HKUST IEMS
INSTITUTE FOR EMERGING MARKET STUDIES

China's Financial System and Economic Imbalances

Xi LI, Yikai WANG, Tong ZHANG

HKUST IEMS Working Paper No. 2018-53

February 2018

HKUST Institute for Emering Market Studies
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
T: +852 3469 2215 E: iems@ust.hk W: iems.ust.hk

China's Financial System and Economic Imbalances

Xi LI, Yikai WANG, Tong ZHANG
HKUST IEMS Working Paper No. 2018-53
February 2018

Abstract

In this paper, we study how the financial market frictions in the Chinese economy, especially the interest rate policies, lead to inefficient resource allocations and economic imbalances. First, the repressed low interest rate for household savings induce them to increase saving in order to prepare for future necessary expenditures. Consequently consumption share is low and the economic imbalance of consumption and saving emerges. Second, the government provides explicit or implicit guarantees for state firms, so banks prefer to lend to state firms which are less productive. Private firms get less financial resource and operate at sub-optimal levels. The lower aggregate productivity implies the lower household income and consumption and worsen the imbalance. Due to the financial market frictions, traditional consumption stimulating policies, e.g., reducing the interest rate, may actually results in the opposite: even lower consumption and a more imbalanced economy. Reforms towards market-determined interest rates can help to rebalance the economy.

Keywords: China, consumption, economic imbalance, interest rate

Author's contact information

Xi Li

Department of Accounting
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong
E: acli@ust.hk

Yikai Wang

Department of Economic
University of Oslo
Box 1095 Blindern,
0317 Oslo, Norway
E: yikai.wang@econ.uio.no

Tong Zhang

Department of Economics
University of Zurich
Schönberggasse 1, 8001 Zurich
Switzerland

The work described in this paper was supported by Project no. IEMS16BM03 funded by the HKUST Institute for Emerging Market Studies with support from EY.

1 Introduction

The economic imbalances and financial market frictions are important phenomena in the Chinese economy. One of imbalances often discussed is the low consumption share and the high investment share of the GDP. The financial market frictions in the economy include the financial repression - the low interest rate for deposits, and the private and the state firms' different access to the financial resource. The consumption imbalance and the financial market frictions are in fact connected: the financial frictions lead to a lower household income, which forces households to reduce consumption and save for future necessary expenditures. Understanding this link is crucial for finding out the efficient policy reforms to stimulate the consumption and to rebalance the economy.

One traditional policy to increase the consumption and to stimulate the economy, based on the experience in developed economies such as the U.S., is lowering the interest rate. It is believed that a lower interest rate can reduce the household saving and increase the consumption. However, in China, the opposite happens when the interest rate goes down. As we can see from Figure 1, a lower real interest rate is associated with a lower consumption share in China, compared to a higher consumption share in the U.S. The reason for the positive relation in China is that the Chinese households need to save a large fraction of the income for future necessary expenditures such as medical expenditures and housing purchase. These necessary expenditures are important in China because the household income is relatively low and the public insurance for these expenditures is weak. In a country with a high income, good social security and medical insurance, and an efficient rental market, households do not have to save too much for the old-age necessary expenditures. So in China, the saving motive for these expenditures is strong and this saving target imply that the low interest rate for household savings is associated with a higher saving rate. The financial repression, i.e., the low interest rate for household savings, is one reason for the economic imbalance.

Moreover, the friction in the financial resource allocation between state and private

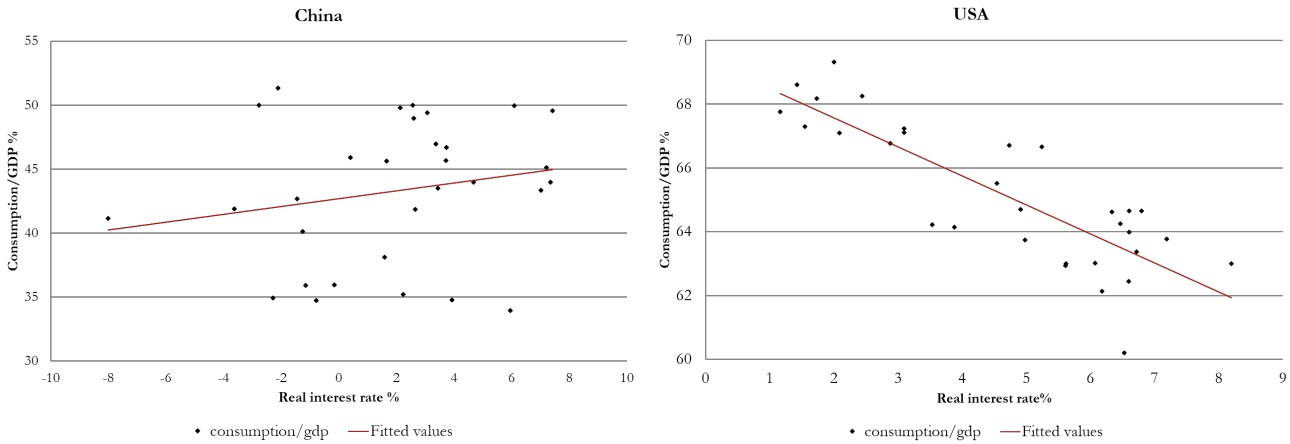

Figure 1: Consumption Share and Interest Rate

firms also aggravates the imbalance. The government provides explicit and implicit guarantees for state firms and also some private firms in sectors that are crucial for the government, e.g., the real estate. In this case, banks know that these government-backed firms will always be able to pay back loans, even if they are not very productive or profitable. Then banks prefer to lend to these safe borrowers, and have no incentives to find other borrowers who are potentially more productive. Many inefficient firms in the sunset industries should have stopped producing but they still survive only because of this government guarantee and bank loans. Those are the so-called "zombie firms". The inefficient allocation of bank loans and capital reduces the aggregate firm productivity and household income. The low household income forces households to consume even less to save up for the future necessary expenditures. Moreover, the connected firms with abundant financial resource find the investment in the production gives low return, they may divert some investment out of production and into speculation, e.g., into the real estate market. This creates a price boom in the real estate. Households who save for housing purchase need to save even more and consume even less. The economic imbalance is worsen.

