletten and Zilibotti (2014). In the second panel, the dashed line connects ratios of state firms' labor productivity to private firms' in 1998 and 2007, calculated by Hsieh and Song (2015), and in the third panel, ratios of capital productivity are shown.



Figure 4: Targeted Moments: 1998 - 2017

First, in the calibration, the initial assets and the time discount factor of entrepreneurs are chosen to match the state employment share in 1998 and its decline from 1998 to 2008, as shown in panel 1 of Figure 4. Notice that the decline of the state employment share speeds up gradually between 1998 and 2004, and the decline slows afterward and eventually stops in 2008. This pattern - initially fast but then slow decline of the state employment share - is also generated by this model: The differences between the state employment shares from 1998 to 2008 in the data and the counterparts in the model are quite small: on average only -0.006. This pattern cannot be generated by a competitive equilibrium model. The competitive equilibrium generates an exponential growth of private sector capital and labor, and the decline of the state employment speeds up over time. In my model, before the state-to-private labor reallocation finishes, the government already wants to slow growth of the private sector, because slowing growth of the private sector reduces future costs of sustaining the regime. Second, choosing the cost parameter of transfers in 1998 and 2007 allows the model to match the labor productivity gaps between state firms and private firms, estimated by Hsieh and Song (2015). As we can see from panel 2 of 4, the labor productivity gaps are matched well: The ratio between the labor productivity of state firms and that of private firms in 1998 and in 2007 generated by the model are also very close to the numbers in the data.²⁴ The matched labor productivity of state firms relative to that of private firms implies that the capital productivity is automatically

²⁴Capital productivity ratios are automatically matched, given that labor productivity ratios are matched and the TFP gaps between state and private firms are backed out using labor and capital productivity ratios.
matched, because the TFP gap between state and private firms is chosen according to the labor and capital productivity gaps, as we can see from panel 3 of Figure 4.

Untargeted Moments The dynamics of some variable before 2008 are targeted in the calibration while the dynamics after 2008 are not, e.g., labor and capital productivity gaps and state employment share. The time series of some other variables from 1998 to 2017 are not targeted, e.g., output growth and invest share of state firms. We can compare these untargeted time series generated by the model with the historical data to discuss the validity of the model. Let us first look at state employment share and the labor and capital productivity after 2008 in Figure 4. Not surprisingly, state employment share stays at around 38.5%, as the data from Storesletten and Zilibotti (2014) suggest, because the stop of privatization is one of the motivating facts of this paper and the model is built to reflect this fact. The labor productivity of state firms keeps converging toward the counterpart of private firms after 2008, but at a slower speed, while state firms' capital productivity stops increasing when it reaches a bit more than 50% of private firms' capital productivity. This feature is qualitatively consistent with the findings in Hsieh and Song (2015) with one difference in capital productivity right after 2008. Hsieh and Song (2015) state that "... average labor productivity of the state-owned firms continued to increase from 2007 to 2012 relative to incumbent private firms, albeit at a lower rate than in the 1998 to 2007 period. ... there is little convergence in capital productivity after 2007." In my model, the labor productivity of state firms keeps increasing after 2008 but at a lower rate because TFP of state firms keeps converging to that of private firms but at a lower speed. The other reason for the increasing labor productivity is that the cost of transfers increases, which implies decreasing transfers and decreasing redundant labor in state firms. The capital productivity stills keeps increasing a bit in the model because TFP growth of state firms is assumed to smoothly decrease, and the model can generate the stop of state firms' capital productivity growth between 2007 and 2012, together with slower growth of their labor



Figure 5: Untargeted Moments: 1998 - 2017

productivity, if we assume that the growth rate of state-firm TFP is discontinuous and drops to about the same rate of private-firm TFP, which is possible while assuming smooth convergence to the steady state is more standard in a model. The qualitative features of the capital productivity growth after 2008 is still consistent with the empirical findings: the capital productivity of state firms relative to that of private firms grows much slower than the labor productivity, and stops when there is still a large gap. The reason is that the cost of transfers increases, implying that redundant labor in state firms decreases, which harms the capital productivity of state firms but helps the labor productivity.

Now we can look into the other variables from 1998 to 2017. The most important one is output growth. The solid line in panel 1 of Figure 5 shows the output growth rate in the model: It stays high - above 10% - before 2008, and decreases a bit to a still high level of above 5% in 2017. The dashed line shows the growth rate of PPP adjusted GDP per capita in China, computed by the World Bank. Here I plot growth of output per capita in the aggregate economy because the counterpart in the manufacturing data used by Hsieh and Song (2015) is not available. Moreover, it is useful to compare the model's implications with the counterparts in the aggregate economy and understand to which the extend the results of the benchmark model can be generalized to the aggregate economy.²⁵ As we can see, both in the model and in the data, output growth is fast before 2008 and slows after

²⁵Ideally, the dashed line should be the growth of output per worker in manufacturing firms, obtained from the dataset used by Hsieh and Song (2015), because the benchmark model is calibrated to the manufacturing sector. However, that dataset is not publicly available. Moreover, though the benchmark model is calibrated to the manufacturing sector, because of better estimated statistics, the mechanism in the model applies to the whole urban China.

2008, but still stays relatively high. The growth rate from 1998 to 2008 in the data is a bit lower than that in the model. This is reasonable, because 1997 Asian financial crisis slowed China's growth, and this business cycle shock is missing in the model. Not surprisingly, my model, like any other neoclassic growth model, generates a non-increasing output growth rate converging to the steady state level and cannot generate an increasing output growth rate unless additional shocks are introduced. Growth after 2008 generated by the model is very close to the counterpart in the data: It steadily decreases from about 13% in 2008 to above 5% in 2017.

In panel 2 of Figure 5, the solid line shows state investment share since 1999 in the model, and the dashed line shows the state investment share in the aggregate economy computed by Knight and Ding (2010) and the dotted line shows that in the urban China computed by Brandt and Zhu (2010). In the model, the investment share of state sector decreases from about 70% in 1999 to about 40% in 2008, and afterward, it stops declining and even increases to about 50% and stays around that level, which is higher than the state employment share. Similarly in the data, the state employment share starts from a high level and declines, especially in Knight and Ding (2010); while differently, the decline documented by Brandt and Zhu (2010) is less dramatic, only from about 60% to 50%. The quantitative difference in investment share reflects the difference in state employment share. Brandt and Zhu (2010) focus on the aggregate economy, construct the state employment share differently, and document that the fast decline of the state employment in the aggregate economy already happened since 1994 and became slower after 1999. The qualitative patterns are the same: Brandt and Zhu (2010) also argue that the state sector's investment share declines as its employment share declines, but the investment share stays higher than the employment share, which is consistent with the implication of the model in this paper. Moreover, the pattern that the state investment share stays high and even slightly increase after the state-private labor reallocations ended in 2008 is consistent with the recent discussion on the phenomenon that "the state

advances as the private sector retreats."

It is also interesting to look at transfers implied by the model. As panel 3 of Figure 5 shows, transfers increase from 34% of state sector wages in 1998 to 45% in 2004 and then gradually decline to 33% in 2017. Transfers are relatively low initially because the government wants to make the private sector grow fast: higher transfers result in more redundant labor in the state sector, less labor in the private sector, and slower private sector growth. As the private sector becomes larger, transfers increase, to balance the cost of maintaining the political support and growth of the private sector. Eventually, transfers decrease, because transfers become more costly as the cost parameter b increases over time. It is difficult to measure the exact counterpart of the variable T in the data, because some policies such as propaganda cannot be measured, but it is useful to compare the variable in the model with the value of subsidies and in-kind wages enjoyed by state workers documented by Brandt and Zhu (2010). They argue that the value of subsidies and in-kind wages is about 25% of cash wages in 2000s. The number in the model is larger, because the variable T represents not only non-cash benefits and subsidies to state workers, but also labor subsidies to state firms and propaganda and other government policies which affect state workers' expected incomes in the oligarchy and in the democracy, so the difference corresponds to the value of policies not captured by the subsidies and in-kind wages enjoyed by state workers. The dynamics of transfers - the initial increase and the later decline - generated by the model is also reasonable. The increase of T in the 2000s imply that as the private sector grows, the government becomes more willing to use transfers and similar policies such as political propaganda to maintain the political support. Cantoni et al. (2017) document that the textbook reform in China since 2004 resulted in positive views of the government, changed views on democracy, and more skepticism toward free markets. Eventually, the decline of transfers will happen if the cost of transfers and similar policies increases. For example, if the cost of political propaganda increases, its effectiveness decreases, and then the government increases the expenditures



Figure 6: Dynamics of The Economy: 1998 – 2040

on propaganda but its effects decline.

Three-Stage Politico-Economic Transition Now we can extend the focus to also include the dynamics in the future and study the whole picture of the transition. Figure 6 extend the sequences of some variables studied above, i.e., state employment share, the labor and capital productivity, and output growth until 2040, while Figure 7 shows the dynamics of more variables in the equilibrium.

