An Equilibrium Model of Entrusted Loans

Ying Liu

July 4, 2019

Abstract

Entrusted loan is a type of inter-corporate loan and a major component of shadow banking in China. In a model with entrepreneurial moral hazard and bank moral hazard, entrusted loans arise endogenously. Entrusted loans involve a lending chain in which high-capitalized firms channel bank loans into medium-capitalized firms. High-capitalized firms obtain cheap bank loans and over-borrow to form shadow banks with the extra capital. Medium-capitalized firms simultaneously borrow from both banks and shadow banks, while low and semi-highly capitalized firms borrow only from banks. As a result of lower bank monitoring, entrusted loans improve the total welfare of banks and firms. However, entrusted loans destroy firms’ value because firms earn reduced expected profits. Default risk is increased and real efficiency reduced. The model can explain the rapid growth of entrusted loans in China since the economic stimulus plan of 2009-2010: credit expansion policies lower the funding cost of banks and further trigger the growth of entrusted loans.

Keywords: Shadow Banking, Entrusted Loans, Moral Hazard, Corporate Structure

JEL: G21, G28, G32, G38

*I would like to thank Norman Schürhoff for his helpful comments and advice. I am also grateful for comments received from Theodosios Dimopoulos, John Duca, Lei Mao, Artem Neklyudov, Steven Ongena, Jean-Charles Rochet, and from seminar and conference participants at IFABS Ningbo conference, UNIL brown bag seminar, HEC Paris Finance PhD workshop, FMA European Conference, Queensland University of Technology, The Chinese University of HongKong(ShenZhen), Cass Business School, The Vanguard Group.

§Shanghai University of Finance and Economics; E-mail: liuying@mail.shufe.edu.cn.
1 Introduction

Since the 2007-2008 global financial crisis, shadow banking in China has undergone rapid growth. The size of this sector accounted for 39 percent of China’s GDP at the end of 2012, and in 2016, it already amounted to about $8.5 trillion, equivalent to 80 percent of China’s GDP, as reported by Moody’s. A dominant form of China’s shadow banking is entrusted loans.\footnote{Shadow banking has different forms in China, the largest sectors are, entrusted loans, wealth management products and trust loans (Elliott, Kroeber, and Qiao (2015), Elliott and Qiao (2015), Ehlers, Kong, and Zhu (2018), Hachem (2018)).} Entrusted loans are inter-corporate loans in which a financial institution acts as a trustee, and the lender and borrower directly negotiate the terms of the loan. The trustee merely acts as a middleman to facilitate the transaction for legal reasons. Classical banking theories posit that banks differ from other credit suppliers in that banks screen or monitor borrowers.\footnote{See, for example, Diamond (1984).} Entrusted loans are associated with direct inter-corporate lending, so the borrowing firms are subject to little or no monitoring by the lending firms. Despite being suboptimal to bank loans regarding monitoring, entrusted loans are becoming prevalent in China.\footnote{Entrusted loans were the largest component of China’s shadow banking sector until 2014. Since 2015, entrusted loans has been the second largest component and wealth management products become the largest one. As documented by Allen, Qian, Tu, and Yu (2019), entrusted loans constituted 32% of shadow banking in terms of RMB amount at the end of 2012. The outstanding amount of entrusted loans has jumped to 13.2 trillion RMB ($1.92 trillion) in 2016.} This phenomenon raises a puzzle. Why do firms rely on other firms to obtain financing, rather than specialized financial intermediaries such as banks? On top of that, what is the driven force of the fast growth of entrusted loans in China since the financial crisis?

In this paper, I provide an economic rationale for why entrusted loans endogenously emerge as a form of shadow banking. I show that the interplay of entrepreneurial moral hazard and bank moral hazard results in a lending chain in which firms with access to cheap bank loans recycle the loans by granting entrusted loans to other firms. A reduction of
bank funding cost is accompanied by an expansion of shadow banking sector. In addition to characterizing the feature of firms engaged in entrusted loans, my model shows that entrusted loans improve the total welfare of banks and firms but destroy firms’ value and reduce the real efficiency.

I develop the model based on Besanko and Kanatas (1993) with firms differing in their capitalization, as in Holmstrom and Tirole (1997). Each firm invests its own capital and raises the remaining part through bank loans or entrusted loans to undertake risky investment project. Limited liability induces entrepreneur to choose effort that influences the success probability of investment project. Banks grant loans and monitor firms so that entrepreneurs exert more effort than they would without bank monitoring. The new element I introduce is that firms may form shadow banks and issue entrusted loans to other firms. The difference between banking and shadow banking, apart from the source of financing, is that shadow banks exert no monitoring efforts compared to banks. Shadow banking is, thus, a different lending technology.

When the banks also face moral hazard problem, there exists an equilibrium in which high-capitalized firms over-borrow from banks and form shadow banks with their extra capital. The reason is the following. Since banks do not commit to provide monitoring, banks associate the monitoring with the amount of loans issued to firms. Given that bank monitoring is costly, the monitoring costs must be incorporated into loan interest rates to incentivize banks to provide monitoring. Because the demand of bank loans is high when the firm has few capital, the bank has to charge a high interest rate to break even. The bank monitoring cost is an increasing, convex function of monitoring intensity, so bank interest rate has increasing return to scale. Therefore, high-capitalized firms have the advantage of obtaining cheap bank loans and choose to over-borrow, then re-lend extra capital to other firms that incur expensive bank loans.

Firms which borrow entrusted loans also borrow from banks at the same time. What
is interesting is that the borrowing firms are medium-capitalized firms, rather than low-capitalized firms. To be precise, the demand of entrusted loans is a hump-shaped function of firm capitalization. Medium-capitalized firms borrow the most entrusted loans, while low and semi-highly capitalized firms only borrow from banks. The reason is as follows. When choosing the amount of entrusted loans to take, entrepreneur of the firm faces a trade-off: a marginal increase of entrusted loans reduces the success probability of investment project because of no monitoring from shadow banks; on the other hand, entrepreneur obtains higher payoff since shadow banks charge lower interest rate than banks. Low-capitalized firms are highly leveraged and entail severe entrepreneurial moral hazard, a marginal increase of entrusted loans causes large reduction of success probability which dominates the gain of entrepreneurial payoff, so these firms are better off not taking entrusted loans. Semi-highly capitalized firms entail low entrepreneurial moral hazard, so the yield difference between bank loans and entrusted loans is small, the increase of entrepreneur payoff by replacing bank loans with entrusted loans is lower than the reduction of success probability, so entrepreneurs are better off not taking entrusted loans. Therefore, the demand for entrusted loans displays a hump-shaped pattern over firm capitalizations.

To study the impact of entrusted loans on social surplus and real efficiency, I examine a benchmark model in which entrusted loans are prohibited. In the benchmark, firms only take the amount of bank loans which is just enough to finance their investment projects. Comparing the model with benchmark, it shows that total welfare of banks and firms is improved in the presence of entrusted loans. The reason is that total bank monitoring cost is reduced when high-capitalized firms recycle bank loans to medium-capitalized firms since lending firms do not monitor borrowing firms. Essentially, excess bank monitoring is avoid in the presence of shadow banking. However, firms’ value are reduced. Because entrepreneurs maximize their utilities rather than firms’ profits. Bank monitoring forces entrepreneurs to exert more effort than they would do without bank monitoring. By participating in shadow
banking, entrepreneurs can shirk and exert less effort. As a result, the success probabilities of firm investment projects are reduced and thereby, real efficiency declines.

This model is then extended by differentiating firms by unrestricted and restricted industries. The restricted industries refer to firms which are prohibited from bank lending.\textsuperscript{4} The lenders of entrusted loans are still high-capitalized unrestricted firms; the borrowers become medium-capitalized unrestricted firms and high-capitalized restricted firms. Shadow banks are the only external financing source of restricted firms, and thus, behave as a monopolist. Welfare analysis shows that by allowing shadow banks to lend to restricted industries reduces the total welfare of banks and firms. The reason is that restricted firms crowd out the borrowing firms from unrestricted industries, but restricted firms generate even lower utility than unrestricted firms because of the lack of monitoring from shadow banks and reduced return due to the high market power of shadow banks.

The contribution of this paper is twofold. First, the model links the growth of entrusted loans with the credit market. The emergence of entrusted loans could be an unintended consequence of loosen monetary policy. This result is in line with the observation that entrusted loans started to grow very rapidly in China after the “4 Trillion RMB Stimulus Plan” from 2009-2010.\textsuperscript{5} One component of the stimulus plan is to increase the lending of commercial banks by lower bank reserve requirements and lower benchmark lending rates.

\textsuperscript{4}In China, such industries include the real estate, coal mining and shipbuilding companies which are regarded as industries with over-capacity, and banks are discouraged to lend to firms in such industries. (Elliott, Kroeber, and Qiao (2015)). Allen, Qian, Tu, and Yu (2019) find that close to half of non-affiliated entrusted loans flew into the real estate and construction industry. A Wall Street Journal article “A Partial Primer to China’s Biggest Shadow: Entrusted Loans(May 2, 2014)” reports that from 2004 to 2013, around 20\% of entrusted loans went to the property sector.

\textsuperscript{5}In November 2008, the Chinese State Council approved an economic stimulus plan to invest 4 trillion RMB by 2010. The central government provides 1.2 trillion RMB, the rest of the fund is provided by the provincial budget and local government. Another component of the policy is a set of credit expansion, including lower the reserve requirement of banks and lower the benchmark lending rate. Bai, Hsieh, and Song (2016) document that the required reserve ratios were adjusted from 17.5\% to 16\% and from 16.5\% to 13.5\% for large and small financial institutions, respectively, in the fourth quarter of 2008. The official benchmark one-year loan rate dropped from 7.2\% to 5.31\%. More studies on the policy can be found from Bai et al. (2016); Cong, Gao, Ponticelli, and Yang (2018); Chen, He, and Liu (2018)
New bank loans suddenly increased from 15% of GDP in 2008 to 28% of GDP in 2009. The funding cost of banks are significantly reduced during the stimulus plan, at the same time, new entrusted loans also started to grow, as shown in Figure 1.