To rebalance the economy and achieve the more efficient allocation of financial resources, the financial reform towards market-determined interest rates is necessary. Poli-

cies directly stimulating the economy may lead to the opposite. First, as we discussed above, lowering the interest rate to boost consumption may work in the opposite direction in China. Second, some other traditional stimulations such as increasing supply of money, bank loans, and liquidity do not necessarily increase investment in efficient production and the output. The injected liquidity and bank loans are mostly allocated to the connected firms, which already have enough financial resources. They tend to invest the additional resource to speculation, while the unconnected firms which want to expand production do not benefit from the injection of liquidity. This is a new type of "liquidity trap", different from the traditional type due to the inability to lower the nominal interest rate.

In this paper, we build an overlapping generation model to study the financial frictions and the economic imbalance. Based on the overlapping generation model built by Song et al. (2011), we introduce the following special features in Chinese economy: the price scissors in interest rates for deposits and loans, the financial guarantee for state firms, and necessary expenditures for the old households. They capture the key features in China's financial market and the economy. This model is also useful to evaluate policy reforms to rebalance the economy, in addition to some other important issues such as the efficiency of resource allocations and effects of simulation policies.

This paper contributes to two strands of literature. The first is the literature on China's growth and structural change with the state-private dichotomy. Song et al. (2011) study how the state-private dichotomy affects China's economic growth, foreign reserve, and so on. They show that given the credit constraint on private firms' borrowing, the high household savings cannot be channeled to private firms and eventually lead to a large foreign reserve. This paper contributes by showing that the constraint on private firms' borrowing and the inefficient capital allocation between state and private firms also contribute to the high household saving because this inefficiency results in a low income, and then a high saving rate is necessary for covering the old-age expenditures. It also provides

the micro-foundation for private firms' borrowing constraint and state firms' easy access to bank loans by showing how the government guarantee on state firms' loans changes the incentive of state banks in loan allocation. The second strand of literature is on China's high household saving rate. Various saving motives have been identified, including saving for housing (Chamon and Prasad (2010)), marriage (Wei and Zhang (2014)), old-age support (Banerjee et al. (2014)) and so on. This paper incorporates these saving motives into a macro model, and show how the macro environment and policies - for example the low interest rate - strengthen these saving motives and result in the high saving rate. It also studies which policies can mitigate these saving motives and rebalance the economy.

The rest of the paper is organized as follows. Section 2 discusses crucial facts in China's economy and the financial system which motivate the theory. In section 3, we build the benchmark model to capture the key features in China's financial system. The model is used to understand the empirical facts and also to study the implications of government policies. In Section 4, we extend the benchmark model to study more features such as the growing shadow banking system and to evaluate more policies, e.g., stimulus packages. Section 5 concludes.

2 Motivating Facts

In this section, we discuss a couple of important motivating facts in China's financial system and the imbalanced economy, including the investment share, the correlation between consumption and interest rate, and the leverage ratios of different firms.

2.1 China's High Investment Share

China's investment share has always been the highest among all countries. Korea and Japan have also experienced a couple of years of rapid growth and high investment, but their investment shares of GDP in the rapid-growing periods are still lower than the in-

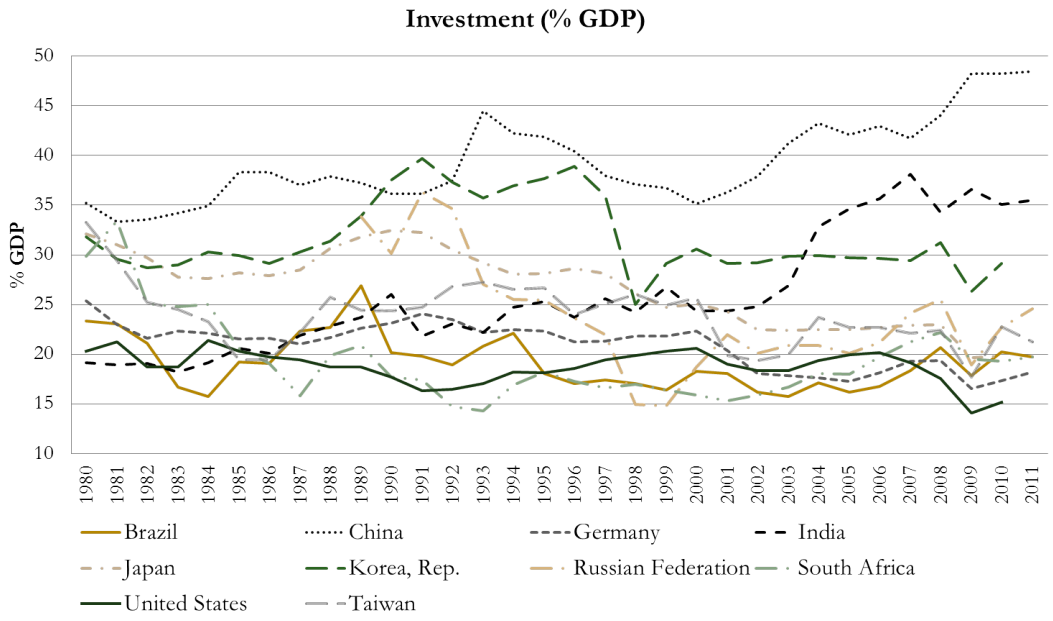

Figure 2: Investment Share of GDP

vestment share in recent China. As shown in Figure (2), China's investment share of GDP is above 35%, and it even increases to almost 50% in recent years.

2.2 Consumption Share and Interest Rate

The high investment rate implies the low consumption share of GDP. As we can see from Figure (1), the consumption equals about 45% percent of GDP in China, while it is around 65% in the U.S. The other contrast is that in China the consumption share is positively correlated to the real interest rate, while it is negative in the U.S. In other words, when the interest rate is lower, Chinese households do the opposite of the U.S. households: consume less and save more. This negative relation between the real interest rate and the saving of Chinese households is also confirmed by Nabar (2011) using a household level panel data over the period 1996–2009. Moreover, we can see that real interest rates in China fluctuate around the level of 0%, much lower than the real interest rates in the U.S., which average at about 4-5%. This is related to the financial repression that reduces the interest rate of the household savings.