From 1998 to 2008, the state employment share declines from 85% to 38.5%. Capital and labor in the private sector grow very rapidly (see Figure 6). This is the first stage of the development - *rapid growth*. In this stage, the main driving force is private sector growth, in terms of both the relative size (employment share) and the absolute size (capital level). The output growth stays above 10%. The growth of the private sector TFP also contributes significantly to output growth. Though the state sector TFP grows faster, the relative size of the state sector declines, so the contribution of the private sector TFP to output growth becomes relatively more important. The labor productivity of state firms relative to that of private firms grows from 60% to 75%, while the capital productivity grows from 35% to 47%, as documented in Hsieh and Song (2015). The labor and capital productivity of state firms grows following the growth of the state sector TFP. In addition



Figure 7: Dynamics of The Economy: More Variables

to TFP growth, which increases both the labor and the capital productivity, transfers also affect the labor and the capital productivity differently: It directly reduces wages and the labor productivity of state firms and indirectly increases the capital productivity through the lower capital-labor ratio. In this stage, transfers as a share of wages increase and then decrease, but roughly stay around a high level - 40%, as we can see from the third panel of Figure 6; so state firms' labor and capital productivity grows as rapidly as state firms' TFP. Why do transfers increase and decrease around 40%? There are two counterforces affecting the dynamics of transfers. First, the cost of transfers increases over time because both the cost parameter and workers' incomes grow, creating incentives for the government to reduce transfers as a percentage of incomes. Second, as we discussed above, given the initially small private sector, the government prefers an initially low level of transfers to help the private sector to grow, because a higher level of transfers implies more redundant labor in the state sector and reduces the private sector's labor and output. As the private sector grows over time, the government has a lower incentive to increase the size of the private sector, so it worries less about redundant labor in the state sector and may want to increase the level of transfers. The two forces together generate a relatively stable level of transfers, going slightly up and then down in this stage. So the dynamics of the labor wedge is mainly determined by the dynamics of TFP gaps but not transfers. So is the capital wedge. In later stages, when transfers as a percentage of wages decline, the labor productivity growth differs from the capital productivity growth.

After 2008, state employment stops declining and stays around 38.5%, implying that

the economy enters a different stage of development: state capitalism. In this stage from 2008 to 2027 - private sector capital keeps growing, but at a slower speed. State sector capital increases proportionally to private sector capital to maintain state sector employment. Private sector's capital growth slows because its employment share does not grow anymore and wages are pushed up by the large investment in the state sector. Private sector growth is still an important direct driving force of output growth, but now the private sector's share in the economy does not grow anymore. Private sector capital still keep growing, while state sector capital grows proportionally, and the overinvestment in the state sector keeps the aggregate economy's growth rate still high. From 2008 to 2017, output growth declines but stays above a reasonably high level of 5%. Another important feature in this stage is that the growth of labor productivity in the state sector differs from the growth of capital productivity, because transfers as a share of incomes decrease when the cost of transfers grows as both the cost parameter *b* and incomes grow over time. When transfers from the government are lower, wages paid by state firms are higher, and state firms hire less redundant labor. The declining transfers and the growing TFP together generate the growth of labor productivity in the state sector. The reduction of redundant labor in the state sector implies a higher capital-labor ratio and the lower capital productivity. The over-investment in the state sector is the other reason for low capital productivity. As we can see from Figure 6, the model generates a continuous growth of state firms' labor productivity. In comparison, state firms' capital productivity does not increase much and stays around 50% of private firms' capital productivity. This prediction is qualitatively consistent with the finding in Hsieh and Song (2015). They show that from 2007 to 2012, the labor productivity of state firms relative to that of private firms keeps growing but at a slower rate than before, while the capital productivity stops increasing.²⁶

²⁶There are some qualitative differences from the findings in Hsieh and Song (2015): Growth of labor and capital productivity after 2007 generated by my model is faster than that documented by Hsieh and Song (2015). The main reason is that TFP growth after 2007 used in my paper is faster. The quite slow growth of productivity after 2007 found by Hsieh and Song (2015) implies a large drop of the TFP growth rate after

As the entrepreneurs' assets and private sector capital keep growing, taxes from the private sector can become lower than the cost of maintaining the regime. Then the elite chooses to directly slow the growth of private sector capital by reducing lending to private firms. In the model, this reduction happens in 2028, and then private sector capital stops growing. The economy enters the third stage: *the middle-income trap*. The credit constraint on private firms becomes even tighter and private firms find it even harder to grow. The long-run private sector output is lower than the efficient level in the democracy, for two reasons. First, private firms' capital and loans are reduced by the government. Second, private sector employment is lower than in the democracy because of the political constraint. Eventually, output growth slows and then stays at the steady-state level - 2%, before it reaches the level in the democracy.

To sum up, the model generates a three-stage transition. The first two stages are consistent with China's growth experience since 1998. The model accounts for the rapid privatization that started in 1998 and stopped in 2008. It can also account for the dynamics of the labor and the capital market frictions: for the declining labor productivity gap between state firms and private firms, and for a larger capital productivity gap which initially declined but recently stopped decreasing. It also generates reasonable output growth until today. The model offers the following explanation for the dynamics of frictions and the growth since 1998: the political constraint - buying support from state sector workers using wages and transfers - creates redundant labor and over-investment in the state sector, but the private sector still grows and contributes to the rapid growth. Finally, the model predicts that frictions will persist and that private sector capital growth will slow. The rapid growth in the current regime is not sustainable and the long-run output is lower than that in the democracy.

²⁰⁰⁷ to about one fourth of the rate before 2007. In my paper, the state sector TFP growth is assumed to decline smoothly, so the changes in state firms' productivity growth are not that dramatic.

3.4 Alternative Development Path: Sustained Growth

In the benchmark calibration, capital market frictions persist, because they help the elite to sustain the regime and maximize its income. These frictions harm growth of the private sector and the aggregate economy. The alternative case is that the elite chooses to undertake political reforms, including liberalizing the financial market and democratization, and then private firms can enjoy better access to the financial market and grow faster. Political reforms do not happen in the benchmark model because the cost of sustaining the regime is low: The government can invest in the state sector at a low cost because it can borrow from the international market at a low interest rate, and it can use transfers to buy support also at a low cost. If the costs are large enough, the elite chooses to democratize. For example, if the interest rate r is high enough or the elite cannot borrow from the international market, then investing in the state sector is more costly and the elite may stop supporting the oligarchic regime. If the cost of transfers is large enough, democratization may also happen.²⁷

To study the consequences of democratization, here I consider the case that the government chooses to democratize and remove all labor and capital market frictions instead of choosing to restrict private sector growth when the economy enters the third stage.²⁸ In the benchmark model, the economy enters the third stage in 2028, so Figure 8 shows the dynamics if democratization happens in 2028 in solid lines, and in comparison, the dynamics in the benchmark model in dashed lines. As we can see, after democratization, labor and capital market frictions disappear. First, the leverage of private firms stays at the highest level. Second, the state sector's labor productivity immediately becomes as

²⁷As we discussed above, transfers in the model can be more generally interpreted as all policies that increase workers' expected incomes in the oligarchy relative to the counterparts in the democracy, including political propaganda and control. If economic development increases education, demand for equal rights, and better organized civil society, then it becomes increasingly costly for the government to use transfers and similar policies to buy support for the oligarchic regime. Eventually if the cost grows to a large enough level, democratization happens. This possibility is what modernization theory (see Lipset (1959)) focuses on.

²⁸It might choose to do so if the cost of transfers becomes large enough, e.g., the cost becomes $\underline{b} + b_t T_t^2$, and \underline{b} is sufficiently large.



Figure 8: Dynamics After Democratization

high as the private sector's labor productivity. Third, the employment share of the state sector declines again. As we can see from panels 4 and 5, private sector capital grows faster than in the benchmark, and state sector capital initially increases but then quickly decreases to 0. The private sector grows faster because now lending to private firms and private employment are not restricted by the government. State sector capital initially increases because the capital level is now determined by the competitive market, not by the government which is the monopolizer of the state capital supply and prefers to set state sector capital to a lower level. Later, state sector capital decreases as private sector capital grows. As shown in panel 6, output after democratization grows faster than in the benchmark, especially in two eras: Right after democratization and right after all state firms exit the market, which happens around 2042. In the first era, growth rate increases because democratization removes restrictions on private firms and they grow faster than before. In the second period, growth speeds up again because state sector capital declines to zero and there is no further decline of state sector capital, and therefore, total capital in the economy grows faster. Afterward, growth is driven by private sector capital accumulation. Eventually, output in the democracy converges to a higher level

than that in the oligarchy because of the removal of capital and labor market frictions.

3.5 Calibration to China's Urban Economy

The mechanism in the theory also applies to China's urban economy including both the manufacturing and the service sectors. Because of the data availability, in the benchmark discussed above, the model is first calibrated to manufacturing. Now given the statistics and parameters backed out from the benchmark calibration, I can extend the model to study the urban economy since 1994 under the assumption that the missing statistics and parameters in the urban economy are in the same pattern of those in manufacturing.

The time series of state employment share in the urban economy is taken from Song et al. (2011). The pattern of the state employment share is similar to that in manufacturing, but with some quantitative difference: The decline started earlier than the counterpart in manufacturing, became slower than the latter after 2001, and stayed parallel to the latter after 2004. The labor and capital productivity gaps between state and private firms after 1998 are assumed to be the same as the counterparts in the manufacturing sector, estimated by Hsieh and Song (2015). TFP series of state firms and private firms before 1998 are obtained by extrapolating the series after 1998. In the calibration, the endogenous parameters - \underline{L} , β , b_{98} , b_{ss} - are reset to match the changes in the above targeted moments, and other parameters are kept the same as those in the benchmark.