Figure 1: New entrusted loans to GDP, Data source: CEIC

Secondly, this paper provides policy implications of the impact of entrusted loans. Although shadow banks can not monitor borrowing firms as well as banks, the total welfare is actually improved when allowing firms to conduct entrusted loans. However, the benefit of entrusted loans to total welfare is harmed if allowing firms to lend to restricted industries which are prohibited from bank loans. Therefore, entrusted loans should not be totally cracked down, but regulations should prevent the loans from flowing to restricted industries. On the other hand, entrusted loans could threaten the financial stability. Firms which participate in entrusted loans have higher default risk compared to the situation should entrusted loans been prohibited. Thus, regulators should consider corresponding policies to prevent loan default.

This paper also provides interpretations for some stylized facts. First, the result is consistent with the findings from Allen, Qian, Tu, and Yu (2019) that the lenders of entrusted
loans are high-capitalized firms that can obtain cheap bank loans, and the borrowers are medium-capitalized firms. Second, Allen, Qian, Tu, and Yu (2019) and He, Lu, and Ongena (2016) find that lenders of entrusted loans usually charge very high interest rates to non-affiliated borrowing firms. The model assumes that the lenders do not monitor the borrowers, which can be interpreted as the lenders and borrowers being non-affiliated, so the lenders do not have the knowledge or ability to monitor the borrowers. The lack of monitoring makes the borrowers riskier, and thus the lenders have to charge high interest rates. Third, the lenders simultaneously borrow from banks and re-lend to the borrowers in the model. This “borrow in order to lend” activity is consistent with findings in Shin and Zhao (2013). Forth, Allen, Qian, Tu, and Yu (2019) and He, Lu, and Ongena (2016) show that the average abnormal return for the lenders of non-affiliated entrusted loans is negative after the entrusted loans announcement, indicating a negative outlook of lenders. This paper shows that entrusted loans indeed destroy firm value and severely increase firm default risk because entrepreneurs do not behave diligently.

The rest of this paper is organized as follows. Section 2 relates this paper to previous research. In Section 3, I describe the model and solve the benchmark. In Section 4, I introduce the shadow banking sector and derive the equilibrium results. I discuss welfare implications and list the theoretical predictions in Section 5. In Section 6, I give one extension of the model. Section 7 concludes the paper. All proofs are given in the appendix.

2 Literature Review

To the best of my knowledge, this is the first theoretical model on entrusted loans. The most related paper is that of Allen, Qian, Tu, and Yu (2019), in which the authors examine

Allen, Qian, Tu, and Yu (2019) find that lenders are well-capitalized firms, the average debt ratio is 16.7% for non-affiliated lenders, 23.5% for affiliated lenders. The borrowers are medium-sized companies with a median value of 0.4 billion RMB.
entrusted loans with transaction-level data. They divide entrusted loans into affiliated loans and non-affiliated loans on the basis of firm relationships. They show that the loan interest rates incorporate both fundamental and informational risks. Another related paper is He, Lu, and Ongena (2016), in which the authors assess the valuation effects of entrusted loans by measuring the abnormal returns of stock prices after entrusted loan announcements. Chen, Ren, and Zha (2018) focus on the impact of monetary policy on entrusted loans and emphasize the different roles of state-owned banks and non-state-owned banks that work as trustees of entrusted loans. Yu, Lee, and Fok (2015) find a significant and positive correlation between high-interest entrusted loans and firms’ last period cash holdings, suggesting that firms would reserve cash for speculative entrusted loan activity.

This paper is also related to the theoretical studies on credit rationing and constrained bank lending. My paper belongs to one strand of papers featuring moral hazard problem, including Besanko and Kanatas (1993), Petersen and Rajan (1995), Holmstrom and Tirole (1997), Repullo and Suarez (2000) and Allen, Carletti, and Marquez (2011). Another strand of paper focus on asymmetric information, such as Stiglitz and Weiss (1981), Besanko and Thakor (1987a), Besanko and Thakor (1987b). Different from the above papers, I introduce the channel of shadow banking, and focus on the endogenous emergence of shadow banking in the model. Moreover, by focusing on the moral hazard problem, I can study the impact of shadow banking on entrepreneurial behavior and real efficiency of firms.

This paper also belongs to the branch of literature on China’s shadow banking. Ehlers et al. (2018) and Hachem (2018) provide very detailed documentation of the whole shadow banking sector in China. Acharya, Qian, and Yang (2017) document another large component of shadow banking in China, wealth management products (WMP). They show that when small and medium-sized banks face strong competition from the ‘Big Four’ banks to acquire deposits, they significantly increased the issuance of WMPs after the financial crisis. Dang, Wang, and Yao (2014) explain the growth of shadow banking with “information sen-
sitivity”, which is a measure of tail risks. The rise of shadow banking is associated with the asymmetric perception of information sensitivity from banks, shadow banks, and investors. Both Hachem and Song (2017) and Chen et al. (2018) link shadow banking activities to fiscal policy. Hachem and Song (2017) argue that the shadow banking system is related to the competition between small and medium-sized banks and the ‘Big Four’ banks. Shadow banking is an unintended consequence of the liquidity rules. Chen, He, and Liu (2018) show that the growth of shadow banking after 2012 is a hangover effect of the “4 trillion stimulus package”.

Other theoretical studies regard shadow banking as a problem of regulatory arbitrage. Buchak, Matvos, Piskorski, and Seru (2018) find that the increased regulatory requirements of banks account for 60% of shadow bank growth, and 30% of the growth comes from the development of financial technology. Farhi and Tirole (2017) show that financial intermediaries can migrate to shadow banking in response to regulatory requirements, and they also study optimal regulation in the presence of the shadow banking sector. In my paper, shadow banking is not caused by the regulatory burden of traditional banks, and the banks do not have capital constraints in my model. Shadow banking is a market reaction to the loosen monetary policy. Moreover, entrusted loans are very similar to trade credit as documented by Petersen and Rajan (1997) and Biais and Gollier (1997) in the sense that both are intercorporate financing. However, entrusted loans are not necessarily granted between suppliers and customers, and there are no goods or service purchases involved in entrusted loans.

3 Model

The model is built on Besanko and Kanatas (1993) with two innovations. First, entrusted loans are introduced as an outside option, besides firm’s own investment project. Firms have to decide whether to participate in entrusted loans or not. Second, firms have different
capitalizations, as in Holmstrom and Tirole (1997). When taking risky investment projects, firms invest all own capital and borrow the remaining from external sources. The introduction of firm heterogeneity allows me to study the firm characteristics that are engaging in entrusted loans.

The model contains two types of agents: firms and banks. All agents are assumed to be risk neutral. There are two dates in the model: date 0 and 1. At date 0, each firm tries to locate funding for a risky investment project. At date 1, returns from projects are realized, firms are liquidated, and banks are paid off.

There is a continuum of firms. Each firm is characterized by the capital \( w \in [0, 1] \) it holds. The set of firms is described by the probability density function \( g(w) \). All firms have access to the same technology and undertake the same risky project in the model. The project requires investment 1 at date 0. If \( w < 1 \), the firm needs at least \( 1 - w \) external financing to start the project. At date 1, the project generates return either \( Q > 1 \) with probability \( p \) (success) or 0 with probability \( 1 - p \) (failure).

Firms are run by entrepreneurs. The probability that the project succeeds depends on the effort the entrepreneur spends on the project. The cost of effort required to achieve a success probability \( p \) is \( \frac{p^2}{\beta} \), where \( \beta \in (0, \frac{1}{2}) \) is the reciprocal of the marginal cost of effort. For simplicity, \( p \) is also the total effort from the entrepreneur. If the entrepreneur wants to increase the success probability of the project, she has to expend more effort, which will reduce her utility. Moral hazard arises since the effort of the entrepreneur is un-observable by the market.

When the firm is not able to be self-financed, it has to take a loan from the bank. The function of banks is twofold: providing financial credit to the firm; offering monitoring services to improve entrepreneur performance. Banks have a comparative advantage at monitoring firms, which alleviates the entrepreneur moral hazard problem. Monitoring carries a
cost to the bank, which is assumed to be

\[
C(p - p_0) = \begin{cases} 
(p - p_0)^2/2m, & \text{if } p \geq p_0, \\
0, & \text{if } p < p_0.
\end{cases}
\]

\(p\) is the effort level desired by the bank, \(p_0\) denotes the entrepreneur effort were there no bank monitoring and thus called entrepreneur non-monitoring effort, \(m \in (0, \beta)\) denotes the reciprocal of the marginal cost of monitoring. The bank has advantage in monitoring, the marginal monitoring cost of the bank is thus smaller than the entrepreneur, so \(m < \beta\). Bank monitoring raises the success probability from \(p_0\) to \(p\). In case the bank desired level \(p\) is lower than the entrepreneur non-monitoring effort level \(p_0\), the bank does not monitor and the total effort \(p\) equals to non-monitoring effort \(p_0\).

Without loss of generality, the gross risk-free interest rate is assumed to be one. The funding cost of bank is \(M \geq 1\). \(M\) can be interpreted as the average rate of return of the deposit and equity of banks\(^7\).

### 3.1 Entrepreneur effort and bank monitoring

The timing of the model is as follows. Banks set up loan contracts which specify the loan size \(L_B\) and corresponding repayment \(R_B\). Firms decide how much loans to take and invest in the project. Then entrepreneurs choose effort level \(p_0\) and banks choose monitoring intensity which equals to \(p - p_0\).

By back-ward induction, I first solve the optimal entrepreneur non-monitoring effort \(p_0\) and bank monitoring \(p - p_0\) for given bank loan size \(L_B\) and repayment \(R_B\). The entrepreneur

\(^7\)\(M\) can also be an indication of market power of the bank, as in Petersen and Rajan (1995). Highly concentrated banking sector would require a high rate of return, while perfectly competitive banking sector requires a rate of return \(M = 1\).
chooses the effort to maximize her expected utility that is given by

\[ p_0^* = \arg \max_p p(Q - R_B) - \frac{p^2}{2\beta} + L_B + w - 1 = \min\{\beta(Q - R_B), 1\}. \]

For ease of exposition, two assumptions are made:

**Assumption 1.** \( \beta Q \leq 1 \).

**Assumption 2.** \( \frac{\beta Q^2}{2} - 1 > 0 \).

Assumption 1 states that the first-best level of effort, equivalently the success probability is not bigger than 1. Under this assumption, the non-monitoring effort is thus

\[ p_0^* = \beta(Q - R_B). \] (3.1)

Assumption 2 ensures that the maximized NPV of self-financed project is strictly positive, otherwise the entrepreneur will not undertake the investment project. Assumption 1 and 2 together require that \( \sqrt{\frac{2}{3}} < Q \leq \frac{1}{\beta} \).