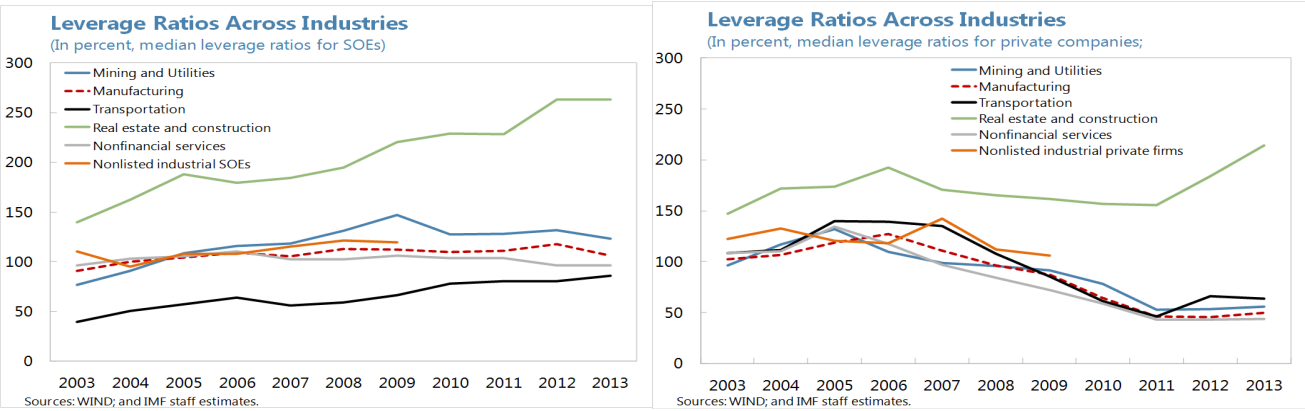

Figure 3: Leverage Ratios

2.3 Leverage Ratios

The other important feature in China's financial system is that the state and private firms are very different in their abilities of getting bank loans. State firms can borrow relatively easily and cheaply from state banks. As shown in the left panel of Figure 3, the leverage ratios of state firms, have been high, and have been growing rapidly in recent years to the level of about 100%, especially in the real estate and construction sector. The right panel shows a different pattern for POEs, who leverage ratios have been declining down to around 50%, except for the real estate and construction sector.

3 The Benchmark Model

We first build a two-period overlapping generation (OLG) model to illustrate the main relation between the financial frictions and the economic imbalance. Later we extend the benchmark model to discuss more features in the economy and more policy implications.

3.1 Preferences, Technology and Markets

Households Time is discrete and there are infinite periods. Each cohort of households live for two periods: young and old. The utility of the representative household of the

cohort born in period t is:

$$U_t = \log\left(C_t^y\right) + \beta \log\left(C_{t+1}^o - \underline{C}^o\right).$$

She discounts the future at the rate β and the coefficient of relative risk aversion equals 1. $\underline{C}^o$ is the subsistence level of consumption of the old, representing the necessary old-age expenditures. The necessary expenditures include medical expenditures, housing purchases, and so on. The young supplies one unit of labor and receives labor income w_t. She also pays a lumpsum tax T_t. She consumes C_t^y and saves A_{t+1}. The old does not work. She receives return on her saving from the last period and consumes it.

Firms In each period, a firm i maximizes its expected profit. We can omit the t subscript for firms' problem when there is no confusion. Firm i produces with capital K_i and labor N_i. The capital is financed through bank loans. The labor is hired in the competitive labor market. The realization of firm i's production can be good or bad. With probability p_i, the productivity shock is high, and the output is $Z^H K_i^\alpha N_i^{1-\alpha}$; with probability $1 - p_i$, the output is $Z^L K_i^\alpha N_i^{1-\alpha}$, where $Z^L < Z^H$.

There are two types of firms - the financially connected firms (F firms) and the entrepreneurial firms (E firms). They are different in two aspects. First, an F firm is connected to and backed up by the government. The government provides explicit or implicit guarantee to the firm's borrowing, i.e., if the firm is not able to pay back the bank loan, the government steps in and pays for the loan. F firms represent state firms and also some private firms in the industries that the government strategically supports and backs up, e.g., real estate and construction. Second, F firms are less productive than E firms, which means that their probability of receiving the high productivity shock is lower: $p_F < p_E$.

Government The government helps to pay for the loans of F firms, and this payment can be financed using profits from banks, and taxes on young workers. Here we assume

that the government budget is balanced every period.

The Bank The representative bank is state-owned. In the benchmark, the interest rate for household deposits in the bank - R - is set by the government. It is set at a level that is lower than the interest rate of bank loans to firms and this gap is the so-called price scissor. The interest rate for bank loans to F firms and E firms are determined by the market demand and supply, given the government's financial support. The interest rates for loans to F firms R_F and to E firms R_E can be potentially different. The bank take the interest rates as given. This is consistent with the features of China's financial market in recent years. [1] The final payment to the bank from a firm depends on the realization of the productivity shock and the government's financial support. Denote the actually interest rates paid back to the bank as ρ_F and ρ_E, from F firms and E firms respectively. If a firm i is not able to pay back the loan according to the promised interest rate, then the residual value of the firm, after paying the wage bill, is claimed by the bank. In this case, $\rho_i < R_i$.

The objective of the manager of a state bank is not only to maximize the expected profit, but also to reduce the probability of loan default, as follows

$$\max_{L_F, L_E} \mathbb{E}\left(\rho_F L_F + \rho_E L_E\right) - \eta\left(P\left(\rho_F < R_F\right) L_F + P\left(\rho_E < R_E\right) L_E\right),$$
$$\text{s.t. } L_F + L_E \leq A.$$

η can be considered as the additional punishment to the manager of the bank when firms default on the loans. In China, this additional cost is relevant, because the state bank managers are also government officers. If a bank manager gives a loan to a firm which ends up defaulting on the loan and claiming bankruptcy, the manager will be held accountable,

[1] An recent reform in 2015 allows banks to freely set the interest rate for deposits with more than one-year maturity. Still, other interest rates are regulated and influenced by the government. The return to household savings is still kept at a low level. So the benchmark setup still represent the current market well.

and this will have a negative impact on her promotion.