The model calibrated to the urban economy gives a qualitatively similar three-stage transition, and the prediction for future political development is the same: the third stage of the transition is the middle-income trap, as shown in Figure 9 and discussed in more details in the Online Appendix. The state-to-private transition in the urban economy is quantitatively different from that in manufacturing. The most important difference is the slower decline of the state employment share after 2001 than that in the manufacturing sector, suggesting that the state-to-private transition is slower in the service sector. The slower transition is generated by changing the exogenous parameters, e.g., the discount



Figure 9: The Model Calibrated to the Urban Economy: 1994 – 2040

factor of entrepreneurs, and also by the endogenously changed policies, e.g., the reduction of lending to private firms, which starts earlier in this calibration than in the benchmark. The leverage of private firms decreases in 2005 and triggers a decline of output. In reality, the decline of leverage less dramatic and so is the decline of output. Another possible reason for the slower transition in the urban economy is that the TFP gap between state and private firms is larger than that in manufacturing, and the gap declines slower. In the calibration to manufacturing, the gap is assumed to keep declining after 2008, and this increase of state firms' productivity does not increase the state employment share, because the growth of the private sector is fast. However, in the calibration to the urban economy, given the slower growth of the private sector, the increase of state firms' productivity implies that the state employment share may increase a bit after 2008, though eventually it will decline to the lower bound. This pattern - slight increase of state employment share after it has stabled around the lower bound - can be seen in panel 1 of Figure 9, but it is only because the TFP growth of state firms after 2008, borrowed from the benchmark calibration, is too fast. If the TFP growth of state firms after 2008 is slower, which is the likely case in the urban economy, the state employment share should stay at the lower bound after 2008. To minimize this unnecessary pattern due to the assumption on the

missing statistics, I adjust the cost of transfers after 2008 up slightly, which reduces the state employment share, with the tradeoff that the state firms' labor productivity misses the target slightly. The calibration to the urban economy can be done better after further research provides the necessary statistics and data moments.

3.6 Discussion

In this subsection, I discuss in greater detail the model's mechanics, extensions, and implications. I first discuss the core mechanism that generates the rapid growth by analyzing efficiency and redistribution in the oligarchy. Then I discuss an extension of the model which assumes that the economy is closed. Finally, I study the implications of the model for the middle-income trap and for middle-class activism.

Efficiency vs. Redistribution In this model, growth in the oligarchy is faster in the beginning but then slower in the long run, whereas in the democracy it is slower in the beginning but faster in the long run. This result and the intuition behind it are similar to those in Acemoglu (2008): In the beginning, the elite's interests are in line with growth and government policies are growth-enhancing, but in the long run, the elite's interests conflict with growth and policies are set to restrict growth. However, the mechanics and the implications for efficiency and redistribution are different. In Acemoglu (2008), an oligarchic society produces more than a democratic society does in the beginning because it achieves higher efficiency given that the elite - the major producers - can protect its property better than it can in a democracy. Differently, in this paper, initial growth is rapid because of inefficiencies - labor and capital market distortions. First, the economy starts from a low output level resulting from labor market distortion, so it is not difficult to generate rapid output growth as long as the distortion declines. In this sense, output efficiency in an oligarchy is lower, though growth is faster. Figure 10 compares the dynamics in a democracy, starting from the first period (blue lines), with



Figure 10: Dynamics in Democracy

the dynamics in the benchmark economy in an oligarchy (red dashed lines). We can see that labor productivity of state firms always equals that of private firms (panel 1). This higher efficiency in the democracy implies that the initial output level is higher than in the oligarchy (panel 5), though output growth is lower (panel 6 shows that growth in the democracy is around 10% from 1998 to 2008 and then declines to about 4% in 2017). Second, output growth in the state capitalism stage is also rapid, driven by inefficient over-investment in the state sector. In this stage, the output level in the oligarchy can catch up with the output level in the democracy. We can see from Figure 10 that state sector capital grows fast in the oligarchy and it catches up with the level in the democracy around 2028 (panel 4), while the output level in the oligarchy catches up with the level in the democracy and gets very close in 2028 (panel 5).²⁹ However, this high output is mainly due to the very high level of state sector capital, but not high productivity of capital

²⁹In alternative calibrations, output in the oligarchy can be larger than the counterpart in the democracy in the state capitalism stage, if state sector TFP is high enough. Also notice that the capital level is initially lower in the oligarchy than that in the democracy. There are two reasons for the low level of state sector capital in the beginning. First, the government is the monopolizer of state sector capital; so it prefers to set state sector capital at the low monopolistic level to obtain a higher return to capital. Second, inefficient labor allocation implies a lower return to capital; so there is a lower incentive for investment. However, when the private sector's employment share reaches the critical level, the government must set state sector capital high enough to maintain enough supporters. Afterward, state sector capital in the oligarchy exceeds the level in the democracy.

and labor. If we measure efficiency by output, we can state that efficiency in the oligarchy can be close to or even higher than in the democracy, because the institutions and policies in the oligarchy are very efficient in generating fast capital accumulation. However, if we measure efficiency by the productivity of resources, for example, the capital, then efficiency in the oligarchy is still lower.

When we focus on efficiency in achieving a high social welfare, how aggregate output is distributed matters.³⁰ The large capital in the state sector is intended to sustain the regime, so it generates a high income for the elite. However, private sector workers' incomes are higher in the democracy; and state sector workers expect at least the the same level of incomes in the democracy. The population size of workers is much larger than that of the elite class, so the aggregate welfare in the oligarchy is lower because of lower incomes for the majority. In other words, welfare and efficiency in the oligarchy are lower than in the democracy because of the larger inequality.

The Agricultural Sector In the benchmark and the extension, the model is used to study the manufacturing sector and the urban economy which includes also the service sector. The rural economy is not in the model because the main mechanism - buying support from state sector workers by providing them high incomes - does not apply to the rural area, where different policies, e.g., propaganda, are use to maintain the support among the rural population as a group. The presence of the agricultural sector matters for the model because the structural transformation from the agriculture sector to the manufacturing and the service sectors increases the workforce in the latter two sectors. In the model, this can be captured by introducing population growth of workers. Suppose that population of workers increases at the rate g_w , then as in a neoclassic growth model with population growth, we can normalize the variables, e.g., use capital per capita

³⁰An economy can obtain a high output level but a low social welfare level. Consider, for example, a neoclassic growth model with a positive capital subsidy. The output level is higher than the counterpart in the competitive equilibrium without the subsidy but the competitive equilibrium without distortions achieve the highest social welfare.

instead of capital, and focus on the steady state of per capita variables. The population growth is similar to the productivity growth, except that in the representative household's problem, the time discount factor also needs to be adjusted accordingly. In my model, the normalization is even simpler, because the workers are myopic and there is no need to set up workers' problems or adjust their discount factor. Then the agriculture to manufacturing transformation and the growth of workers' population are equivalent to adding the growth rate of labor to the productivity growth rate, i.e., $g = \hat{g} + g_w$, where \hat{g} is the original TFP growth rate. We can still solve the model like we did above, and the interpretations of g and employment are slightly different. First, decline of the state employment can be also because of the increase of the labor force, and decline of the state employment can be much slower than decline of the state employment share. Second, the decline of state employment share implies that a larger fraction of increment labor is allocated to the private sector. For example, if the growth rate of labor exactly equals the decline of the state employment share in a certain year, then the state employment stays the same and all new workers work in the private sector.

The Closed Economy If we assume that the elite cannot borrow from the international market and can use only domestic savings to invest in the state sector, then the cost of sustaining the regime changes and the elite may want to influence the domestic savings. Let us first consider the case in which the elite does not manipulate domestic savings and the corresponding interest rate. Then the elite has to use its own savings to finance state sector investment, and the elite's assets becomes a state variable of the equilibrium. As entrepreneurs' assets grow and the economy enters the state capitalism stage, the elite must increase its savings to maintain enough capital in the state sector. If the time discount factor of the elite is small, the elite's savings can be low in the open economy, but in the closed economy, the elite is forced to save more and consume less. The cost of maintaining the oligarchy increases. Two possible outcomes may emerge: One is that the elite finds it

too costly to sustain the oligarchy and chooses to democratize; the other is that the elite chooses to sustain the oligarchy and adopts even harsher policies to restrict private sector growth. The two outcomes have an interesting implication: International sanctions on non-democratic countries, which forbid the governments from accessing the international financial market, may create two opposite results. Some governments stop maintaining the regime, for example, the Soviet Union. In some other countries, for example Iran, the governments respond by tightening their control over their economies and then the societies become even less democratic. This outcome is the opposite of what the sanctions aim to achieve. The different outcomes of sanctions are consistent with the empirical findings in Grauvogel and von Soest (2013) and von Soest and Wahman (2015).

If we consider the case in which the government can influence domestic savings and the interest rate, then the government wants to increase savings in the domestic banking system and also wants to reduce the corresponding interest rate, which is the cost of using the savings to sustain the regime. Compulsory saving, for example through the pension system, can help to increase control over the politico-economic system, while regulated low interest rates for household savings, can reduce the cost of maintaining the regime. If we consider a half-closed economy where the population can have access to the international financial market, under the regulation of the government, then capital control and the recent tightening of capital outflows in China are the policies which help the government to control more financial resources and to maintain the regime.