Banks do not commit to monitor firms, and bank monitoring is also unobservable to the market, so banks also incur moral hazard problem. The bank monitoring intensity is chosen to maximize the expected profit of the bank, for given non-monitoring effort \( p_0^* \), that is

\[ \max_p pR_B - \frac{(p - p_0^*)^2}{2m}. \]

The optimal bank monitoring is derived as

\[ p^* - p_0^* = mR_B. \] (3.2)

Equation (3.1) and (3.2) show that, both entrepreneur non-monitoring effort and bank monitoring expenditure are proportional to their payoffs. The payoff to entrepreneur is the total
investment output deducted by repayment to the bank, so entrepreneurs of high-capitalized firms are willing to exert more effort. Conversely, bank monitoring intensity is higher for low-capitalized firms. Overall, the total optimal entrepreneur effort is given as

\[ p^* = \beta Q + (m - \beta)R_B. \]

3.2 Benchmark: bank lending with no shadow banking

For comparison, I start with the benchmark in which shadow banking is prohibited. In the benchmark, banks are the only outside financing source for all firms. Then the equilibrium bank lending problem is equivalent to the question in Besanko and Kanatas (1993), except that firms have different initial capital.

The equilibrium loan size \( L_B \) and bank repayment \( R_B \) are chosen to maximize the entrepreneur’s expected utility, subject to three constraints: participation constraint of the bank, and two participation constraints of firms. The optimization problem of entrepreneur is given as

\[
\max_{L_B, R_B} U = p^*(Q - R_B) - \frac{p^{*2}}{2\beta} + L_B + w - 1, \tag{3.3}
\]

subject to

\[
p^*R_B - \frac{(p^* - p_0^*)^2}{2m} - ML_B \geq 0, \quad \text{(bank participation constraint)}
\]
\[L_B + w \geq 1, \quad \text{(firm participation constraint 1)}
\]
\[R_B \leq Q. \quad \text{(firm participation constraint 2)}
\]

The entrepreneur utility is composed by three parts: the expected firm profit \( p^*(Q - R_B) \); the cost of entrepreneur effort \( -\frac{p^{*2}}{2\beta} \); and retained capital \( L_B + w - 1 \). The participation constraint of bank states that the expected profit of the bank should be non-negative. Firm
participation constraint 1 states that total capital of each firm is enough to invest in the project. If not, the firm will return the loan \( L_B \) to the bank and give up the project. Participation constraint 2 ensures that the payoff to the entrepreneur should be non-negative to incentivize the expenditure of effort on the project. A third assumption is further required: 

**Assumption 3.** \( m < \frac{\beta}{2M} \).

Assumption 3 ensures that the second order condition of the optimization problem is negative, so there exists a solution to the maximization problem.

Solving the optimization problem of the entrepreneur yields the following results.

**Proposition 1.** In equilibrium, only firms with capital \( w \in [w^N, 1) \) borrow from banks, where

\[
\begin{align*}
  w^N &\equiv 1 - \frac{\beta^2 Q^2}{2M(2\beta - m)}.
\end{align*}
\]  

(3.4)

The loan size is

\[
\begin{align*}
  L^*_B &= 1 - w \equiv L^{OB}_B.
\end{align*}
\]  

(3.5)

The required repayment to banks is

\[
\begin{align*}
  R^*_B &= \frac{\beta Q - \sqrt{\beta^2 Q^2 - 2M(2\beta - m)(1 - w)}}{2\beta - m} \equiv R^{OB}_B.
\end{align*}
\]  

(3.6)

Firms with capital \( w \in [0, w^N) \) are not able to obtain bank loans.

Note that \( w^N \leq 0 \) when \( Q^2 \geq \frac{2}{\beta}(2 - \frac{m}{\beta})M \), that is, the profitability of investment project is high enough such that all firms are financed. While \( w^N > 0 \) when \( Q^2 < \frac{2}{\beta}(2 - \frac{m}{\beta})M \), then firms with initial capital \( w < w^N \) are not able to raise funding from banks. Moreover, \( w^N \) increases with \( M \). That is, when the funding cost of bank increases, there will be less firms receiving bank loans and credit rationing will become more severe.

Equation (3.5) states that entrepreneur only issues the minimum amount of debt that is just enough to finance the project. Because the marginal cost of debt increases with the
size of the loan, so it’s getting even more expensive to raise one more unit of bank loan. Moreover, the bank loan interest rate, given in equation (3.7), is higher than the marginal return of retained capital, which is 1. So there is no need to hold extra capital.

\[
\frac{R_{OB}^{OB}}{L_{OB}^{OB}} = \frac{\beta Q - \sqrt{\beta^2 Q^2 - 2M(2\beta - m)(1 - w)}}{(2\beta - m)(1 - w)}. \tag{3.7}
\]

The interest rate increases with the funding cost \( M \) of the bank. If the bank has to repay high cost to the creditors, the bank correspondingly charges high cost to its debt borrowers. The interest rate decreases in firm capitalization \( w \). That is, banks charge lower interest rate on high-capitalized firms. There are two reasons. First, banks price loans by incorporating risks of investment projects. When the firm has less own capital to invest in the risky project, the entrepreneur’s incentive to exert effort declines, which results in a high probability of investment failure. Banks thus charge a higher interest rate to compensate the risk of default. Secondly, banks also incorporate the monitoring cost into the interest rate. Bank’s monitoring cost is a convex function of expenditure; the expenditure is proportional to the exposure \( R_B \) as shown by equation (3.2). So the average monitoring cost increases in loan demand. Low capitalized firms take a large amount of bank loans and thus incur higher monitoring cost. Overall, banks charge a high interest rate on the low capitalized firm.

4 Shadow banking

In this section, shadow banking is introduced. Since shadow banking in this paper refers to inter-corporate entrusted loans, there are still two types of agents in the model: firms and banks. However, among the firms, some endogenously form shadow banks and provide financing to other firms in the form of entrusted loans. To differentiate, the former firms are called lending firms or shadow banks and the latter ones are called borrowing firms.
The first question is to determine characteristics of firms participating in shadow banking. Because all firms are in short of capital for the investment project, the way lending firms can finance borrowing firms is through over-borrowing from banks. Therefore, firms’ decision of whether participating in shadow banking as lenders is equivalent to whether over-borrowing from banks.

Proposition 1 shows that firms optimally choose to borrow the minimum amount of bank loan which is just enough to finance the risky project in the benchmark. That is because retained capital only generates a return of risk-free rate, which is lower than the interest rate of bank loan. However, once firms have access to another option which generates a much higher return, firms might be willing to over-borrow and invest extra capital in the new option. That is the way I model the emergence of shadow banking. Specifically, in the first step, shadow banking is simply formalized by an outside investment option which generates an expected rate of return $r_s$. Firms can invest extra capital $L_B + w - 1$ in the outside option. The value of $r_s$ is larger than 1. Otherwise, firms are better off depositing the extra capital and earn the rate of return 1.

The lenders are firms that choose to over-borrow bank loans. The borrowers are merely firms choosing to take shadow bank loans. After lenders are determined, and shadow banks are formed. All firms except lenders thereby have both banks and shadow banks as outside financing sources. These firms then optimally choose to borrow from both sources.

The existence of shadow banking depends on the rate $r_s$. The equilibrium value of $r_s$ is determined by market clearing condition. That is, the total supply of entrusted loans should be equal to the total demand. Total supply is the aggregate extra capital from lenders, and total demand is the aggregate demand from borrowers. The existence of shadow banking is therefore equivalent to having a solution $r_s > 1$ which solves the market clearing condition.

---

8Purely conducting lending activities without investing in own business is not allowed, otherwise, the firm becomes a bank and has to follow all the requirements of bank and function like a bank.
Although the emergence and lending of shadow banks are sequential, I take the two procedures as simultaneous. Essentially, the two procedures are both external fund-raising activities of firms. In a static model, I can study how banks and shadow banks interact when providing loans to firms. The timeline is given in Figure 2.

![Timeline Diagram](image)

**Figure 2: Time line**

4.1 **Endogenous emergence of shadow bank**

Similar to the benchmark, except that return from extra capital is $r_s$ instead of 1. The bank loan contract $\{L_B, R_B\}$ is set so as to maximize the entrepreneur’s utility, which is given as

$$
\max_{L_B, R_B} U = p^*(Q - R_B) - \frac{p_s^2}{2 \beta} + r_s(L_B + w - 1),
$$

(4.1)
subject to

\[ p^* R_B - \frac{(p^* - p_0^*)^2}{2m} - ML_B \geq 0, \quad (4.2) \]

\[ L_B + w \geq 1, \quad (4.3) \]

\[ R_B \leq Q. \quad (4.4) \]

Solving the optimization problem of entrepreneur, yields the following results.

**Lemma 1.** For given \( r_s \), firms with capital \( w \) over-borrow from banks when

\[ r_s \geq \tilde{r}_s(w) = M \left[ 1 + \frac{m^2 + \beta - m}{2\beta - m} \left( \frac{\beta Q}{\sqrt{\beta^2 Q^2 - 2M(2\beta - m)(1 - w)}} - 1 \right) \right]. \quad (4.5) \]

First, note that \( \tilde{r}_s > M \), that is, firms (shadow banks) require higher rate of return than banks, because the capital that shadow banks lend out are actually loans raised from banks. Secondly, \( \tilde{r}_s \) is firm specific, which is a decreasing function of \( w \). It implies that firms that hold high capitalization tend to over-borrow and thus become lenders of shadow banking. Inverse solution of (4.5) shows that lenders of shadow banking have initial capital \( w \geq 1 - \frac{\beta^2 Q^2}{2M} \left( \frac{r_s - m}{\beta - m} \right) \left[ \sqrt{\beta^2 Q^2 - 2M(2\beta - m)(1 - w)} \right]^2 \). This threshold value decreases in \( r_s \). That is, high expected rate of return of shadow banking drives up the firm participations. The following proposition states the optimal loan contract.