3.2 The Equilibrium

In the beginning of each period, given the savings of the old, the following events happen:

1. The bank allocates the savings of the old as loans to the F firms and E firms.

2. Firms hire workers. Then they receive productivity shocks and produce. Firms pay the wage bills, and then pay for the bank loans, with the government's financial support for F firms.

3. The government taxes the young.

4. The young receive income, consume and save in the bank. The old receive the return from their savings, and consume.

Then the economy enters the next period. The young become the old.

Households The young solve the following problem:

$$\begin{aligned} \max_{A_{t+1}} \log C_t^y + \beta \log \left(C_{t+1}^o - \underline{C}^o \right) \\ \text{s.t. } C_t^y + A_{t+1} = w_t - T_t, \\ C_{t+1}^o = R_{t+1} A_{t+1}. \end{aligned}$$

Under the condition that the income is high enough to cover the necessary expenditure, i.e. $w_t - T_t \geq \frac{\underline{C}^o}{R_{t+1}}$, which we assume to hold throughout the paper, the solution is the following:

$$\begin{aligned} C_t^y &= \frac{1}{1+\beta} \left(w_t - T_t - \frac{\underline{C}^o}{R_{t+1}} \right), \\ A_{t+1} &= \frac{\beta}{1+\beta} \left(w_t - T_t \right) + \frac{1}{1+\beta} \frac{\underline{C}^o}{R_{t+1}}. \end{aligned}$$

The solution shows the relation between consumption and the interest rate, as follows.

Proposition 1 (Consumption and saving)**.** *Consumption of the young is positively related to the interest rate, while the saving is negatively related to the interest rate. If the interest rate goes down, consumption of households - both the young and the old - goes down.*

This proposition offers an explanation to the relation between consumption and the interest rate in China in Figure 1. More specifically, this proposition states that, if the interest rate R_{t+1} is higher, it becomes easier to save up for the necessary expenditures, so the young save less and consume more. When the young become old in the next period, they also consume more, because $C^o_{t+1} = R_{t+1}A_{t+1} = \frac{\beta}{1+\beta}R_{t+1}(w_t - T_t) + \frac{1}{1+\beta}\underline{C}^o$ is also positively correlated to the interest rate. Moreover, if the interest rate R_t becomes higher, the old in period t also consume more. So if both R_t and R_{t+1} go up, the total consumption of households, either we look at the old and the young at period t, or we consider the lifetime consumption of the young born at period t, are positively correlated to the interest rate.

This result is not only valid for log-utility, but also robust to some variations of utility functions if the necessary expenditure is large enough. If the inter-temporal elasticities is smaller than 1, the positive relation between consumption and interest rate becomes even stronger. If the inter-temporal elasticity is larger than 1, the results can still hold given a relatively high level of $\underline{C}^o$. For example, consider a CRRA utility $u(C) = \frac{C^{1-\frac{1}{\theta}}}{1-\frac{1}{\theta}}$ with $\theta > 1$. The optimal consumption of the young is $C^y_t = \frac{w_t - T_t}{1+\beta^\theta R^{\theta-1}_{t+1}} - \frac{\underline{C}^o}{R_{t+1}+\beta^\theta R^\theta_{t+1}}$. The interest rate enters the expression twice and has two effects on the consumption. One refers to the standard substitution and income effects. R_{t+1} shows up in the first term and is negatively related to C^y_t, when $\theta > 1$. The substitution effect dominates the income effect, and the household has the incentives to save more because tomorrow's consumption is cheaper. The other refers to the incentive to save for necessary expenditures. R_{t+1} shows up in the second term and is positively related to C^y_t: the higher interest rate, the easier it is to save up. So the young can consume more. We can see that if (1) $\underline{C}^o$ is large relative to $w_t - T_t$, and (2) $\theta - 1$ is small, the second effect dominates the first one, so C^y_t is still

positively correlated to R_{t+1}. In other words, if $\underline{C}^o$ is small relative to $w_t - T_t$, then C_t^y can be negatively related to R_{t+1}, and a lower interest rate can increase consumption, as the case in the U.S. An important difference between the economy of China and U.S. is the size of necessary expenditures when old, relative to the income level. Naturally, in the U.S., the income level is high and the necessary expenditure's impact on consumption and saving is relatively small. In China, the necessary expenditure is an important part of the relatively low income, so consumption moves towards the opposite direction compared to the U.S. when interest rate moves.

Interest Rates In the equilibrium, the zero profit condition pins down the interest rate for bank loans to F firms: it equals the marginal return of capital of F firms with the high productivity shock. Every F firm is willing to accept this interest rate: if realization of the productivity is high, the loan can be paid back using the capital return, and the firm's profit is 0; if the realization is low, the F firm receive the government support to pay back the loan and the profit is also 0. If the interest rate is lower than this level, every F firm expects a positive net return of capital and wants to borrow infinitely from the banks and the market does not clear. So

$$\rho_F = R^H,$$

where R^H is the return to capital when the productivity shock is high. We denote the return when the productivity is low as R^L.

This interest rate is higher than the level that E firms can afford. Because without the government support, capital return for E firms with low productivity shocks is lower than R^H, and so is the expected return of capital of E firms. So the bank prefers lending to F firms to lending to E firms. Formally speaking, in a bank manager's objective, lending to an F firm gives a safe return R^H , but lending to an E firm at this interest rate implies an expected return of $p_E R^H + (1 - p_E) R^L$, and additionally, the punishment $\eta (1 - p_E)$. The equilibrium bank loan allocation is $L_F = A$, and $L_E = 0$. This implies that all capital

and labor are allocated to F firms: $K_F = A, N_F = 1$. Then the rate of return to capital is: $R^H = \alpha Z^H K_F^{\alpha-1} N_F^{1-\alpha} = \alpha Z^H A^{\alpha-1}$, and the equilibrium wage is $w = (1-\alpha) Z^H A^{\alpha}$. We can summarize the equilibrium interest rates and resource allocation in F and E firms as follows.