The Middle-Income Trap Government policies in the oligarchy, endogenously generated by the model, result in rapid growth in the beginning, but then become detrimental to growth when the elite's interests conflict with private sector growth. The model explains why some governments adopt the right policies to achieve rapid growth out of the poverty but then suddenly implement the "wrong" policies that stop the convergence to rich countries: Those policies are actually not wrong for maximizing the elite's own income,

but just wrong for growth. In the real world, the negative impacts of government policies can be more detrimental than what was described above in my model. In the benchmark calibration, the restriction that the government can put on private firms is moderate: At the maximal, it can set lending to private firms at zero, thus reducing their capital by 24%. The long-run output level is lower than but not that far from the level in the democracy. In reality, the government may implement much harsher policies to restrict private firms and these policies may be much more harmful for growth. The government can directly confiscate capital of some private firms. The government can also forbid private firms from entering or staying in certain industries, which they may have invested in before. If the government implements these two policies, there are two consequences. First, capital of private firms can be even lower than entrepreneurs' assets, if private firms are not allowed to increase their investments in certain industries or if capital of some private firms is confiscated by the government. Second, profits and the productivity of private firms can decline when private firms are excluded from some profitable industries. These policies can be modeled as allowing the government to set $\eta_t < 1$ or to directly reduce z_{Pt} .³¹ If these two modelling setups are allowed, the long-run output in the oligarchy can be even lower than that in the benchmark. If I extend the benchmark model to allow the government to set $\eta_t < 1$, η_t in the long run is set to 0.9 by the government, and this lower η_t in turn reduces the long-run output by another 10%. In the real world, many governments can be much more afraid of private sector growth than the government in my model. For example, if the productivity of state firms or elite-controlled firms in an economy is lower than that in my model, then the government implements harsher policies in controlling private firms, and the long-run output becomes even lower. Such an economy can stop

³¹In my model, the implications of these two policies are similar, for the following reason: the P firm output is $z_{Pt} (\eta_t a_{pt})^{\alpha} L_{Pt}^{1-\alpha} = (z_{Pt}^{\frac{1}{\alpha}} \eta_t a_{pt})^{\alpha} L_{Pt}^{1-\alpha}$, and decreasing z_{Pt} by α % is equivalent to decreasing η by 1% in the impacts on output. They are different only in their respective impacts on the cost of the P firm, because reducing η reduces the borrowing of the P firm and the related cost, while reducing z_{Pt} does not. However, given the relatively low interest rate for borrowing compared to the high rate of return to P firms' capital, this difference is small.

growing when its income is still far away from the income of the United States, and may stay as a lower middle-income economy for very long.

Middle-class Activism An important political implication of this model is that the middle class do not necessarily support democracy. What we learn from European history and also from modernization theory is that as the middle class receive higher incomes and better education, so their demand for democracy grows over time and they are more likely to support democratization. A similar mechanism is described in this paper: As the middle class' incomes increase, the cost of buying support rises, putting more pressure on the oligarchic government. However, this paper shows that the middle class' demand for democracy may not be higher than the lower class, and the demand may not grow to a high enough level that results in democratization, because of two counterforces. First, the government can reduce the demand for democracy from the middle class by providing them well-paid state sector jobs and associated benefits. Then the middle class, being afraid of losing their decent incomes and benefits, become less supportive of democratization than the lower class are. In fact, a large fraction of the middle class is created by the state. The creation of the middle class relying on the state is especially common in countries with historically large state sectors. For example, in a former communist country, many people rely on the state but not the market to receive incomes, goods, and services including education, health, and housing, so it is natural and cheap for the government to create or maintain a large fraction of the middle class relying on the state. As discussed above, China is an example. Rosenfeld (2017) finds a similar pattern for the Russian middle class. He reveals that the middle class from the state sector in Russia are less likely to mobilize against electoral fraud, and argues that potential coalitions in support of democratization are weakened by middle-class growth in state-dependent sectors. Second, even if the demand for democracy by the middle class is large and grows over time, it is not always the case that this demand will be sufficient to pressure the non-democratic government to

democratize. This paper shows the possibility that the government, for a very long period at least, can maintain the non-democratic regime, as long as it is willing to pay the cost, even though the middle class' incomes and demand for democracy have been growing.

4 Conclusion

This paper proposes a political-economic theory to understand China's recent growth experience and to predict developments in China's economic and political transition in the future. The political constraint - maintaining enough supporters in the state sector - creates labor and capital market frictions and results in a three-stage transition. The first two stages are *rapid growth* and *state capitalism*, which are consistent with several salient aspects of China's recent development, including the growth of the private sector, persistent labor and capital market wedges, and rapid output growth. In the future, China is likely to enter a *middle-income trap* with persistent capital market frictions, given the economically and politically powerful state. To switch to the development path that leads to *sustained growth*, economic and political reforms are necessary, though such reforms may endanger the current regime.

Although this paper focuses on China, it is also useful for understanding the development of many other emerging countries and some developed countries having similar patterns. The political constraint in the theory also exists in some other countries such as Kuwait, where the political elite need to gain support from public workers, and Korea before 1980, when politicians needed support from workers in industries tightly connected to the government. In Korea before 1980, large conglomerates (chaebol) were granted privileged access to low-cost credit and the employment share of small and medium enterprises (SMEs) had stagnated. The difference is that after 1980, democratic changes and financial reforms happened together, and the employment share of SMEs increased from 50% to 68% in 1990 and continued in the early 1990s after democracy was consolidated.³²

³²See Song et al. (2011) for more details. They also discuss the case of Taiwan and the drop of SMEs'

This political and economic development path is consistent with the second case in the theory: the case of *sustained growth*.

Moreover, the theory is also useful in pondering an important question in development: Should other developing countries apply the "China model," i.e., the combination of authoritarian politics and state-guided capitalism, to promote economic growth? Some in favor of adopting this model cite its recent success, but the long-run implications should be carefully examined and distinguished from the short-run implications. This paper provides a quantitative framework to evaluate the economic and political consequences of the "China model," which may become the "Korea model" under the right conditions.

References

- Acemoglu, Daron, "Oligarchic Versus Democratic Societies," *Journal of European Economic Association*, 2008, 6 (March), 1–44.
- **Bai, Chong-En Chongen, Chang tai Hsieh, and Yingyi Qian**, "The Return to Capital in China," *Brookings Papers on Economic Activity*, 2006, 37 (2), 61–102.
- **Besley, Timothy and Torsten Persson**, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, aug 2009, *99* (4), 1218–1244.
- Brandt, Loren and Xiaodong Zhu, "Redistribution in a Decentralized Economy growth and Inflation in China under Reform," *Journal of Political Economy*, apr 2000, 108 (2), 422.
- $_$ and $_$, "Accounting for China's Growth," 2010, (4764).
- _ , Trevor Tombe, and Xiaodong Zhu, "Factor market distortions across time, space and sectors in China," *Review of Economic Dynamics*, jan 2012, 16 (1), 39–58.

employment share before 1971 and the increase until 1991.

- Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman, Y Jane Zhang, Leonardo Bursztyn, Thomas Dee, Stefano Dellavigna, Nicola Gennaioli, Paola Giuliano, Torsten Persson, Nancy Qian, Mark Rosenzweig, Scott Rozelle, Monika Schnitzer, Andrei Shleifer, Shing-Yi Wang, Linxiu Zhang, and David Zweig, "Curriculum and Ideology," Journal of Political Economy, 2017, 125 (2).
- Chen, Jie and Chunlong Lu, "Democratization and the Middle Class in China: The Middle Class's Attitudes toward Democracy," *Political Research Quarterly*, aug 2011, 64 (3), 705–719.
- **Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski**, "The Economy of People's Republic of China from 1953," jul 2015.
- Chow, Gregory C., China's Economic Transformation, 3 ed., Wiley-Blackwell, 2015.
- **Ge, Suqin and Dennis Tao Yang**, "Changes in China's Wage Structure," *Journal of the European Economic Association*, apr 2014, 12 (2), 300–336.
- **Grauvogel, Julia and Christian von Soest**, "Claims to Legitimacy Matter: Why Sanctions Fail to Instigate Democratization in Authoritarian Regimes," 2013, (235), 40.
- **Hsieh, Changtai Chang-Tai and Zheng Song**, "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China," 2015.
- Knight, John and Sai Ding, "Why Does China Invest So Much?," *Asian Economic Papers*, 2010, 9 (3).
- Li, Hongbin, Lingsheng Meng, Qian Wang, and Li An Zhou, "Political connections, financing and firm performance: Evidence from Chinese private firms," *Journal of Development Economics*, 2008, 87 (2), 283–299.
- Lipset, Seymour Martin, "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *American Political Science Review*, 1959, 53 (1), 69–105.