**Proposition 2.** For given \( r_s \), firms with initial capital \( w \geq 1 - L_B^s \) over-borrow and form shadow banks with extra capital. Bank loan size and required repayment are

\[ L_B^* = \frac{\beta^2 Q^2}{2M} \left( \frac{r_s M}{M} - 1 \right) \left[ \frac{(2\beta - m) r_s M + 2m^2}{\beta} - m \right] \equiv L_B^S, \]

\[ R_B^* = \frac{\beta Q (r_s M - 1)}{(2\beta - m) \frac{r_s M}{M} + \frac{m^2}{\beta} - \beta} \equiv R_B^S. \]
When entrepreneur chooses to over-borrow, the optimal loan size $L_B^S$ does not depend on firm capitalization $w$ anymore. Instead, all lending firms take the same amount of bank loan. The firm that is indifferent from over-borrowing or not has capital $w = 1 - L_B^S$. Note that the expected profit from outside option $r_s(L_B + w - 1)$ is a linear function of loan size $L_B$, so firms are willing to take as much bank loan as possible. However, taking more debt means entrepreneurs' payoff from investment project will decline because high bank repayment, and further discouraging entrepreneur from behaving diligently, worsening the expected profit of firm’s investment project. Therefore, the optimal amount of bank loan is determined such that the return from shadow banking compensate the reduction of entrepreneur utility.

The required repayment $R_B$ and bank loan interest rate $R_{BL}$ are also independent of $w$. Each lending firm has the same bank loan contract, each entrepreneur pays the same amount of effort and bank expends the same monitoring. Entrepreneurs of lending firms obtain the same amount of utility, which equals to the utility of entrepreneur of the boundary firm $w = 1 - L_B^S$ could get.

### 4.2 Shadow bank lending

In the presence of shadow banks, all firms except lending firms optimally choose loans from banks and shadow banks. The difference between banks and shadow banks are twofold: first, shadow banks do not provide monitoring service. Second, shadow banks require a rate of return to be at least $r_s > 1$, while the funding cost of bank is $M$. Since shadow banks compete with formal banks in providing capital to firms, the rate of return from shadow banks is squeezed to $r_s$. Similarly, the optimal bank loan contract $\{L_B, R_B\}$ and shadow bank loan contract $\{L_S, R_S\}$ are selected to maximize entrepreneur’s utility, which is given as

$$\max_{LB, RB, LS, RS} U = p^*(Q - R_B - R_S) - \frac{p^{*2}}{2\beta} + L_B + L_S + w - 1,$$  

(4.6)
subject to
\[ p^* R_B - \frac{(p^* - p^*_0)^2}{2m} - M L_B \geq 0, \quad (4.7) \]
\[ p^* R_S - r_s L_S = 0, \quad (4.8) \]
\[ L_B + L_S + w - 1 \geq 0, \quad (4.9) \]
\[ R_B + R_S \leq Q. \quad (4.10) \]

The inclusion of shadow banks bring several changes. First, entrepreneur’s optimal non-monitoring effort becomes \( p^*_0 = \beta(Q - R_B - R_S) \). The optimal bank monitoring \( p^* - p^*_0 \) keeps but the total entrepreneur effort becomes \( p^* = \beta(Q - R_S) + (m - \beta)R_B \). Second, if firms decide to take both bank loan and shadow bank loan, that is, \( L_B > 0 \) and \( L_S > 0 \), total loans should be enough to start the project, as indicated by (4.9). Meanwhile, total loan repayments \( R_B + R_S \) should not exceed the investment return \( Q \).

The rate of return from retained capital \( L_B + L_S + w - 1 \) is merely 1, because there is no second-tier shadow banking opportunity. The intuitions are as follows. In one case, suppose firm A over-borrows from both bank and shadow bank, and A wants to lend out extra capital to firm B. Within the capital A obtained, shadow bank requires a rate of return \( r_s \), but when A lends to B, A expects a rate of return which is higher than \( r_s \). However, in this case, B could borrow directly from shadow bank, rather than A. So A loses the borrower because of competition from shadow bank. In another case, suppose firm A only borrows from bank and over-borrows, then A lends to B. In this case, A must have initial capital \( w \geq 1 - L_S^B \), which means A belongs to the group of lenders derived in section 4.1. Therefore, there is no second-tier shadow banking opportunity and the rate of return from retained capital is 1. Solving the entrepreneur’s optimization problem gives the equilibrium of bank and shadow bank lending.

**Proposition 3.** Depending on the value of the expected rate of return of shadow banking \( r_s \),
optimal financing strategies of firms are given as

1. If \( r_s \geq \min \{ \frac{M}{3} (1 + \frac{m}{\beta} + 2\sqrt{1 - \frac{m}{\beta} + \frac{m^2}{\beta^2}}), \frac{\frac{m}{\beta} (\frac{m^2}{\beta^2} - m + 1) M + m - 2}{\frac{m^2}{\beta^2} + 2m - 2} \} \equiv \bar{r}_s \), all firms with capital \( w \in [w^N, 1 - L^S_B] \) take only bank loans. Each firm borrows the minimum amount \( L^*_B = 1 - w \), and bank requires return

\[
R^*_B = \frac{\beta Q - \sqrt{\beta^2 Q^2 - 2M(2\beta - m)(1 - w)}}{2\beta - m} \equiv R^0_B. \quad (4.11)
\]

Firms with capital \( w < w^N \) can not get any financing.

2. If \( r_s < \bar{r}_s \), firms with capital \( w \in [w_1, w_2] \) borrow from both bank and shadow bank, where

\[
w_1 = 1 - \frac{2\beta Q R_1 - (2\beta - m)R_1^2}{2M},
\]

\[
w_2 = 1 - \frac{2\beta Q R_2 - (2\beta - m)R_2^2}{2M},
\]

and

\[
R_1 = \frac{Q (m^2 + \beta^2 - m\beta) M + \beta(2\beta - m)r_s + \beta(2m - 3\beta) + \sqrt{\Delta}}{2 (m^2 + \beta^2 - m\beta)((1 - \frac{m}{\beta}) M + r_s) - (\beta - m)(2\beta - m)},
\]

\[
R_2 = \frac{Q (m^2 + \beta^2 - m\beta) M + \beta(2\beta - m)r_s + \beta(2m - 3\beta) - \sqrt{\Delta}}{2 (m^2 + \beta^2 - m\beta)((1 - \frac{m}{\beta}) M + r_s) - (\beta - m)(2\beta - m)},
\]

\[
\Delta = (m^2 - m\beta + \beta^2) M^2 - 2\beta(2\beta - m)(m^2 - m\beta + \beta^2) M + \beta^4 + m\beta^2 [2(2m - \beta) - 3mr_s] r_s.
\]

The equilibrium bank loan size \( L^*_B \equiv L^M_B \), bank repayment \( R^*_B \equiv R^M_B \), the equilibrium shadow bank loan size \( L^*_S \equiv L^M_S \), shadow bank repayment \( R^*_S \equiv R^M_S \), are solutions to
the equation system

\[
\begin{aligned}
\frac{\partial U}{\partial L_B} &= \frac{\partial U}{\partial L_S} \\
L_B &= \frac{\beta M}{\beta} R_B \left( Q - R_S + \left(\frac{m}{\beta^2} - 1\right) R_B \right) \\
L_S &= \frac{\beta r_s}{r_s} R_S \left( Q - R_S + \left(\frac{m}{\beta} - 1\right) R_B \right) \\
L_B + L_S &= 1 - w \\
L_B \geq 0, L_S \geq 0
\end{aligned}
\] (4.12)

Firm with capital \( w \in [w^N, w_1) \) or \( w \in (w_2, 1 - L_B^S] \) borrow only from banks. The bank loan size is \( L_B^* = 1 - w \) and repayment to bank is \( R_B^* = R_B^{OB} \).

### 4.3 The existence of shadow banking

Proposition 3 shows that no firm would borrow from shadow bank if \( r_s \geq \bar{r}_s \), so the equilibrium rate of return \( r_s^* \) can not be larger than \( \bar{r}_s \). When \( r_s < \bar{r}_s \), firms with capital \( w \in (w_1, w_2) \) borrow from shadow banks. Firms with capital \( w \in [1 - L_S^B, 1] \) over-borrow and form shadow banks with their extra capital. The existence of shadow banking is thus equivalent to the existence of a solution \( r_s \in (M, \bar{r}_s) \) which satisfies the market clearing condition: total supply of shadow bank capital equals to total demand of shadow bank loans.

The supply of shadow bank are the aggregate extra capital of lenders derived in section 4.1, which is

\[
TS = \int_{1-L_B^S}^{1} (L_B^S + w - 1) g(w) dw.
\]

The demand of shadow bank loan is the aggregate demand of the borrowers derived in section 4.2, that is,

\[
TD = \int_{w_1}^{w_2} L_S^M g(w) dw.
\]

21
The following proposition proves the existence of shadow banking.

**Proposition 4.** The existence of shadow banking depends on the cost of capital $M$ of banks:

1. If $M \geq \frac{1}{1 - \frac{m}{\beta^2}}$, there does not exist shadow banking sector.
2. If $M < \frac{1}{1 - \frac{m}{\beta^2}}$, there exists a shadow banking sector, with expected rate of return $r_s \in (M, \bar{r}_s)$ that solves $TD = TS$.

When the funding cost of banks is larger than $\frac{1}{1 - \frac{m}{\beta^2}}$, a shadow banking sector does not exist. In this situation, $M \geq \bar{r}_s$, that is, banks require a high rate of return than shadow banks. While the existence of shadow banking requires $M \leq r_s < \bar{r}_s$. When shadow banks re-lend bank loans, shadow banks require an even higher rate of return. But when the required rate of return from shadow banks is too large, no firms would borrow from shadow banks. Therefore, shadow banking does not exist when the funding cost of banks is large.

Not only the existence of shadow banking depends on the funding cost $M$ of the banks, but also the equilibrium rate of return $r_s^*$ from shadow banking. Further analysis of the relationship between $r_s^*$ and $M$ yields the following result.