Proposition 2 (Interest rate). *The interest rate for bank loans to F firms is $R_F = \alpha Z^H A^{\alpha-1}$ and all bank loans and labor are allocated to F firms.*

The allocation in this proposition is extreme. It represents the qualitative feature of bank loan allocation in favor of government backed firms. In the extensions, there will be allocation of loans to E firms, after we introduce the shadow banking institutions.

The Government The government's income includes tax from young households and profits from banks. The government expenditure is the financial support for F firms. The government budget balances every period:

$$T + \left(R^H - R\right) A = (1 - p_F) \left(R^H - R^L\right) A, \tag{1}$$

where R^L stands for the return to capital when the productivity shock is low, and $R^L = \frac{Y^L - wN^L}{K^L} = \left(Z^L - (1-\alpha) Z^H\right) A^{\alpha-1}$. This pins down the tax level T:

$$\begin{aligned} T &= (1 - p_F) \left(R^H - R^L\right) A - \left(R^H - R\right) A \\ &= \left(R - p_F R^H - (1 - p_F) R^L\right) A \\ &= RA - \left(p_F z^H + (1 - p_F) Z^L - (1-\alpha) Z^H\right) A^{\alpha}. \end{aligned} \tag{2}$$

Essentially, the financial support that the government provides to the low productive F firms goes back to the government as part of bank profits. So eventually the government gets from F firms and banks the output of F firms minus the wage bill. If Z^L is low or p_F is low, this part can be negative, i.e., the output of F firms is lower than the wage bill. The

cost for government is the interest payment to household savings -RA. The difference is financed through the lumpsum tax on the young T.

3.3 Dynamics

Given the above expressions for all variables in each period, now we can calculate the household savings and the dynamics of the economy.

The young household after-tax income is

$$\begin{aligned} w_t - T_t &= (1-\alpha) Z^H A_t^\alpha - \left(R_t A_t - \left(p_F Z^H + (1-p_F) Z^L - (1-\alpha) Z^H \right) A_t^\alpha \right) \\ &= \left(p_F Z^H + (1-p_F) Z^L \right) A_t^\alpha - R_t A_t \\ &= Y_t - R_t A_t. \end{aligned}$$

This is essentially the total output of firms, minus the capital income of the old. The wage bill appears in the household's income and also in the government's cost which enters the tax, so it cancels out. Even though the government support helps F firms to pay higher wage to workers, it does not really help workers in the end, and what matters is still the total output.

Then the aggregate saving evolves according to

$$\begin{aligned} A_{t+1} &= \frac{\beta}{1+\beta} (w_t - T_t) + \frac{1}{1+\beta} \frac{\underline{C}^o}{R_{t+1}} \\ &= \frac{\beta}{1+\beta} (Y_t - R_t A_t) + \frac{1}{1+\beta} \frac{\underline{C}^o}{R_{t+1}}. \end{aligned} \tag{3}$$

The saving rate of the young is

$$\frac{A_{t+1}}{w_t - T_t} = \frac{\beta}{1+\beta} + \frac{1}{1+\beta} \frac{\underline{C}^o / R_{t+1}}{Y_t - R_t A_t}, \tag{4}$$

while the aggregate saving rate is

$$\begin{aligned}\frac{Y_t - C_t^y - C_t^o}{Y_t} &= \frac{Y_t - \frac{1}{1+\beta}\left(Y_t - R_t A_t - \underline{C}^o / R_{t+1}\right) - R_t A_t}{Y_t} \\ &= \frac{\beta}{1+\beta} + \frac{1}{1+\beta}\frac{\underline{C}^o / R_{t+1}}{Y_t} - \frac{\beta}{1+\beta}\frac{R_t A_t}{Y_t}. \end{aligned} \tag{5}$$

Both saving rates are positively related to $\underline{C}^o$, implying that the necessary expenditures increase the savings. They are also negatively related to the next period interest rate R_{t+1}, implying that a lower interest rate for the savings induces more savings for the future necessary expenditures.

3.4 Implications

3.4.1 Understanding the Empirical Facts

The simple benchmark model offers explanations for the phenomena in the Chinese economy, discussed in section (2). First, due to the motivation to save up for future necessary expenditures, household saving rate is high. Formally speaking, from equations 5, we can see that the aggregate saving rate is positively related to $\underline{C}^o / Y_t$. In this interpretation, the necessary expenditure relative to income is high in China. It can be because of the relatively low income in China, the lack of social support for the old, or the cultural difference in home-owning preferences. Moreover, the fast growth, the inflation, and the rising housing price in China make households expecting more expensive medical services and housing expenditures in the future, and these imply that the consumption constraint becomes more prominent. Second, lowering interest rate for household savings increases the saving rate and reduces the consumption share. This can be seen from the negative relation between the saving rate and R_{t+1} in equation 5. This is because that a lower interest rate makes it more difficult for the households to build up enough wealth for the consumption in the later stage of life, so a lower interest rate increases the saving

rate in China, as the opposite of the U.S. The third fact, i.e., the high debt and leverage of SOEs and firms in certain industries connected to the government such as the real estate and construction sector, is the outcome of the government guarantee. This corresponds to the model outcome that F firms receive the government guarantee and all the bank loans.

3.4.2 Policies to Rebalance the Economy

The imbalanced low consumption and high saving in China's economy has been an important issue, and there has been some attempts to mitigate it, including reducing the interest rate in the hope of stimulating the consumption. However, as we see from the empirical correlation and the theoretical analysis, it does not work in China given the structure of the financial system. Here, with the aid of the model, we can study alternative policies and their consequences, and evaluate whether they help to rebalance the economy. Moreover, we can also study which policies can improve the resource allocation among firms.

Removing the Government Guarantee The inefficient capital allocation and the additional cost of providing financial support reduce the total output, and also the household income. If the government support for F firms is removed, and the bank loan allocation is determined by the market, then the more productive E firms can receive more bank loans and the aggregate output can be higher. The household income becomes higher and consequentially, the necessary expenditures become less important. The young can easily save for the future expenditures and the saving rate goes down.

We can formally illustrate this logic with the aid of the model by showing the equilibrium outcome without the government guarantee for F firms. If the government does not guarantee the loan to F firms, then when a bank manager decides whether to give a loan to an F firm or an E firm, she compares the probability that the F firm receives a good enough productivity shock to pay back the loan and the probability for the E firm.