- **Moll, Benjamin**, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 2014, pp. 3186–3221.
- **Morris, Stephen and Hyun Song Shin**, "Unique Equilibrium in a Model of Self-Fufilling Currency Attacks," *American Economic Review*, 2000, *90* (1), 316–318.
- **Rosenfeld, Bryn**, "Reevaluating the Middle-Class Protest Paradigm: A Case-Control Study of Democratic Protest Coalitions in Russia," *American Political Science Review*, 2017, pp. 1–16.
- Shirk, Susan L., The Political Logic of Economic Reform, 1 ed., University of California Press, 1993.
- Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing Like China," American Economic Review, 2011, 101 (February), 196–233.
- **Storesletten, Kjetil and Fabrizio Zilibotti**, "China's Great Convergence and Beyond," *Annual Review of Economics*, aug 2014, 6 (1), 333–362.
- **von Soest, Christian and Michael Wahman**, "Not All Dictators Are Equal: Coups, Fraudulent Elections, and the Selective Targeting of Democratic Sanctions," *Journal of Peace Research*, 2015, 52 (1), 17–31.
- **Zilibotti, Fabrizio**, "Growing and Slowing Down Like China," *Journal of the European Economic Association*, 2017, 15 (October), 0–46.

5 Appendix (Available Online)

In the Online Appendix, I first provide proofs of the propositions. Then Subsection 5.3 presents more details of the model, including the equilibrium in the democracy. How to solve the equilibrium in the oligarchy, including how to solve the elite's problem, is discussed in Subsection 5.4, in two steps: First, the elite's problem is transformed into the recursive form; Second, the algorithm for solving the recursive problem is presented. The solution to the elite's problem essentially gives us the solution of the equilibrium. In Subsection 5.5, an extension of the benchmark model to study the whole urban China is discussed. Finally, I discuss more details of the model, including the micro-foundation of the political constraint, the general interpretation of transfers, and the decomposition of the elite's problem.

5.1 **Proof of Proposition 2**

First, workers' expected incomes in the democracy y_{wt}^D can be written as a function of capital allocations K_{St} and K_{Pt} , combining equations 10 and 11:

$$y_{wt}^{D} = \left(1 + \frac{\tau \alpha}{1 - \alpha}\right) (1 - \alpha) \left(z_{St} K_{St} + z_{Pt} K_{Pt}\right)^{\alpha}$$

So the S sector labor in the oligarchy L_{St} can be expressed as a function of y_{wt}^D , T_t , K_{St} and K_{Pt} , while L_{St}^D as a function of K_{St} and K_{Pt} :

$$L_{St} = ((1-\alpha)z_{St})^{\frac{1}{\alpha}} w_{St}^{-\frac{1}{\alpha}} K_{St}$$

= $((1-\alpha)z_{St})^{\frac{1}{\alpha}} (y_{wt}^D - T_t)^{-\frac{1}{\alpha}} K_{St}$
 $L_{St}^D = \frac{z_{St}^{\frac{1}{\alpha}} K_{St}}{z_{St}^{\frac{1}{\alpha}} K_{St} + z_{Pt}^{\frac{1}{\alpha}} K_{Pt}}.$

Finally, we can compare L_{St} and L_{St}^D and check for which levels of the transfer, we get $L_{St} \ge L_{St}^D$. Let us define $d_t \doteq \frac{T_t}{w_{St}}$, then we have:

$$(1-\alpha)^{\frac{1}{\alpha}} \left(\frac{1}{1+d_t} \left(1 + \frac{\tau\alpha}{1-\alpha} \right) (1-\alpha) \right)^{-\frac{1}{\alpha}} \frac{z_{St}^{\frac{1}{\alpha}} K_{St}}{z_{St}^{\frac{1}{\alpha}} K_{St} + z_{Pt}^{\frac{1}{\alpha}} K_{Pt}} \ge \frac{z_{St}^{\frac{1}{\alpha}} K_{St}}{z_{St}^{\frac{1}{\alpha}} K_{St} + z_{Pt}^{\frac{1}{\alpha}} K_{Pt}},$$
$$\left(\frac{1}{1+d_t} \left(1 + \frac{\tau\alpha}{1-\alpha} \right) \right)^{-\frac{1}{\alpha}} \ge 1,$$
$$1+d_t \ge 1 + \frac{\tau\alpha}{1-\alpha},$$
$$d_t \ge \frac{\tau\alpha}{1-\alpha}.$$

So if the transfer is large enough, i.e., $\frac{T_t}{w_{St}} \ge \frac{\tau \alpha}{1-\alpha}$, then there is redundant labor in the S sector compared to the efficient labor allocation in the democracy: $L_{St} \ge L_{St}^D$, given the capital allocations K_{St} and K_{Pt} . Moreover, we know that the labor allocation in the democracy is determined by the competitive equilibrium and efficient, so the labor productivity of the S and P sector in the democracy are the same: $\frac{Y_{St}^D}{L_{St}^D} = \frac{Y_{Pt}^D}{L_{Pt}^D}$. Then $L_{St} \ge L_{St}^D$ implies a labor productivity gap between the S and the P sector: $\frac{Y_{St}}{L_{St}} \ge \frac{Y_{Pt}^D}{L_{Pt}^D} = \frac{Y_{Pt}}{L_{Pt}^D} \ge \frac{Y_{Pt}}{L_{Pt}}$, given the decreasing productivity of labor.

5.2 **Proof of Proposition 3**

Given μ as an upper bound of T_t/w_{St} , the constraint $T_t + w_{St} \ge y_{wt}^D$ implies that w_{St} can not be too low:

$$w_{St} \ge \frac{y_{wt}^D}{1+\mu}.$$

Then given the not too low w_{St} , the enough state labor constraint $L_{St} \ge \underline{L}$ implies that the state sector capital K_{St} cannot be too low:

$$\begin{split} \left(\frac{1}{1+\mu}\left(1+\frac{\tau\alpha}{1-\alpha}\right)\right)^{-\frac{1}{\alpha}} \frac{z_{St}^{\frac{1}{\alpha}}K_{St}}{z_{St}^{\frac{1}{\alpha}}K_{St}+z_{Pt}^{\frac{1}{\alpha}}K_{Pt}} \ge L, \\ \frac{K_{St}}{K_{St}+(z_{Pt}/z_{St})^{\frac{1}{\alpha}}K_{Pt}} \ge \left(\frac{1}{1+\mu}\left(1+\frac{\tau\alpha}{1-\alpha}\right)\right)^{\frac{1}{\alpha}}L, \\ \frac{K_{St}}{(z_{Pt}/z_{St})^{\frac{1}{\alpha}}K_{Pt}} \ge \frac{\left(\frac{1}{1+\mu}\left(1+\frac{\tau\alpha}{1-\alpha}\right)\right)^{\frac{1}{\alpha}}L}{1-\left(\frac{1}{1+\mu}\left(1+\frac{\tau\alpha}{1-\alpha}\right)\right)^{\frac{1}{\alpha}}L}, \\ K_{St} \ge \frac{\left(1+\frac{\tau\alpha}{1-\alpha}\right)^{\frac{1}{\alpha}}L}{(1+\mu)^{\frac{1}{\alpha}}-\left(1+\frac{\tau\alpha}{1-\alpha}\right)^{\frac{1}{\alpha}}L}\left(\frac{z_{Pt}}{z_{St}}\right)^{\frac{1}{\alpha}}K_{Pt}, \\ \doteq \kappa K_{Pt}. \end{split}$$

More precisely, the state sector capital cannot be too low compared to the private sector capital, if the political constraint, or equivalently, the two economic constraints, are to be satisfied.

5.3 More Details on The Democracy

The equilibrium in the democracy is a decentralized competitive equilibrium given taxes, similar to Song et al. (2011). The competitive labor market implies that wages are the same in the S and the P sectors. The competitive capital market and the elite's infinity borrowing capacity imply that the return of S sector capital to the elite is equal to the interest rate r if the S sector exist. In this case, the raw rate of return to S sector capital, i.e., the marginal productivity of S sector capital, is simply $r/(1-\tau)$, where τ is the tax rate on the elite's income. This helps to pin down the S sector capital labor ratio. The higher productivity of the P firm imply that the capital return in the P sector is higher than in the S sector, while the credit constraint of the entrepreneur implies that the entrepreneur's capital supply is limited by the credit constraint. Given a large enough time preference parameter β and

the corresponding high enough saving rate, the entrepreneur keeps accumulating asset and the P sector keeps growing until all workers move to the P sector while S firms no longer produce. In this case, the elite does not supply capital to the S sector anymore but simply saves its asset in the bank and get the return at the rate r. More details are discussed below.