**Corollary 1.** The equilibrium expected rate of return $r_s^*$ of shadow banks is an increasing function of the funding cost $M$ of banks.

When banks require a high rate of return from issuing loans to firms, highly capitalized firms which issue loans to other firms via shadow banking also require a high rate of return $r_s$ from loans. When the funding cost $M$ of banks reduces, each bank reduces the interest rate of the loans to firms. The reduction of bank loan interest rate induces more firms to participate in shadow banking by re-lending bank loans to other firms. So the increased funding cost of banks drive up competition among shadow banks and lower the expected rate of return $r_s$ of shadow banks.

I solve the model numerically by setting $Q = 2.2$, $\beta = 0.45$, $m = 0.15$ and assuming $w$ follows uniform distribution on $[0, 1]$. The funding cost $M$ of banks are given three values: 1.25, 1.05 and 1 to do comparative analysis. The bank and shadow bank lending
are demonstrated in Figure 3. The gray region plots the initial capital of firms, green region plots the bank loans, red region plots the shadow bank loans (entrusted loans).

First, note that credit rationing exists in all three cases: firms with capital $w < w^N$ cannot get any kind of financing. Second, the bank and shadow bank lending vary with the funding cost $M$ of banks. When the funding cost of bank is high, e.g. $M = 1.25$, shadow banking does not exist. Each firm with the capital $w \geq w^N$ only borrows the minimum amount from the bank. When $M$ reduces, shadow bank appears. The firms with the capital $w \geq 1 - L^S_B$ over-borrow and become lenders of shadow banking, as plotted by the red stripes. These firms then re-lend the extra bank loans to firms with capital $w \in (w_1, w_2)$. The other firms only borrow the minimum amount from the bank. Moreover, when $M$ reduces from 1.05 to 1, that is, the funding cost of banks decreases, the shadow banking sector expands. More firms participate in shadow banking activity: $w_1$ reduces, $w_2$ increases and $1 - L^S_B$ reduces, which can be seen from the expansion of red region in Figure 3.

Figure 3 also show that the demand of shadow bank loan $L^M_S$ (red region) is a concave function of firm capitalization $w$. When firms simultaneously borrow from banks and
shadow banks, banks provide monitoring which guarantees a high level of success probability. Shadow banks free ride banks’ monitoring and thus charge a lower interest rate because of no monitoring. So entrepreneurs are willing to take shadow bank loans and thereby reduce the demand of bank loans. However, the increase of shadow bank loan reduces the entrepreneur effort $p$ because the bank monitoring declines. So entrepreneurs have to balance the tradeoff between increased payoff and reduced success probability. For firms with low capitalization $w \in [w^N, w_1)$, the reduction in $p$ is much larger than the increase in entrepreneur payoff, so entrepreneurs do not take shadow bank loans. For firms with high capitalization $w \in (w_2, 1 - L_S^B]$, the increase in entrepreneur payoff is smaller because these firms already enjoy low bank interest rate, the rate of shadow bank loan is not as competitive as the rate of bank loan, so entrepreneurs choose not to take shadow bank loans. For firms which take both bank loans and shadow bank loans, the two loan demands are determined that the marginal profit of bank loans and shadow bank loans be equal, as shown by equation (4.12).

I then compute and plot the interest rates of bank loan and shadow bank loan (only for the case $M = 1$) in Figure 4. Figure (4a) shows the bank loan rates with and without shadow banking (benchmark), as well as the rate of shadow bank loan. The green line represents the equilibrium bank lending rate, the gray dotted line plots the bank lending rate in benchmark, the blue line plots the shadow bank lending rate. Figure (4b) gives a clear illustration of how the bank lending rate changes compared with the benchmark. Firms with the capital $w \in [1 - L_S^B, 1]$ are lending firms, firms with capital $w \in [w_1, w_2]$ are borrowing firms of shadow banking, other firms do not participate in shadow banking.

First, note that both banks and shadow banks charge lower interest rate to high-capitalized firms and higher interest rate to medium-capitalized firms. The interest rates are mainly determined by the default risk caused by moral hazard problem. Next, shadow bank interest rates are smaller than bank interest rates. The difference results from the monitoring specialty of banks. Banks monitor firms by paying a convex cost, so banks incorporate the
The lending rates

\[ R_B = \frac{1}{p} \left( M + \frac{(p - p_0)^2}{2mL_B} \right). \]

Shadow banks do not monitor firms, so shadow banks can charge a lower rate than bank lending rate, as given by (4.8):

\[ R_S = \frac{1}{p} r_s. \]

Since it is costly to monitor firms, the average monitoring cost increases in firm capitalization \( w \), so is the difference between bank and shadow bank lending rates, as demonstrated by the green line and the blue line.

Moreover, in the presence of shadow banking, banks increase the lending rate to both lending firms and borrowing firms, which is illustrated in Figure 4b. The adjustment is also positively correlated with the involvement of shadow banking activities. That is because participating of shadow banking discourages entrepreneur from performing diligently. The
default risk \((1 - p)\) of both lenders and borrowers thus increase, so banks incorporate the high risk into the interest rates.

5 Welfare Analysis

In this section, I study the impact of shadow banking on three aspects: the total welfare of banks and firms; the value of firms; and the real efficiency. Because the endogenously emerged shadow banking is an unintended consequence of loosen monetary policy, I first study the how the reduced funding cost of banking sector would affect the social welfare in the benchmark where shadow banking is assumed to be non-exist. Then I compare the equilibrium results in the presence of shadow banking with the benchmark, to learn how shadow banking affects the social welfare.

5.1 Total welfare

I first compute the total welfare of the banks and firms in the benchmark where shadow banking is prohibited and study how the total welfare varies when the funding cost \(M\) changes. Then I compute the total welfare of the banks and firms with shadow banking sector, and compare the total welfare to the benchmark, keeping the same level of funding cost \(M\).

In the benchmark, there does not exist a shadow banking sector, either because the funding cost \(M\) of banks is large, or shadow banking is assumed to be prohibited. In this situation, banks only issue loans to firms with initial \(w \geq w^N\), as described in Proposition 1. Each bank issues \(L^{OB}_B\) amount of loan to each firm and requiring a rate of return \(M\) for each unit of loan. The total welfare of banks is given as

\[
TWB^{OB} = \int_{w^N}^{1} \left[ p^* R^{OB}_B - \frac{(p^* - p^*_0)^2}{2m} - M L^{OB}_B \right] g(w)dw = 0. \tag{5.1}
\]
The total welfare of firms is calculated as

\[ TWF^{OB} = \int_{w^N}^{1} \left[ p(Q - R_B^{OB}) - \frac{p^2}{2\beta} + L_B^{OB} + w - 1 \right] g(w) dw. \quad (5.2) \]

Replacing the equilibrium bank loan \( L_B^{OB} \) and sum up the total welfare of banks and firms gives the total welfare as

\[ TW^{OB} = \frac{\beta}{2} \int_{w^N}^{1} \left[ Q^2 - 2QR_B^{OB} + \frac{\beta^2 - m^2}{\beta^2} (R_B^{OB})^2 \right] g(w) dw, \quad (5.3) \]

where \( R_B^{OB} \) is the equilibrium bank loan return which is derived from Proposition 1.

The bank funding cost \( M \) affects the total welfare \( TW^{OB} \) through two channels: the credit rationing threshold \( w^N \) and the equilibrium bank return \( R_B^{OB} \). The effect of \( M \) on total welfare \( TW^{OB} \) can be derived by taking the derivatives, and the following results apply.

**Lemma 2.** Conditional on non-existence of shadow banking, the total welfare of banks and firms increases when the funding cost of banks decreases. That is, \( \frac{\partial TW^{OB}}{\partial M} < 0 \).

Reducing the funding cost of banks alleviates the credit rationing problem, so more firms are able to raise bank loans and conduct investment project. Moreover, banks correspondingly reduce the interest rate, so external financing becomes cheaper for firms. Overall, the total welfare is improved under a loosen monetary policy.

Next, I study how the existence of shadow banking affect the total welfare. In the presence of shadow banking, two types of firms are affected: highly capitalized firm with capital \( w \in [1 - L_B^S, 1] \) which become the lender of entrusted loans, and medium capitalized firms with capital \( w \in [w_1, w_2] \) which become the borrowers of entrusted loans. The welfare of banks is again zero. The welfare of firms is

\[ TWF^S = \int_{w^N}^{w_1} U_F(L_B^{OB}) g(w) dw + \int_{w_1}^{w_2} U_F(L_B^M) g(w) dw + \int_{w_2}^{1-L_B^S} U_F(L_B^{OB}) g(w) dw + \int_{1-L_B^S}^{1} U_F(L_B^S) g(w) dw, \]
where $U_F$ is the utility of firm. The total welfare with a shadow banking sector is the sum of welfare of banks and welfare of firms:

$$TW^S = TW^BS + TW^FS.$$  \hspace{1cm} (5.4)

Comparing the difference between the total welfare $TW^{OB}$ without shadow banking sector to that $TW^S$ with shadow banking sector, yields the following results.

**Proposition 5.** The total welfare with a shadow banking sector is larger than that without a shadow banking sector, that is, $TW^S > TW^{OB}$.

Shadow banking benefits the economy by improving the social welfare of banks and firms. Comparing the benchmark result in proposition 1 with the equilibrium result derived in sector 4, one can see that the threshold value of credit rationing $w^N$ is independent of the existence of shadow banking. That is, the total external financing is the same in both cases, so the difference between the total welfare is not due to the availability of credit.

The role of shadow banks (lending firms) in this paper is transmitting bank loans to borrowing firms. Although the total amount of bank loans to both types of firms do not change, more loans flow to lending firms and in turn be transmitted to borrowing firms. The difference between banks and shadow banks is that shadow banks do not monitor the entrepreneurs. So monitoring flows away during the transmission of loans from lending firms to borrowing firms. Even though the lending firms receive more bank monitoring because they over-borrow, the increased cost of monitoring on lending firms is smaller than the reduced cost of monitoring on borrowing firms. That is because the marginal monitoring cost on lending firms is smaller than the marginal monitoring cost of borrowing firms, due to the entrepreneur moral hazard problem and the convex bank monitoring cost. Overall, the total cost of bank monitoring is reduced in the presence of shadow banking, so the total welfare is improved.
Corollary 2. Conditional on the existence of a shadow banking sector, the total welfare increases when the funding cost of banks reduces, that is, $\frac{\partial TW}{\partial M} < 0$.