Of course, the interest rates of loans to F firms and E firms also matter. In the equilibrium, the interest rate for E firms is the return to capital when the an E firm receives the high productivity shock, i.e. , R^H, under the condition that the return to capital when it receives the low productivity shock is low enough.[2] The highest interest rate that F firms can accept, without government support, is also R^H, but it is less attractive for a bank manager to lend to an F firm compared to an E firm, because now the risk of lending to F firms is higher: $p_F < p_E$. Now all financial resources are allocated to E firms: $K_E = L_E = A$, and $N_E = 1$. The total output is

$$Y = \left(p_E Z^H + (1 - p_E) Z^L\right) A^\alpha.$$

Now the government does not spend on providing financial support for F firms, so the budget contains only the profits from the banks and the tax income. Through the tax on the young, eventually, like before, the young's after-tax income equals the total output minus the capital income of the old. Like before, we have $w_t - T_t = Y_t - R_t A_t$. The income of the young is now higher, because the total output Y_t is larger than in the case that the government provides guarantee for F firms. The necessary expenditure relative to income becomes smaller. So households can consume more now and save less. This income effect helps to increase the consumption and to rebalance the economy. Formally speaking, from equation 4, we can see that the young's saving rate is lower when Y_t is higher. Similarly, from equation 5, we can see that a higher Y_t implies that the aggregate saving as a share of output becomes lower and the consumption share higher.

Market-Determined Interest Rate for Deposits In the benchmark setup, the interest rate for household savings in the banks is set by the government at a low level. If this

[2]Of course, a bank has the choice to charge the E firm the interest rate that equals the return to capital when the productivity is low, so that the E firm can always pay back. However, if Z^L is low enough, the return to capital can not even cover the cost of loans for the bank, i.e., $R^L < R$, which is possible, for example, if $Z^L = 0 < R$.

interest rate is determined by the market - the zero profit condition, then it is at the level of the expected return to capital: $R = p_E R^H + (1 - p_E) R^L$, under the condition that the government does not provide financial support to F firms. If this level is higher than the previous level that the government sets, then the saving rate goes down, as we can see from equations 4 and 5.

The mechanism is the following: the higher return to household savings relaxes the households from the burden of saving up for future necessary expenditures so that they can consume a larger share of the income and save a lower share. From this analysis, we can also see that the traditional policy of reducing interest rate to stimulate consumption would lead to even lower consumption and more severe imbalance in the economy.

4 The Extension with Shadow Bank Institutions

In this section, we extend the benchmark model by introducing the shadow bank institutions. With this extension, we can study more implications of financial market frictions and the effects of more policy reforms.

4.1 Shadow Bank Institutions

We first introduce into the model the shadow banking institutions (SBIs), which are different from the state banks in two aspects: (1) a SBI maximizes the expected profit, without considering the additional cost of unpaid loans that the state bank adds on top of the expected loss; (2) a SBI cannot diversify the risk by lending to many firms. More specifically, the first difference means that while a state bank manager not only cares the expected profits but also tries to avoid the additional cost of unpaid loans, e.g., the punishment in her promotion, a manager in a SBI only cares about the expected profit. The second difference exists because state banks are large and operate at the national level, so they can lend to different firms in different regions to diversify the risk. So the return

to a household's saving in a state bank does not depends on the return of the loan to a particular firm but the average return of all loans. In contrast, SBIs are relatively small, and often operate at regional levels. They can not fully diversify risks by lending to many different firms in many different regions. If a SBI lends to one firm, then so the return to the savings in a SBI equals the return of the loan to the particular firm which receives the loan from the SBI, subject to the risk of productivity shock to the firm.[3]

In the model, a household i decides the saving in a SBI - $a^S_{i,t+1}$, in addition to the saving to the state bank - $a^B_{i,t+1}$. A SBI can give a loan to an E firm. [4]This means that the return to the shadow bank's loan is subject to the productivity shock to this E firm. With probability p_E, the E firm receives a high productivity shock and gives return R^H, and with probability $1-p_E$, R^L. The equilibrium interest rate that the shadow banks charge to E firms is $R_E = R^H$, and if an E firm with the low return cannot pay back the loan, the shadow bank claims the residual and the return to the loan is R^L. Now saving in a shadow bank becomes a risky asset for the household, but this asset gives a higher expected return than the safe asset - deposit in the state bank - which gives a return under the condition $R < p_E R^H + (1-p_E) R^L$. Given the higher expected return, the representative household always allocates a positive fraction of the total savings in the shadow banking system. She solves the following problem:

$$\max_{C^Y_t, A^S_{t+1}, A^B_{t+1}} \log C^Y_t + \beta \left(p_E \log \left(C^{OH}_{t+1} - \underline{C}^O \right) + (1-p_E) \log \left(C^{OL}_{t+1} - \underline{C}^O \right) \right) \tag{6}$$

$$\text{s.t. } C^Y_t = w_t - T_t - A^S_{t+1} - A^B_{t+1}$$

$$C^{Oj}_{t+1} = R^j_{t+1} A^S_{t+1} + R_{t+1} A^B_{t+1},$$

[3]If the SBI lends to firms in one region, the return to the savings in the SBI is subject to the risk of the shock to the regional productivity.

[4]We can also allow the SBIs to give loans to F firms, but in the equilibrium, they choose not to do so. This is because SBIs have no advantage in channeling household savings to capital of F firms, compared to the state bank, which is better in risk diversification. SBIs have comparative advantages in lending to E firms because the state bank does not want to take the risk to lend to E firms.

where $j \in \{H, L\}$. The optimal choice for C_t^Y is the same as before: $C_t^Y = \frac{w_t - T_t - \underline{C}^o / R_{t+1}}{1+\beta}$, i.e., a constant fraction of the income after deducting the necessary expenditures. Then the rest of the income is the total saving - A_{t+1}, which is also a constant fraction of the income minus necessary expenditures. The saving of a household is allocated in the state bank and a SBI. We can denote the fraction of the saving in the SBI as μ and express $A_{t+1}^S = \mu A_{t+1}$, and then the saving in the bank is simply $A_{t+1}^B = (1-\mu) A_{t+1}$. The optimal fraction of the saving in the SBI is

$$\mu = \frac{p_E}{1 - R^L/R} - \frac{1 - p_E}{R^H/R - 1}.$$

The investment in the SBI is positively related to p_E, R^H and R^L, but negatively related to R. This result is intuitive: the higher expected return from saving in the SBI, the more attractive the savings in the SBI is. The opposite is true for the return from the saving in the bank. Notice that R is set by the government which is not affected by the equilibrium wage, while R^H and R^L are the returns of the firms, which are negatively related to the wage.