First, in the democracy, if the S firm still exists, then the rate of return of S firm capital to the elite has to be r. If it is greater than r, each elite member wants to supply infinite capital and it is not an equilibrium; if it is lower than r, the elite does not want to get any loan or supply any capital to S sector, but prefers to put all its asset in the bank and get the return at rate r. In other words, competition of S sector capital supply implies that the net rate of return to the elite, denoted as ρ_{et}^D , equals the marginal cost:

$$\rho_{et}^D = (1-\tau) \alpha z_{St} K_{St}^{\alpha-1} \left(L_{St}^D \right)^{1-\alpha} - \delta = r.$$

This determines *S* sector capital labor ratio and the wage:

$$\begin{aligned} \frac{K_{St}}{L_{St}^{D}} &= \left(\frac{r+\delta}{(1-\tau)\alpha z_{st}}\right)^{\frac{1}{\alpha-1}} \Rightarrow \\ w_{t}^{D} &= (1-\alpha)z_{St} \left(\frac{K_{St}}{L_{St}^{D}}\right)^{\alpha} = (1-\alpha)z_{St}^{\frac{1}{1-\alpha}} \left(\frac{(1-\tau)\alpha}{r+\delta}\right)^{\frac{\alpha}{1-\alpha}}. \end{aligned}$$

The wage pins down the private sector capital labor ratio, the labor given capital and the rate of return to capital:

$$\frac{K_{Pt}}{L_{Pt}^{D}} = \left(\frac{w_{t}^{D}}{(1-\alpha)z_{Pt}}\right)^{\frac{1}{\alpha}} \Rightarrow$$

$$L_{Pt}^{D} = \left(\frac{w_{t}^{D}}{(1-\alpha)z_{Pt}}\right)^{-\frac{1}{\alpha}} K_{Pt},$$
(22)

$$r_{Pt}^{D} = z_{Pt}^{\frac{1}{\alpha}} \alpha \left(\frac{w_{t}^{D}}{1 - \alpha} \right)^{\frac{\alpha - 1}{\alpha}}.$$
(23)

The elite in the democracy receives no transfers from the government, so it only replies on the return to the asset at the rate *r*, and its income from other sources is simply 0.

The entrepreneur maximizes her lifetime utility by optimally choose every period capital supply and saving, taking rates of returns to P sector capital and the borrowing limits as given, as follows:

$$\max_{\{K_{Pt}, a_{pt+1}, j_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} \log c_{pt}$$

s.t. $K_{Pt} \le \eta_{t} a_{pt}$,
 $y_{pt} = \rho_{pt}^{D} K_{Pt} - r K_{Pt}$
 $a_{pt+1} = Ra_{pt} + y_{pt} - c_{pt}$

where ρ_{pt}^D is the net rate of return to capital for the entrepreneur, which is the rate of return to P sector capital to the P firm minus the tax on the entrepreneur, i.e., $\rho_{pt}^D = (1 - \tau) r_{pt}^D$. The entrepreneur's problem in the oligarchy is in fact the same problem, except that ρ_{pt}^D is generally different from the level in the democracy. The reason is that in both regimes, the representative entrepreneur takes prices as given. Moll (2014) solves a similar entrepreneur's problem. Because the utility function is logarithm and the income is proportional to the assets, the solution of the entrepreneur's lifetime problem is quite simple: the entrepreneur maximizes each period's income and saves a constant fraction of the income to the next period. In the following we formally prove that the solution described above is optimal, or the readers can look into Moll (2014), who solves the problem in the recursive form.

First, we show that the entrepreneur maximizes income in every period. Suppose that the sequence $\{a_{pt}^*, K_{Pt}^*\}_{t=0}^{\infty}$ is the optimal solution to the sequential problem. Then $\forall t, K_{Pt}^*$ must maximize y_{pt} , given a_{pt}^* . Otherwise, $\exists \hat{K}_{Pt}$ such that $\rho_{pt}^D \hat{K}_{Pt} - r \hat{K}_{Pt} > \rho_{pt'} K_{Pt'}^* - r K_{Pt'}^*$. Then we can simply construct a new sequence of $\{a_{pt}, K_{Pt}\}$ by replacing K_{Pt}^* by \hat{K}_{Pt} while keeping all other $K_{Pt'}^*$ for all $t' \neq t$ and all other $a_{pt'+1}^*$. The new sequence is feasible and implies $c_{pt} > c_{pt}^*$ and $\forall t' \neq t$, $c_{pt'} = c_{pt'}^*$. The lifetime utility of the new sequence is higher. By contradiction, we know that K_{Pt}^* must maximize y_{pt} . The logic behind this proof is the following: the current period capital decision only has a direct effect on current period income, but no direct effect on future variables, so there is no dynamic tradeoff in choosing K_{Pt} and no reason for not maximizing y_{pt} . If the return to capital for the entrepreneur is higher than the cost, i.e., $\rho_{pt}^D > r$, it is optimal for the entrepreneur to borrow as much as possible and set K_{Pt} as high as possible. However, the choice of K_{Pt} is subject to the borrowing constraint, so the optimal K_{Pt}^* to maximize y_{pt} given a_{pt}^* is simple:

$$K_{Pt}^{*} \begin{cases} = \eta_{t} a_{pt}^{*} & \text{if } \rho_{pt}^{D} > r, \\ \in \left[0, \eta_{t} a_{pt}^{*}\right] & \text{if } \rho_{pt}^{D} = r, \\ = 0 & \text{if } \rho_{pt}^{D} < r. \end{cases}$$

In the first case, which happens when a_{pt} is small and r_{Pt} and ρ_{pt}^D are large, $K_{Pt}^* = \eta_t a_{pt}^*$ and $y_{pt}^* = (\rho_{pt}^D - r)\eta a_{pt}^*$. Adding ra_{pt} , we get the total income, which all come from the return on assets: $y_{pt}^{tot} = Ra_{pt}^* + y_{pt}^* = (R + (\rho_{pt}^D - r)\eta_t)a_{pt}^*$. So $\rho_{pt}^{tot,D} = R + (\rho_{pt}^D - r)\eta_t$ can be considered as the total rate of return to the entrepreneur's asset and the asset return is the entrepreneur's only income source. In the other two cases, where $\rho_{pt}^D = r$ or $\rho_{pt}^D < r$, the total return to the entrepreneur's asset is simply r. Finally, given the return to the asset, the entrepreneur optimally chooses the asset, i.e., a_{pt+1} , taking the total return to asset $\rho_{pt}^{tot,D}$ as given:

$$\max_{\{a_{pt+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c_{pt}$$

s.t. $a_{pt+1} = \rho_{pt}^{tot,D} a_{pt} - c_{pt}$

Given the log-utility, the substitution effect of the return to saving exactly cancels the wealth effect, so in each period, the agent saves β fraction of the total resource to the next

period, i.e., $a_{pt+1}^* = \beta \rho_{pt}^{tot,D} a_{pt}^*$, no matter how high or low $\rho_{pt+1}^{tot,D}$ is.

Given the entrepreneur's optimal choices on asset accumulation, the dynamics in the democracy is the following, similar to Song et al. (2011). Starting from a low a_{pt} and related high r_{Pt} and $\rho_{pt}^{tot,D}$, we have $\beta \rho_{pt}^{tot,D} > 1$, so a_{pt} and K_{Pt} keep growing. Meanwhile, L_{Pt}^{D} keeps growing according to (22), and L_{St}^{D} declines gradually. When a_{pt} is large enough $-a_{pt} \ge 1/\bar{\eta} \left(w_{t}^{D} / ((1-\alpha) z_{Pt}) \right)^{\frac{1}{\alpha}}$, L_{Pt}^{D} reaches 1, and the *S* sector vanishes. Afterwards, the entrepreneur keeps accumulating asset, and the economy behaves like a neoclassic growth model, until $\beta \rho_{pt}^{tot,D} = 1$ and the economy reaches the steady state.

5.4 Solving the Equilibrium in the Oligarchy

First, the entrepreneur's problem in the oligarchy is basically the same as the entrepreneur's problem in the democracy, as she takes prices, i.e., the returns to P sector capital and her asset, and policies - the borrowing limit - as given. Though the returns are different in the democracy and in the oligarchy, the basic structure of the entrepreneur's problem and the solution is the same: the entrepreneur borrows as much as possible subject to the borrowing constraint if the return to P sector capital is higher than *r* to maximize her income in each period, and then sames β fraction of her total resource in hand - asset and its return - to the next period.

Second, the elite takes the entrepreneur's optimal choices into account and solves the problem described by equation 15 and the associated constraints. The elite's problem is essentially a planner's problem, though the planner only cares about her own utility. In a standard neoclassic growth model, the planner's problem can be solved sequentially by transforming the maximization of the lifetime utility into sequences of first order conditions in which the planner's optimal choice in certain period can be characterized by the relation of variables in nearby periods. For example, the Euler equation characterizes the optimal choices of saving by the relation of consumptions in consecutive two periods. However, in this model, this approach does not work, because the elite's choice of the polit-

ical system is discrete, and it depends on the lifetime utility in different political systems. So we need to know the elite's lifetime utility in the oligarchy and in the democracy to characterize its political decision, and the problem can be solve backwards and recursively.