In the presence of a shadow banking sector, the total welfare of banks and firms increase when the funding cost of banks reduces. The welfare improvement arises from two channels. On one hand, more firms can obtain bank loans when the funding cost of banks reduces. That is, the credit rationing threshold $w^N$ reduces when $M$ declines. On the other hand, the shadow banking sector expands when the funding cost of banks reduces. As shown in Corollary 1, the equilibrium required rate of return $r^*_{s}$ of shadow banks reduces when $M$ reduces. When $r^*_{s}$ reduces and shadow bank loan thus becomes cheaper, the demand of shadow bank loan increases. In addition, more firms borrow from shadow banks, as seen from Figure 3 that the range of $(w_1, w_2)$ increases when $M$ reduces.

5.2 The value of firm

The value of firms is important because it affects the stock price of the firm. In this paper, the value of each firm is equivalent to the profit $\pi$ of firm’s project, which is different from the utility of the entrepreneur:

$$U = \pi - \frac{p^2}{2\beta}.$$  \hspace{1cm} (5.5)

The entrepreneur’s decision improves her utility, but not necessarily the firm’s profit.

I first calculate the firm’s profit in the benchmark, which is given as

$$\pi_{F}^{OB} = \beta(Q - R_{B}^{OB}) \left[ Q - (1 - \frac{m}{\beta})R_{B}^{OB} \right], \quad w^N \leq w < 1.$$  \hspace{1cm} (5.6)

Then I calculate the firm’s profit in the presence of shadow banking. Shadow banking affects two types of firms: firms with capitalization $w \in (1 - L_{S}^{B}, 1]$ that become the lenders of entrusted loans, and firms with capitalization $w \in (w_1, w_2)$ that become the borrowers. Using results from proposition 2 and 3, firm’s profit in the presence of shadow banking is
computed as

\[
\pi_F = \begin{cases} 
\pi_{OB}^F = \beta \left[ (Q - R_{OB}^F)(Q - (1 - \frac{m}{\beta})R_{OB}^F) \right], & w^N \leq w < w_1 \\
\pi_{OB}^M = \beta \left[ (Q - R_{OB}^M)(Q - (1 - \frac{m}{\beta})R_{OB}^M) \right], & w_1 \leq w < w_2 \\
\pi_{OB}^S = \beta \left[ (Q - R_{OB}^S)(Q - (1 - \frac{m}{\beta})R_{OB}^S) + \frac{r_s}{M} R_{OB}^S(Q - (1 - \frac{m}{\beta})R_{OB}^S) + r_s(w - 1) \right], & 1 - L_B^S \leq w \leq 1 . 
\end{cases}
\]

First, I study the impact of bank funding cost on firm’s value, in the absence of shadow banking.

**Lemma 3.** Conditional on non-existence of shadow banking, the firm’s value is increased when the funding cost of banks reduces. That is, \( \frac{\partial \pi_{OB}^F}{\partial M} < 0 \).

Reducing the funding cost of banks benefit the firms, because the bank charges a lower interest rate to the firms. The firm takes the same investment project but the external financing becomes cheaper, so the firm obtains higher expected profit.

Next, I study how the existence of shadow banking affects firm’s value. \( \pi_F^S \) denotes the value of lending firm of entrusted loans, \( \pi_F^M \) denotes the value of borrowing firm of entrusted loans.

**Proposition 6.** In the presence of shadow banking, both the value of lending firms and borrowing firms are reduced: \( \pi_F^S \leq \pi_{OB}^F \), \( \pi_F^M \leq \pi_{OB}^F \).

The reduction of firms’ values is due to the agency problem. Entrepreneur chooses the debt level to maximize her utility, which equals to firm’s profit minus the cost of entrepreneur effort, as shown by equation (5.5). Since banks also incur moral hazard problem, banks choose the monitoring effort to maximize their profits, instead of entrepreneur utility. So the bank’s monitoring forces entrepreneur to exert more effort than her desired level. More specifically, banks force entrepreneurs to exert effort \( p \), which is higher than entrepreneur’s desired level \( p_0 \). In the presence of entrusted loans, the lending firms replace bank loans
with entrusted loans and avoid the corresponding bank monitoring. The lending firms of entrusted loans take excess loans which result in higher debt level, so the entrepreneurs also behave less diligently because of moral hazard. Overall, by engaging in entrusted loans, entrepreneurs obtain higher welfare, but firms earn lower expected profits as entrepreneurs exert less effort than the benchmark.

I then compute the profits with the numerical results and plot the percentage change of profit \( 100 \times (\pi_F - \pi^{OB}_F)/\pi^{OB}_F \) in Figure 5. The profits are reduced in the presence of entrusted loans. The reduction is very significant for high-capitalized lending firms because these firms invest a large amount in shadow banking. For the borrowers of entrusted loans, the reduction is a concave function of firm capitalization, which is proportional to the demand of entrusted loans of these firms. In other words, firms which participate more in shadow banking lose more profit.

![Figure 5: The difference of firm’s profit with and without entrusted loans](image)

5.3 Real efficiency

The probabilities of success \( p \), also the total efforts from entrepreneur are endogenously determined by the bank loan and shadow bank loan contract. In the absence of shadow
banking, the success probability is

\[ p^{OB} = \beta Q - (\beta - m)R^{OB}_B, \quad w^N \leq w < 1. \]  \tag{5.8}

In the presence of shadow banking, the success probability becomes

\[
p = \begin{cases} 
  p^{OB} = \beta Q - (\beta - m)R^{OB}_B, & w^N \leq w < w_1 \\
  p^M = \beta (Q - R^M_S) - (\beta - m)R^M_B, & w_1 \leq w < w_2 \\
  p^{OB} = \beta Q - (\beta - m)R^{OB}_B, & w_2 \leq w < 1 - L^S_B \\
  p^S = \beta Q - (\beta - m)R^S_B, & 1 - L^S_B \leq w \leq 1
\end{cases} \tag{5.9}
\]

where \( p^M \) denotes the success probability of borrowing firms of entrusted loans, \( p^S \) denotes that of lending firms. I first study the impact of banking competition on the success probability \( p^{OB} \) in the absence of shadow banking, and then study how the existence of shadow banking affects the success probability.

**Corollary 3.** Condition on non-existence of shadow banking, reducing the funding cost of banks increases the real efficiency, that is, \( \frac{\partial p^{OB}}{\partial M} < 0 \). In the presence of shadow banking, the real efficiency is reduced:

\[ p^S < p^{OB}, \quad p^M < p^{OB}. \]

The decline of bank funding cost enhances real efficiency, but the emergence of shadow banking reduces real efficiency. The success probability \( p \) is also the total entrepreneur effort, which is endogenously determined by the loan contract. When the bank requires a high return, the return to entrepreneur becomes low and entrepreneur’s incentive to exert effort on the investment project is reduced, resulting in a low success probability. Low funding cost of banks reduces the bank required return, as indicated by equation (3.6). The return to entrepreneur thus increases and the total entrepreneur effort also increases. So the success
probability $p$ is higher when the funding cost of banks reduces.

The success probability $p$ is plotted in Figure 6. Figure 6a compares the success probability with and without shadow banking, Figure 6b plots the percentage change of $p$ compared with benchmark: $100 \times (p - p^{OB})/p^{OB}$. The total entrepreneur effort declines in the presence of entrusted loans. Similarly, the reduction is positively correlated with the involvement of shadow banking and more significant for the lending firms. However, the reductions are due to different factors for the lending firms and borrowing firms. Note that the total entrepreneur effort is the sum of entrepreneur non-monitoring effort $p_0$ and bank monitoring effort $p - p_0$. The reduction of lenders comes from the decreased entrepreneur non-monitoring effort, while the reduction of borrowers is from the decreased bank monitoring effort. The lenders increase the demand for the bank loan because part of the loan is invested in shadow banking, so the repayment to bank also increase and the investment return to entrepreneur is reduced. Entrepreneur thus reduces the non-monitoring effort since it is proportional to the return: $p_0 = \beta(Q - R_B)$. For the borrowing firms, the demands of bank loans are reduced since part of bank loans are replaced by shadow bank loans. Banks cut the monitoring effort because banks receive a reduced payoff. While shadow banks do not provide monitoring service, the total entrepreneur effort thus is decreased. Taken together, the reduction of entrepreneur effort of both lending firms and borrowing firms makes these firms riskier. Therefore, banks increase the lending rates of these firms, as shown in Figure 4b.

In conclusion, conditional on the non-existence of shadow banking, loosing monetary policy, i.e. reducing the funding cost of banks improves the total welfare of banks and firms, and increases firm’s value and the real efficiency. But reducing the funding cost of banks could trigger the emergence of shadow banking. In the presence of shadow banking, the total welfare of banks and firms are also improved, because the total cost of bank monitoring is reduced. However, firms' values are reduced and real efficiency is harmed. Because
participating in shadow banking activity avoid excess bank monitoring to entrepreneurs, so that the entrepreneurs exert less effort on firms’ investment projects. Therefore, firms’ expected profits and the success probabilities are both reduced.

5.4 Theoretical predictions

The theoretical model has shown that the existence of entrusted loan is related to the funding environment of banking sector. Proposition 4 shows that entrusted loans do not exist when bank funding cost is high. When the funding cost reduces, entrusted loans arise endogenously. A reduction of bank funding cost by exogenous shock also triggers the growth of entrusted loans. So I postulate the first hypothesis:

**Hypothesis 1.** The growth of entrusted loans market is associated with the decline of bank funding cost.

This hypothesis can be tested using the “4 Trillion RMB Stimulus Plan” event in China from 2009-2010. One policy is to lower bank reserve requirements and lower benchmark lending rates of commercial banks. One needs to test that the new issued entrusted
loans are ultimately bank loans that the lending firms borrowed. That is, firms which take significantly more bank loans during the event are more likely to issue entrusted loans.

The next two hypotheses are related to the welfare analysis.