Given the capital allocation in firms, the wage is simply the expected marginal return of labor, which is equal in the E firms and in the F firms. Then we can express the wage as the expected marginal return of labor given the aggregate effective capital:

$$w_t = (1-\alpha) \left(\mathbb{E} Z_F^{\frac{1}{\alpha}} K_{Ft} + \mathbb{E} Z_E^{\frac{1}{\alpha}} K_{Et} \right)^{\alpha},$$

where $\mathbb{E} Z_j^{\frac{1}{\alpha}} = p_j \left(Z^H\right)^{\frac{1}{\alpha}} + \left(1 - p_j\right) \left(Z^L\right)^{\frac{1}{\alpha}}, j \in \{E, F\}$, represents the expected efficiency of capital.

The SBIs provide a channel for the financial resource to flow from households to E firms. Given the same amount of asset A_t, now a fraction of them are allocated into the more productive E firms, i.e., $K_{Et} > 0$, then the total output is higher compared to the case that all savings are allocated to F firms if there are no SBIs. Same for the wage w_t.

A higher wage, as we discussed before, reduces the relative importance of the necessary expenditure, reduces the saving and increases the current period consumption. We can see that the introducing of the SBIs can reduce the economic imbalance. However, the equilibrium is still not efficient, given the undiversified risks in the shadow banking system, and the non-zero capital in the inefficient F firms. There are policies that can reduce SBIs' risks and improve their efficiency, and can help rebalancing the economy. Below, we study the effects of these policies.

4.2 Deposit Insurance

To reduce the risks of household savings in SBIs, introducing deposit insurance is an effective approach. The constraint for SBIs, or generally the relatively small regional banks is that they can not diversify risks at a large scale. Deposit insurance allows for that. It pools the risks to the national level, so deposits in the regional banks, up to certain limit, are guaranteed to be paid back. Consider the case that the savings in the SBIs are covered by the deposit insurance. Though the returns to some SBIs' loans are low when the E firms receive negative shocks, the returns to the household deposits in these SBIs may not be affected: the deposit insurance guarantees that their deposits will be paid back, up to the limit. Then the optimal choice for households is to first put savings, up to the limit that is insured, in the SBIs, and then allocate the rest of the savings, in the safe state banks and the risky SBIs according to the expected return and risk like described above. When the limit is high enough such that all household savings can be covered by the insurance, households can safely put all savings in the SBIs, as long as the regulated interest rate in state banks is low. Under the high enough limit condition, all the household savings in the SBIs receive the same return independent of the realizations of the related firms' productivities:

$$R^S_{t+1} = R^H_{t+1} = R^L_{t+1} = \left(p_E Z^H + (1 - p_E) Z^L\right) \alpha A^{\alpha-1}_{t+1}.$$

Given that the interest rate for deposit in the state bank is lower than this level, the solution to the household's problem 6 is simply $A^B_{t+1} = 0, A^S_{t+1} = A_{t+1}$, where the total saving A_{t+1} is still a constant fraction of the young's income minus the necessary expenditure, as in the benchmark. The total output is higher than in the benchmark because now the expected productivity of capital is $p_E Z^H + (1 - p_E) Z^L > p_F Z^H + (1 - p_F) Z^L$. The higher output and income imply a lower saving rate of the young and also a lower aggregate saving rate. So the deposit insurance can increase the income and help to rebalance the economy. Moreover, the deposit insurance also reduces the risk of the household savings. In the model with the log utility, this change in the risk does not directly affect the saving, but if households are prudent enough, the lower risk can reduce the household saving and can help to increase the consumption share.

China introduces the deposit insurance in 2015. It is estimated to cover 99.6% depositors but only 46% of the money on deposit. This policy helps small banks reducing the risk and attracting household savings. However, still more than half of the deposits are not insured and the big depositors may choose to save them in the large state banks. If the deposit insurance limit and coverage - in both types and sizes of assets - are extended, there will be more resource allocated to productive firms through the small banks and SBIs.

4.3 Stimulus Policies

We discussed above how lower interest rate for deposits in the state banks affect household savings, and consequentially affect the capital and interest rate for loans through the equilibrium effects. In addition to that, the government can also directly lower the interest rate for loans or increase the loan supply to stimulate the economy. The 2008 four-trillion economic stimulus plan is an example of directly increasing the supply of liquidity and loans. We can study the effects of these types of policies by looking at the comparative statics of L, R_E, and R_F.

Consider the case that the government can directly increase the loan supply by injecting liquidity into the state bank. This means that the state bank loan supply is set by the government and it can be larger than the household saving in the state bank: $L^B = A^B + I^g$, where $I^g \geq 0$ is financed by the government borrowing from the international market at the cost of R^I. In the SBIs, the total loan size still equals the deposit size: $L^S = A^S$. Recall that in the equilibrium, state banks only lend to F firms and SBIs to E firms. Then the additional financial resource injected into the economy is all allocated to F firms. Now the resource allocation becomes more in favor of F firms which expand their production more. The E firm capital allocation is not directly affected the expansion policies. However, E firms are hurt by the policy because the expansion of the F firm production increases the wage. Now the equilibrium wage is

$$w_t = (1-\alpha)\left(\mathbb{E}Z_F^{\frac{1}{\alpha}}\left(A_t^B + I_t^g\right) + \mathbb{E}Z_E^{\frac{1}{\alpha}}A_t^S\right)^{\alpha},$$

and it increases when I_t^g increases. Given the higher wage and the same amount of capital, E firms hire less workers and get a lower return to capital. The aggregate output of firms increases, because of the higher output of F firms, but the efficiency and the TFP decrease. In fact, the net output - domestic output minus the cost of borrowing from the international market - may decline, depending on the difference of the return to capital in F firms and the interest rate in the international market. The net output can be expressed as

$$Y_t - R^I I_t^g = \left(\mathbb{E}Z_F^{\frac{1}{\alpha}}\left(A_t^B + I_t^g\right) + \mathbb{E}Z_E^{\frac{1}{\alpha}}A_t^S\right)^{\alpha} - R^I I_t^g,$$

and the marginal change is

$$\frac{\partial\left(Y_t - R^I I_t^g\right)}{\partial I_t^g} = R_F^K - R^I,$$

where R_F^K is the expected marginal return to capital in F firms. If the return to F firm

capital is already low such that $R_F^K < R^I$, the stimulation can reduce the net output.