In the following, we write down the recursive form of the elite's problem and then we can discuss the algorithm for solving the recursive problem. Generally speaking, the value functions and policy functions in the recursive problem can depend on time, because parameters, e.g., TFP levels of the S firm and the P firm, vary over time. So we need to keep *t* as the subscript for the value functions and policy functions. Once all parameters reach their steady state levels, the value functions and policy functions do not depend on time anymore. In this case, we can simply drop the *t* subscript. Let us first write down the elite's problem when all parameters are at the steady state levels. If the elite chooses oligarchy, the lifetime utility is denoted as the value function $W^O(a_e, a_p)$; if democracy, $W^D(a_e, a_p)$; and before making the political choice, $W(a_e, a_p)$. First, the Bellman equation given that the elite chooses oligarchy is:

$$W^{O}(a_{e}, a_{p}) = \max_{K_{S}, w_{S}, T, \eta, c_{e}, a_{e}} \log c_{e} + \beta W(a_{e}', a_{p}'), \qquad (24)$$

s.t. $c_{e} = Ra_{e} + y_{e} - a_{e}',$
 $y_{e} = \alpha z_{S} K_{S}^{\alpha} L_{S}^{1-\alpha} - rK_{S} + \tau z_{P} K_{P}^{\alpha} L_{S}^{1-\alpha} - bT^{2} L_{S},$
 $w_{S} + T \ge y_{w}^{D} = (1 - \alpha + \tau \alpha) \left(z_{S}^{\frac{1}{\alpha}} K_{S} + z_{P}^{\frac{1}{\alpha}} K_{P} \right)^{\alpha},$
 $L_{S} \ge \underline{L},$
 $K_{P} = \eta a_{p},$
 $a_{p}' = \beta \left(ra_{p} + (1 - \tau) r_{P} \left(K_{P} - a_{p} \right) \right),$

where $L_S = (1 - \alpha) z_S^{\frac{1}{\alpha}} (w_S)^{-\frac{1}{\alpha}} K_S$, $L_P = 1 - L_S$, and $r_P = (1 - \tau) \alpha K_P^{\alpha - 1} L_P^{1 - \alpha}$. Notice that I do not include the expressions of L_S , L_P and r_P , or equivalently, the firms' maximization problems and market clearing conditions, into the constraints but treat them as known

functions of other variables, to simplify the expressions of the elite's problem. I also simply write the entrepreneur's capital decision as $K_P = \eta a_p$. Precisely speaking, one can include the complete expression of K_P depending on whether r_P is greater or smaller than r. In the calibrated model, it is always the case that $r_P > r$ in the oligarchy, so it is without loss of generality to simplify the expression of K_P here. Second, the lifetime utility if the elite chooses to democratize - $W^D(a_e, a_p)$ - is given by the equilibrium in the democracy and can be taken as given here. ³³ Finally, we can write the elite's political decision as:

$$W(a_e, a_p) = \max\left\{W^O(a'_e, a'_p), W^D(a'_e, a'_p)\right\}.$$

34

There are two state variables in the elite's problem, the elite's asset a_e and the entrepreneur's asset a_p . The entrepreneur's asset matters for the elite because it affects the P sector capital and labor, and eventually the elite's policies and the equilibrium outcomes. The elite's assets matter for its consumption sequence for sure. Does it also matter for the policies and the equilibrium outcomes other than the elite's consumption sequence? The answer is no, because the elite has deep pockets. Consider two cases with the same level of the entrepreneur's asset, but in the first one the elite has a high level of assets, and in the second case a low level of assets. In the second case the elite can copy the policies in the first case and achieve the same level of lifetime income, and achieve a sequence of consumption that is lower than the consumption sequence of in the first case reflecting only the difference in the initial assets of the elite. There is no reason for the elite in the second

³³Later, when we transform the elite's problem from lifetime utility maximization to its equivalent form of lifetime income maximization, we can see that the lifetime income of the elite in democracy is in fact very simple and we even do not need to compute the elite's lifetime income or utility in the democracy.

³⁴There is in fact a third choice for the elite, i.e., staying in the oligarchy while not respecting the political constraint. In this case, revolution happens and the regime switches to the democracy. Because the revolution happens after the capital allocation is determined by the elite, while democratization allows the capital to be determined by a competitive equilibrium, the equilibrium outcomes and the elite's income in the case of revolution can be different from the voluntary democratization. However, in the calibrated model, the lifetime utility of the elite in the revolution is always dominated by democratization or sustained oligarchy, we can simplify from this case here. In the quantitative model, I do keep track of this case.

case to obtain a lower lifetime income than the first elite. Formally speaking, the elite's problem can be decomposed into two subproblems: (1) conditional on a_p , maximization of the lifetime income; and (2) given a_e , maximization of the lifetime utility. The solution from the decomposed problems gives the same solution of the combined problem. I leave the formal proof to Subsection 5.8. The elite's lifetime income maximization problem, again, when all parameters are at their steady state levels, can be written as:

$$V(a_{p}) = \max \left\{ V^{O}(a_{p}), V^{D}(a_{p}) \right\},$$

$$V^{O}(a_{p}) = \max_{\eta, K_{S}, w_{S}, T} y_{e} + \frac{1}{R} V(a'_{p})$$
s.t. $y_{e} = \alpha z_{S} K^{\alpha}_{S} L^{1-\alpha}_{S} - rK_{S} + \tau z_{P} K^{\alpha}_{P} L^{1-\alpha}_{S} - bT^{2} L_{S},$

$$w_{S} + T \ge y^{D}_{w} = (1 - \alpha + \tau \alpha) \left(z^{\frac{1}{\alpha}}_{S} K_{S} + z^{\frac{1}{\alpha}}_{P} K_{P} \right)^{\alpha},$$

$$L_{S} \ge \underline{L},$$

$$a'_{p} = \beta \left(ra_{p} + (1 - \tau) r_{P} \left(K_{P} - a_{p} \right) \right),$$
(25)

where *V* is the value function representing the discounted lifetime income of the elite excluding the return to its initial asset Ra_e , and V^O and V^D stand for the value functions when the elite chooses sustaining oligarchy and democratization. In fact, $V^D = 0$ because the elite has no other income except the return from its asset at the rate *r*, as we discussed in the solution in the democracy. Given the maximized lifetime income, we can recover the elite's lifetime utility by solving the second subproblem:

$$W(a_e, a_p) = \max_{\{c_{et}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c_{et}$$
s.t.
$$\sum_{t=0}^{\infty} \frac{c_{et}}{R^t} = Ra_e + V(a_p).$$
(26)

where $V(a_p)$ is the maximized lifetime income calculated from the first subproblem.

For periods before parameters reach their steady state values, the logic is the same and

the only difference is that we need to keep subscript t for time-varying parameters, value functions and policy functions. For example, we need to write the elite's lifetime income maximization problem as

$$V_{t}(a_{p}) = \max \{V_{t}^{O}(a_{p}), V^{D} = 0\},$$

$$V_{t}^{O}(a_{p}) = \max_{\eta, K_{S}, w_{S}, T} y_{et} + \frac{1}{R} V_{t+1}(a'_{p})$$
s.t. $y_{et} = \alpha z_{St} K_{S}^{\alpha} L_{S}^{1-\alpha} - rK_{S} + \tau z_{Pt} K_{P}^{\alpha} L_{S}^{1-\alpha} - b_{t} T^{2} L_{S},$

$$w_{S} + T \ge y_{wt}^{D} = (1 - \alpha + \tau \alpha) \left(z_{St}^{\frac{1}{\alpha}} K_{S} + z_{Pt}^{\frac{1}{\alpha}} K_{P} \right)^{\alpha},$$

$$L_{S} \ge \underline{L},$$

$$a'_{p} = \beta \left(ra_{p} + (1 - \tau) r_{Pt} \left(K_{P} - a_{p} \right) \right).$$
(27)

As we can see, though some variables are not time varying, e.g., V^D , some other parameters such as z_{St} , z_{Pt} , and b_t are time varying, and consequentially, related variables y_{et} and r_{Pt} depend also on time. We keep the subscript t for these variables and the value functions because they can be different in different periods even given the same state variable. For example, in an early period where the S firm TFP relative to P firm TFP is lower than the counterpart in the steady state, even given the same level of entrepreneur's asset, so the elite may need to invest more in the S sector to sustain the oligarchy, and the elite's lifetime utility can be different.

Given the recursive formation of the elite's problem, we can use value function iterations from backwards to numerically solve the elite's problem and eventually the equilibrium in the oligarchy. First, we solve for the elite's problem when all parameters have reached the steady state values, i.e., the value functions V^O and V and the Bellman equation 25. This recursive problem is solved using the standard value function iteration method with grid search. Starting from an initial guess of V^O and V, we iterate using the Bellman equation until the value functions converge. All parameters converge to

the steady state in period *T*, or year 2040 in the calibrated model, so we know the value functions and policy functions for $t \ge T$. The second step is to solve for the value functions in period t = T - 1 to t = 0 from backwards. Since we know $V_T(a_p) = V(a_p)$, we can solve for $V_{T-1}^O(a_p)$ from equation 27, and also $V_{T-1}(a_p)$. Then we can continue on backward induction until we get $V_0^O(a_p)$ and the associated policy functions in all periods. Finally, now that we know what the elite would choose given any level of a_p in period 0, we can simulate the economy starting from any a_{p0} . The policy functions in period 0 gives us corresponding K_{S0} , w_{S0} and η_0 , which allows us to compute L_{P0} and a_{p1} . Then the policy functions in period 1 allows us to obtain variables in that period and also a_{p2} . We can continue this process until all variables converge. Then an equilibrium path given a a_{p0} is obtained. There is one level of a_{p0} that gives us $L_{P0} = 15\%$, corresponding to the initial private sector employment share in 1998, which is the starting year of the benchmark model. Starting from this level of a_{p0} , we can compute the equilibrium path that correspond to our benchmark economy.

Notice that we even do not need to compute the elite's consumption and saving to back out other equilibrium variables that we care about, e.g., L_{Pt} , η_t , etc. Again, this is because that the elite has a deep pocket and its policies can be independent of its asset. Still, if one is interested, one can back out the elite's consumption and saving by solving the second subproblem in equation 26 and associated constraints.