**Hypothesis 2.** *For firms with the same capitalization, the bank lending rates are higher to firms that engage in shadow banking activities, than firms that do not.*

**Hypothesis 3.** *For firms with the same capitalization, the default risk of firms that engage in shadow banking activities is higher than firms that do not.*

Hypothesis 2 is based on proposition 3. Fixing the firm capitalization, the lenders of entrusted loans have higher bank loan demands, the lending firms have to pay a higher interest rate to the bank. By testing whether the bank loan interest rate is increased could verify whether the entrusted loans from the lending firms are naturally the over-borrowed bank loans. Hypothesis 3 tests the negative impact of the entrusted loans to the involved firms. The model studies the efficiency problem from the perspective of entrepreneur moral hazard. Participating in entrusted either deters the entrepreneur’s incentive of exerting effort on firm’s investment project, or reduces the monitoring from banks, both result in higher default risk.

### 6 Extensions

To extend the model, I include restricted industries which refer to the firms that are prohibited from bank lending. Since 2005, the China Banking Regulatory Commission has put a more severe restriction on bank lending to certain industries, including real-estate, core mining, shipbuilders and local government financing platform, etc. The limited access to the bank loan, in turn, creates demand for entrusted loans. Unlike the formal banks, shadow banks are not under regulatory supervision. Firms that can issue entrusted loans fill the gap by lending to restricted industries.
In the model, it is assumed that a proportion $q \in [0, 0.5)$ of firms belong to the restricted industries among all the firms. These firms are not able to obtain bank loans. Note that the firms in restricted industries are different from the firms from un-restricted industries which are credit rationed from banks. Credit rationing in section 3.2 is an endogenous decision of banks not lending to firms with low capitalization. But the inability of restricted firms to raise bank loans is caused by an exogenous policy imposed on the bank, and it is irrelevant from firm capitalization.

The equilibrium derived in section 3 and 4 is a special case when $q = 0$. If $q > 0$, all firms are divided into two types: restricted firms and unrestricted firms. Unrestricted firms behave same as described in section 3 and 4: high-capitalized firms over-borrow from banks, other firms optimally choose to borrow from banks and shadow banks. The restricted firms can only borrow from shadow banks. In equilibrium: the lenders of the entrusted loan are high-capitalized unrestricted firms, and the borrowers are medium-capitalized unrestricted firms and high-capitalized restricted firms.

I start by deriving the shadow bank loan contract of restricted firms. The bank loan and shadow bank loan contracts to unrestricted firms follow the results from section 4. Then I show the existence of shadow banking and derive the equilibrium expected rate of return $r_s^*$ of shadow banks.

For the restricted firms, entrusted loans are the only external financing source. Therefore, shadow bank has monopoly power. The entrusted loan contract, including loan size $L_S$ and return $R_S$, are set to maximize the profit of shadow banks.

Since shadow bank does not monitor the firm, the optimal level of entrepreneur effort equals to the non-monitoring effort level, that is,

$$p^* = p_0^* = \beta(Q - R_S).$$

(6.1)
For given effort level $p^*$, shadow bank chooses the loan size $L_S$ and payoff $R_S$ to maximize the expected profit:

$$\max_{L_S, R_S} p^* R_S - L_S. \quad (6.2)$$

subject to

$$p^*(Q - R_S) - \frac{p^*}{2\beta} + L_S + w - 1 \geq 0, \quad (6.3)$$

$$L_S + w \geq 1, \quad (6.4)$$

$$p^* R_S \geq r_s L_S. \quad (6.5)$$

Equation (6.3) and (6.4) are the participation constraints of the firm. Equation (6.3) indicates that the firm gets non-negative profit, and equation (6.4) ensures that the firm obtains enough financing to undertake the project. The last constraint (6.5) is the participation constraint of the shadow bank, which requires that the rate of shadow banks should be at least $r_s$.

The optimal loan size is solved as

$$L^*_S = 1 - w, \quad (6.6)$$

and the optimal return as

$$R^*_S = \frac{Q}{2}. \quad (6.7)$$

With monopoly power, the shadow bank sets a higher lending rate, so each firm only borrows the minimum amount $1 - w$ which is just enough to undertake the investment project. Shadow bank also charges very high return $R_S$ from borrowing firms, compared with that of the bank from equation (3.6). Consequently, the success probability of the borrowing firms from
restricted industry is lower than that of the borrowing firms with the same capitalization from unrestricted industry, because of the low return to entrepreneur and the lack of monitoring from shadow banks. Furthermore, the participation constraint of the shadow bank requires that

$$w \geq 1 - \frac{\beta Q^2}{4r_s} \equiv w^S. \quad (6.8)$$

Hence, only firms with an initial capital $w \geq w^S$ can get entrusted loans. Note that the threshold $w^S > w^N$ since $r_s > M$, which implies that the credit rationing with shadow banks is more severe than the one with formal banks.

**Lemma 4.** To firms from the restricted industries, shadow bank is the only outside financing source. Firms that have initial capital $w \in [w^S, 1]$ borrow from the shadow banks. The optimal loan size is $L^*_S = 1 - w$, and the required return to shadow bank is $R^*_S = \frac{Q}{2}$. Firms that have initial capital $w < w^S$ cannot get any financing.

The lending to restricted firms constitutes part of the entrusted loans demand. Another part comes from the unrestricted firms, which is derived in section 4.3. So the total demand of entrusted loans is

$$\tilde{T}D = (1 - q) \int_{w_1}^{w_2} L^M_S g(w)dw + q \int_{w^S}^{1} (1 - w)g(w)dw.$$ 

The first part is the total demand from un-restricted industries, the second part is the total demand from restricted industries. The supply of entrusted loan comes from the unrestricted firms, which is

$$\tilde{T}S = (1 - q) \int_{1-L^*_B}^{1} (L^S_B + w - 1)g(w)dw.$$ 

The equilibrium rate of return of shadow banking, denoted as $\tilde{r}^*_s$ to differentiate from the equilibrium rate of return $r^*_s$ when $q = 0$, is derived by equalizing the total demand to total supply. Shadow banking exists if and only if there exists a solution $r_s \in (M, +\infty)$ which
solves $\bar{T}D = \bar{T}S$. Based on Proposition 4, one has the following conclusion.

**Lemma 5.** *If there is a proportion* $q \in [0, 0.5)$ *of firms being in restricted industries, the equilibrium expected rate of return* $\bar{r}^*_s$ *is an increasing function of* $q$, *so* $\bar{r}^*_s > r^*_s$. Shadow banking sector grows when $q$ increases. When $q > \bar{q}$, the unrestricted firms are crowded out by the restricted firms, so only restricted firms get loans from shadow banks.*

The inclusion of restricted industries increases the demand of entrusted loans, further driving up the interest rates of entrusted loans. Therefore, the expected rate of return $r_s$ increases in $q$. Note that the lenders of entrusted loans are unrestricted firms which have initial capital $w \in (1 - L^S_B, 1]$. When $q$ increases, $r_s$ increases and $L^S_B$ also increase since it’s an increasing function of $r_s$. So the range of lenders, which is $(1 - L^S_B, 1]$ expands, meaning that the shadow banking sector expands with $q$. However, the demand for entrusted loans from the unrestricted firms shrinks when $r_s$ increases. When $r_s$ is above $\bar{r}_s$, no firm from unrestricted industries takes entrusted loans, so the demand side is only composed of the restricted firms. $\bar{q}$ solves the equation $\bar{r}^*_s(\bar{q}) = \bar{r}_s$, so when $q > \bar{q}$, one has $\bar{r}^*_s > \bar{r}_s$, and only firms from restricted industries borrow from shadow bank. Figure 7 plots the equilibrium lending with restricted industries.

Next, I study the impact of restricted industries to the total welfare of banks and firms. I calculate the total welfare in two cases: (1) there exists a proportion $1 - q$ of unrestricted firms and a proportion $q$ of restricted firms; (2) there exists only a proportion $1 - q$ of unrestricted firms. Case (2) can be interpreted as prohibiting lending of entrusted loans to restricted firms. By comparing these two cases, I can show how the restricted industry affects the shadow banking sector, in terms of the total social welfare. The total welfare in the first case is calculated as

$$TW_1 = (1 - q)TW^S_s(\bar{r}^*_s) + q \int_{\underline{w}^S}^{1} \frac{\beta Q^2}{8} g(w)dw.$$  \hfill (6.9)
Figure 7: The bank and shadow bank lending with restricted industries

(a) when $q = 0.03$

(b) when $q = 0.1$
The total welfare in the second case is calculated as

\[ TW_2 = (1 - q)TW^S(r^*_s), \]  

(6.10)

where \( TW^S(\cdot) \) is the total welfare function (5.4), \( \tilde{r}^*_s \) and \( r^*_s \) are the equilibrium rate of return of shadow banks in the first and second case, respectively.

Comparing the total welfare in the first and second case yields the following conclusion.

**Proposition 7.** Allowing shadow banks to lend to restricted industries reduces the total welfare. That is,

\[ TW_1 \leq TW_2. \]

Allowing shadow banks to lend to restricted industries extends the coverage of external credit but surprisingly reduces the total welfare. The inclusion of restricted industry affects the unrestricted industry indirectly. Allowing shadow banks to lend to restricted industry increases the total demand of entrusted loans, further increases the equilibrium rate of return \( \tilde{r}^*_s \) of shadow banks. Consequently, the demand of entrusted loans from borrowers of the unrestricted industry is reduced, the total amount of borrowing firms is also decreased since the range of \( (w_1, w_2) \) becomes smaller when \( \tilde{r}^*_s \) increases. So fewer borrowing firms from unrestricted industry would benefit from shadow banking since shadow banking increases the total welfare as proved in proposition 5. Although firms from restricted industry are financed by shadow banks and are able to conduct investment project, the welfare created by restricted firms is smaller than the welfare lost from unrestricted firms. On one hand, because the restricted firms could only obtain expensive entrusted loans due to the monopoly market power of shadow banks. On the other hand, restricted firms do not receive any monitoring, and thus the default risk is high. Therefore, the expected utility of restricted firms is very low. Overall, the total welfare is reduced by allowing shadow banks to lend to
the restricted industry. The welfare loss mainly comes from the significantly high default risk of restricted firms as a result of no monitoring from shadow banks.