Next, we can study the impacts of the stimulation on household income, consumption, and saving. As we have seen above, the total output may decrease if the cost of the stimulus dominates the return of from the investment in F firms. In this case, the income of the young goes down, even though the wage goes up. Intuitively, this is because the expenditures of the government - now including the cost of stimulation with liquidity injection - are paid by the young households through the tax. Then given the necessary expenditure constraint, the saving rate goes up and consumption share of GDP declines. The economy becomes more imbalanced during the stimulus era.

Another stimulus policy that the government can implement is lowering the interest rate of bank loans. This policy can effectively affect the interest rate of state bank loans to F firms, but not the interest rate in the SBIs. Then the consequence is similar. F firms benefit from the low interest rate and increase borrowing and investment. Then the state bank loan size adjusts according to the increasing demand from F firms, and this low loan interest rate policy has the same effects as the loan expansion policy described above.

The Chinese stimulus program contains both: the government injects liquidity and increases the investment, especially in infrastructure and housing sectors, and also reduces interest rate for bank loans, hoping to expand productions. However, these policies may reduce the production efficiency. Eventually they can harm the consumption and domestic demand. Moreover, they increase the financial burden of government. If the government debt increases, the stimulus policies create long-lasting effects in the future, even after the policies stop. Below, we introduce the government debt into the model and discuss the long-run impacts.

4.4 Government Debt

In the analysis above, we assume that the government budget is balanced every period. The government taxes the young to pay for the financial support for F firms and the

stimulus program. Alternatively, the government can fund the stimulus program using the debt, to smooth the taxes over time. In this case, the cost of the stimulation policies may lead to increasing government debt. Eventually, the debt will be paid by the taxes on the future generations. The stimulus program can have long-lasting effects on the income and therefore the consumption of many cohorts.

We can formally model this by setting $T_t = \eta_t \left(R^I D_t + \left(R^H_{Ft} - \mathbb{E}R^K_{Ft}\right) K_F + \left(R^I - \mathbb{E}R^K_{Ft}\right) I^g_t\right)$, and $D_{t+1} = (1-\eta_t) \left(R^I D_t + \left(R^H_{Ft} - \mathbb{E}R^K_{Ft}\right) K_F + \left(R^I - \mathbb{E}R^K_{Ft}\right) I^g_t\right)$, which imply that in period t, η_t fraction of the government expenditure and old debt is paid back by the tax, while the rest is financed through accumulated debt D_{t+1}. Obviously, higher the government expenditure in period t, higher the government debt. The accumulated debt will be paid back by the future cohorts. This implies lower incomes for these households. According to the analysis above, the consumption shares will also be lower for these cohorts. The imbalance of the economy can persistent even after the stimulus program stops.

The analysis above considers the debt of the central government and assumes that it is always able to pay back the debt. Some debts are at the local government levels. Depending on different amounts of government financial support and stimulus expenditures, local governments may be different in their debt levels and their abilities in paying back the debts. If a local government fails in paying back its debt, it can not provide further financial support for the loss of F firms, then the local F firms may bankrupt. Household savings in F firms also become risky. The higher uncertainty can increase household savings and enlarge the economic imbalance.

4.5 F Firms' Investments in Speculation

In the benchmark model, F firms invest all in the production. In reality, there are also non-productive sectors that F firms can invest in. For example, they may purchase houses if the housing price increases fast enough and the return in the housing market is high enough. Firms do not derive utilities from housing like the households do, but purchase

them only for the expected price increase. When firms hold housing for the speculation purpose, the supply of housing to household is reduced and the housing price becomes higher. If the necessary expenditure is related to the housing purchase, then the necessary expenditure also becomes larger when the housing price increases as F firms enter the housing market. Then young households need to save more for future expenditures, and this leads to lower consumption and increases the economic imbalance.

5 Conclusion

In this paper, we build an over-lapping generation model to study China's financial system and economy. This model helps to explain crucial phenomena in Chinese economy, including the low consumption and the high saving, the positive correlation between consumption share and the interest rate, and the high leverage ratios in particular sectors. The model is also used to evaluate policy reforms. We find that traditional stimulus policies such as lowering interest rate and injecting liquidity can not increase output, income and consumption but may have the opposite effects in China, given the frictions in the financial system. They may lower the consumption, increase the saving rate, and enlarge the economic imbalance. The solution lies in reducing the financial market frictions and allowing the market to determine the interest rates. Effective policies include eliminating the price scissors in the state bank, removing the government guarantee for state firms, and providing deposit insurance. They help rebalancing the economy, and improve the efficiency of resource allocation and production. The framework provided in this paper allows for further quantitative exercises to evaluate the effects of these policies.

References

Banerjee, Abhijit, Xin Meng, Tommaso Porzio, and Nancy Qian, "Aggregate Fertility

and Household Savings: A General Equilibrium Analysis Using Micro Data," 2014.

Chamon, Marcos D. and Eswar S. Prasad, "Why are saving rates of urban households in China rising?," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 93–130.

Nabar, Malhar, "Targets, Interest Rates, and Household Saving in Urban China," 2011.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing Like China," *American Economic Review*, 2011, *101* (February), 196–233.

Wei, Shangjin and Xiaobo Zhang, "The Competitive Saving Motive : Evidence from Rising Sex Ratios and Savings Rates in China," 2014, *119* (3), 511–564.