5.5 More Details on the Calibration to China's Urban Economy

The model is calibrated to China's urban economy since 1994. The state employment share in the urban economy is documented by Song et al. (2011). The decline of the state employment share is in a similar pattern as that in the manufacturing sector, with some quantitative differences: It started earlier, and became slower than that in the manufacturing sector after 2001 and parallel to the latter after 2004. I extend the state employment share in the urban economy to 2012, assuming that it stays parallel to that in



Figure 11: Sequences of Parameters in The Calibration to the Urban Economy

manufacturing which is available until 2012 and documented in Storesletten and Zilibotti (2014). Since the state employment share in manufacturing does not change much after 2008, the extended state employment share in the urban economy also does not change much after 2008. This assumption used to extend the state employment share data is essentially assuming that the state employment share in the urban economy stays constant after 2008, which is reasonable and consistent with the policies and discussions of "advance of the state and retreat of the private". The labor and capital productivity gaps between state and private firms after 1998 are assumed to be the same as the counterparts in the manufacturing sector, estimated by Hsieh and Song (2015). I extrapolate TFP growth of private firms and state firms to the years before 1998. Now given all the exogenous variables from 1994, we can set the endogenous parameters - L, β, b_{98}, b_{ss} - to match the changes in the above targeted moments. The time series of inputs are shown in Figure 11. I keep other parameters the same as those in the benchmark. Results from the newly calibrated model are similar to those from the benchmark: There is a three-stage transition, whose third stage is *middle-income trap*, as shown in Figure 9.

5.6 General Interpretation of Transfers

In the model, variable *T* is referred as transfers while it can actually represent a broader set of government policies that increase state workers' incomes in the oligarchy relative to their incomes in the democracy but not the wages. These policies can be classified as
three groups: cash compensation, non-cash benefits, and policies that change workers' expectations. Here we formally illustrate how these policies affect the equilibrium and why they are equivalent to transfers.

The first set of policies includes cash compensation to state workers paid by the government, including not only direct cash transfers but also subsidies paid by the government to state firms. The former is the narrow interpretation of variable T and is denoted as T^{1c} . The latter is denoted as T^{1s} . The second set includes non-cash benefits, including housing, education and tax benefits that state sector workers enjoy. We can model them as increasing state workers' final income by T^2 . Third, political propaganda can change workers' expected income in the oligarchy if they support the regime and their income in the democracy. Let us denote the policies that increase workers' expected income in the oligarchy if they support the regime, e.g., the reward of loyalty, as T^{3+} , policies that decrease their expected income if they do not support the regime, e.g., cost of revolution, as T^{3-} , and policies that decrease their expected income in democracy, e.g., cost of transition to democracy, as T^D . Then a state sector worker supports the regime if

$$w_{S} + T^{1c} + T^{1s} + T^{2} + T^{3+} \ge y_{w}^{D} - T^{3-} - T^{D}.$$
(28)

We can then denote $T = T^{1c} + T^{1s} + T^2 + T^{3+} + T^{3-} + T^D$ and rearrange the constraint as

$$w_S + T \ge y_w^D$$

Factors that do not affect the income but the expected utility can also be represented by T, if their impacts on utility are additive to income. This can include other non-monetary benefits of democracy such as more civil rights, for example.

5.7 Microfoundation of the Political Constraint and the Equilibrium in Oligarchy

In the main text, to keep the model simple and to focus on the quantitative exercise, the political constraint - enough supporters from the state sector - is given as a feature of the oligarchic regime. It can be rigorously modeled as a political game of all workers and the elite. Below, I extend the model to incorporate the microfoundation of the political constraint.

The government faces the political constraint: it needs support from sufficiently many workers to keep the regime stable. In each period, each worker *i*, decides whether to support the oligarchic regime ($m_i = 1$) or not ($m_i = 0$). The aggregate mass of supportive workers is $M_w = \int_0^1 m_i di$. If it is larger than a crucial threshold \underline{L} , the regime survives this period, otherwise democratization occurs.³⁵

The political decision of a worker is made after she gets employed - either by the S firm or the P firm - and before she receives her wage and final income. The expected final income of a worker from sector $j \in \{S, P\}$ in the oligarchy is denoted as y_{wj} , and after democratization y_{wj}^D . They are endogenously determined by economic factors, which will be explained later in this subsection. ³⁶ So the payoffs can be summarized as in Table 3. Obviously, for a myopic worker *i* in sector *S*, one (weakly) dominating pure strategy is to support oligarchy if and only if the expected income is higher than that in the democracy, i.e., $m_i = 1$ if $y_{wS} \ge y_{wS}^D$ and $m_i = 0$ if $y_{wS} < y_S^D$. Same for a P sector worker. This strategy expresses the voters' sincere preferences. Without loss of generality, I assume that workers

³⁵The setting that only the support from workers counts is without loss of generality, for two reasons. First, the population size of the workers are much larger than the other two groups, so they should count. Second, the elite and the entrepreneur's welfare is generally higher in oligarchy in all most all cases, so the model is robust to whether considering their political support or not.

In the extreme case that the elite holds dominating political power, and needs little support from workers - e.g., it uses mostly military force to control the citizens - \underline{L} can be close to 0. In the other case that the elite needs to win a majority voting, \underline{L} can be 50% if the voting system is fair, or smaller than 50% if the voting system is manipulated in favor of the elite. A regime with a voting system may not necessarily be a democracy. It can still be an oligarchy, and the government serves the interests of the elite.

³⁶Workers are ex-ante identical, so there is no need for subscript i to denote worker i.

	$M_w \ge \underline{L}$	$M_w < \underline{L}$
j = S	y_{wS}	y_{wS}^D
j = P	y_{wP}	y_{wP}^D

Table 3: Payoffs of Workers

use this strategy, similar to the sincere voting assumption in the literature. ³⁷This political game is a simple global game of regime switching, in the spirit of Morris and Shin (2000). The game here is simple because there is no heterogeneous information, no cost of being against the regime, and no punishment for the supporters of the regime after the regime collapses.

5.8 Decomposition of the Elite's Problem

In Subsection 5.4, we claim that the elite's lifetime utility maximization problem can be equivalently written as two subproblems: lifetime income maximization and consumption smoothing, because the elite has a deep pocket and can set policies to maximize its lifetime income independent of its assets. This logic can be easily seen and formally proved in the sequential problem. The proof in the recursive form is also possible but more complicated. Let us prove this claim using the sequential problem first.

Denote the lifetime utility achieved by solving the two sub-problems - first maximizing lifetime income and then maximizing lifetime utility - as U. Remember that the solution to the original one complete problem gives lifetime utility W. First, $U \leq W$. Let us look at the two subproblems. The solution to the first subproblem achieving V can be denoted as $\{\hat{M}_t, \hat{w}_{St}, \hat{K}_{St}, \hat{\eta}_t\}_{t=0}^{\infty}$ in the sequential form. The corresponding consumption and saving decisions obtaining U are denoted as $\{\hat{c}_{et}, \hat{a}_{et+1}\}_{t=0}^{\infty}$. Combined together, the choice $\{\hat{M}_t, \hat{w}_{St}, \hat{K}_{St}, \hat{\eta}_t, \hat{c}_{et}, \hat{a}_{et+1}\}_{t=0}^{\infty}$ achieving U is a feasible choice of the original problem,

³⁷Of course, given that there are a continuum of workers, worker i knows that her action does not affect the aggregate political outcome and feels indifferent about what she does. There are other dominating strategies and equilibria with pure or mixed strategies. However, if there are finite workers and there is some small probability that worker i's choice can be pivotal, then it is wise to follow the sincere strategy described above.

given that in every period the choice set for M_t , w_{St} , K_{St} , η_t is independent of a_{et} . So the optimal solution for the original problem should be at least as good as this candidate choice, i.e., $W \ge U$.

Second, $U \ge W$. Denote the choice that solves the original problem and achieves W with stars, as $\{M_t^*, w_{St}^*, K_{St}^*, \eta_t^*, c_{et}^*, a_{et}^*\}_{t=0}^{\infty}$. Let us compare the lifetime income given $\{M_t^*, w_{St}^*, K_{St}^*, \eta_t^*\}_{t=0}^{\infty}$ with the counterpart given the solution of the first subproblem achieving V: $\{\hat{M}_t, \hat{w}_{St}, \hat{K}_{St}, \hat{\eta}_t\}_{t=0}^{\infty}$. Obviously, the latter is at least as high as the former: $\hat{V} = \sum \hat{y}_{et}/R^t \ge \sum y_{et}^*/R^t = V^*$, given that $\{M_t^*, w_{St}^*, K_{St}^*, \eta_t^*\}_{t=0}^{\infty}$ is also a candidate solution to the first subproblem, again because of the independence of policies from the elite asset. Then, in the second sub-problem, choosing $\hat{c}_{e0} = c_{e0}^* + \hat{V} - V^*$ and $\{\hat{c}_{et}, \hat{a}_{et}\}_{t=1}^{\infty} = \{c_{et}^*, a_{et}^*\}_{t=1}^{\infty}$ gives at least as high lifetime utility as $\{c_{et}^*, a_{et}^*\}_{t=0}^{\infty}$. In other words, if the solution from the subproblem gives higher lifetime income, then consuming the extra lifetime income in the first period and following the same consumption path of the solution to the direct lifetime utility as W.

Combing these two results, we have U = W, i.e., solving the original lifetime utility maximization problem is the same as solving the two sub-problems.