7 Conclusion

A theoretical model is built to characterize the endogenous emergence of an important type of shadow banking in China: entrusted loans. The model shows that with entrepreneurial moral hazard and costly bank monitoring, entrusted loans arise under a loose monetary environment. A reduction of bank funding cost is associated with an expansion of entrusted loans market. Lenders of entrusted loans are firms which have high capitalization. Borrowers of entrusted loans are medium-capitalized firms. Lenders can obtain cheap bank loans because of lower entrepreneurial moral hazard and choose to over-borrow and form shadow banks with extra capital, then re-lend to medium-capitalized firms through shadow banking. Medium-capitalized firms simultaneously borrow bank loans and entrusted loans. This model can explain the observation that entrusted loans started to grow very rapidly after the economic stimulus plan in China from 2009-2010. Entrusted loans, as a form of shadow banking, is an unintended consequence of reduced bank funding cost by the credit expansion policy.

Despite being suboptimal to bank loans regarding monitoring, entrusted loans improve the total welfare of banks and firms. Excess bank monitoring is avoided when high-capitalized firms recycle bank loans to medium-capitalized firms. However, firms’ profits are reduced. Moreover, firms’ default risk is increased and real efficiency decreased. The reduction is correlated with the involvement of shadow banking activity, and especially significant for lending firms. This result is consistent with the empirical observation that abnormal stock returns of lender firms are negative after the entrusted loans are publicly announced.

In the extension, I also include the restricted firms which are prohibited from taking
bank loans. Since shadow banks are not regulated, restricted firms can borrow entrusted loans from shadow banks. But in this situation, the total welfare of banks and firms is damaged, because shadow bank acts as a monopolist and does not monitor borrowing firms.

I have assumed that all firms have positive NPV investment projects and lack of capital to undertake the projects. Entrusted loans are essentially bank loans that been passed through a lending chain. It is useful to include the case that entrusted loans are from retained earnings of firms that have run out of good investment projects. In this case, entrusted loans can be beneficial for lending firms, because of efficiently capital allocation. However, the impact on borrowing firms is unclear, and it is worth to have a study.

Last, this model links the growth of entrusted loans with the variation of bank funding cost caused by monetary policy. The result is in line with the observations around and shortly after the economic stimulus period in China. However, data shows that from 2012 to 2015, entrusted loans went through another fast growth, but newly issued bank loans barely change during this period. Therefore, it is necessary to study whether the growth of entrusted loans during 2012 to 2015 is driven by other factors.
References


## Appendices

### A List of Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
<td>Firm capitalization</td>
</tr>
<tr>
<td>$g(w)$</td>
<td>The probability density function of firm capitalization $w$</td>
</tr>
<tr>
<td>$w^N$</td>
<td>Threshold of firm capitalization under which firms are credit rationed</td>
</tr>
<tr>
<td>$(w_1, w_2)$</td>
<td>The range of firms that borrow from both banks and shadow banks</td>
</tr>
<tr>
<td>$p_0$</td>
<td>Non-monitoring effort of entrepreneur</td>
</tr>
<tr>
<td>$p$</td>
<td>Total effort of entrepreneur/ Probability of success of investment project</td>
</tr>
<tr>
<td>$\frac{1}{\beta}$</td>
<td>Marginal cost of entrepreneur effort</td>
</tr>
<tr>
<td>$\frac{1}{m}$</td>
<td>Marginal cost of bank monitoring</td>
</tr>
<tr>
<td>$M$</td>
<td>Funding cost of banks</td>
</tr>
<tr>
<td>$Q$</td>
<td>Return of the investment project if succeed</td>
</tr>
<tr>
<td>$r_s$</td>
<td>Expected return of shadow banking</td>
</tr>
<tr>
<td>$\bar{r}_s$</td>
<td>Upper bound of $r_s$</td>
</tr>
<tr>
<td>$L_B$</td>
<td>Firm’s demand of bank loan</td>
</tr>
<tr>
<td>$R_B$</td>
<td>Return of bank loan</td>
</tr>
<tr>
<td>$L_S$</td>
<td>Firm’s demand of shadow bank loan</td>
</tr>
<tr>
<td>$R_S$</td>
<td>Return of shadow bank loan</td>
</tr>
<tr>
<td>$L_B^S$</td>
<td>Equilibrium bank loan to the firms which become lenders of shadow bank</td>
</tr>
<tr>
<td>$R_B^S$</td>
<td>Equilibrium bank return from the firms which become lenders of shadow banking</td>
</tr>
<tr>
<td>$L_B^M$</td>
<td>Equilibrium bank loan to the firms which borrow from both bank and shadow bank</td>
</tr>
<tr>
<td>$R_B^M$</td>
<td>Equilibrium bank return from the firms which borrow from both bank and shadow bank</td>
</tr>
<tr>
<td>$L_S^M$</td>
<td>Equilibrium shadow banking loan to the firms which borrow from both bank and shadow bank</td>
</tr>
<tr>
<td>$R_S^M$</td>
<td>Equilibrium shadow banking return from the firms which borrow from both bank and shadow bank</td>
</tr>
<tr>
<td>$L_B^O$</td>
<td>Equilibrium bank loan to the firms which only borrow from the bank</td>
</tr>
<tr>
<td>$R_B^O$</td>
<td>Equilibrium bank return from the firms which only borrow from the bank</td>
</tr>
<tr>
<td>$TS$</td>
<td>Total supply of shadow bank</td>
</tr>
<tr>
<td>$TD$</td>
<td>Total demand of shadow bank</td>
</tr>
</tbody>
</table>
B Proof of Proposition 1, Lemma 1 and Proposition 2

The optimization problem of (3.3) is a special case of (4.1) when \( r_s = 1 \), so the proofs are put together. Results of Proposition 1 are derived by taking \( r_s = 1 \).

At first, it will be shown that the participation constraint (4.2) of banks is binding. To solve the optimization problem, we first substitute the optimal value of \( p^*_0 \) and \( p^* \) with equations (3.1) and (3.2), respectively. Then the Lagrangian can be written as

\[
L = \frac{1}{2}[\beta Q + (m - \beta)R_B](Q - \frac{m + \beta}{\beta}R_B) + r_s(L_B + w - 1)
\]

\[
+ \lambda \left[ R_B[\beta Q + (m - \beta)R_B] - \frac{mR_B^2}{2} - ML_B \right] + \mu(L_B + w - 1),
\]

where \( \lambda \) is the Lagrange multiplier associated with (4.2) and \( \mu \) is the Lagrange multiplier associated with (4.3). Differentiating the Lagrangian with respect to \( L_B \) and \( R_B \) gives

\[
\frac{\partial L}{\partial L_B} = r_s - \lambda M + \mu = 0 \quad \text{(B.1)}
\]

\[
\frac{\partial L}{\partial R_B} = -\beta Q - \frac{m^2 - \beta^2}{\beta}R_B + \lambda[\beta Q + (m - 2\beta)R_B] = 0 \quad \text{(B.2)}
\]

First, if we assume constraint (4.2) does not bind, then \( \lambda = 0 \), which implies that \( \mu = -r_s < 0 \) from the first order condition of (B.1). Therefore, constraint (4.2) binds, that is, the bank earns expected zero profit, and thus \( \lambda > 0 \).

In terms of constraint (4.3), we first assume that the constraint binds, then \( L_B = 1 - w \) and \( \mu > 0 \). Equation (4.2) and (4.3) give that \( L_B = \frac{R_B[2\beta Q + (m - 2\beta)R_B]}{2M} = 1 - w \). We can solve the optimal bank repayment as

\[
R^*_B = \frac{\beta Q - \sqrt{\beta^2Q^2 - 2M(2\beta - m)(1 - w)}}{2\beta - m} \equiv R^{OB}_B,
\]

48
which requires that
\[ w \geq 1 - \frac{\beta^2 Q^2}{2M(2\beta - m)} \equiv w^N, \]
so the determinant of the quadratic equation is non-negative. Next, to check whether \( \mu > 0 \) is satisfied. Combining equations (B.1) and (B.2) and substituting \( R^*_B \) into equation (B.2), yields
\[
\mu = \lambda M - r_s = M \left[ 1 + \frac{m^2 + \beta - m}{2\beta - m} \left( \frac{\beta Q}{\sqrt{\beta^2 Q^2 - 2M(2\beta - m)(1 - w)}} - 1 \right) \right] - r_s.
\]

\( \mu > 0 \) requires that \( r_s < M \left[ 1 + \frac{m^2 + \beta - m}{2\beta - m} \left( \frac{\beta Q}{\sqrt{\beta^2 Q^2 - 2M(2\beta - m)(1 - w)}} - 1 \right) \right] \equiv \tilde{r}_s. \) Note that \( \tilde{r}_s > M \geq 1, \) the benchmark applies since \( r_s = 1, \) and the equilibrium results are given as in Proposition 1.

On the other hand, if \( r_s \geq \tilde{r}_s, \) then \( \mu > 0 \) is not satisfied, which implies that constraint (4.3) does not bind. That is, the firm decides to over-borrow from the bank. If constraint (4.3) does not bind, then \( \mu = 0, \) equations (B.1) and (B.2) yield
\[
R^*_B = \frac{(\frac{r_s}{M} - 1)\beta Q}{(2\beta - m)\frac{r_s}{M} + \frac{m^2}{\beta} - \beta} \equiv R^S_B,
\]
which is smaller than \( Q \) when \( r_s > \tilde{r}_s. \) And optimal bank loan size is derived from (4.2) as
\[
L^*_B = \frac{(m - 2\beta)R^*_B + 2\beta Q(R^*_B)^2}{2M} = \frac{\beta^2 Q^2 \left( \frac{r_s}{M} - 1 \right) \left( (2\beta - m)\frac{r_s}{M} + \frac{2m^2}{\beta} - m \right)}{2M \left[ (2\beta - m)\frac{r_s}{M} + \frac{m^2}{\beta} - \beta \right]^2} \equiv L^S_B.
\]
Those are the results of Lemma 1 and Proposition 2.

49
C Proof of proposition 3

To solve the optimization problem (4.6), note that $L_S$ can be replaced by equation (4.8). In addition, substituting $p^* = p_0^* + mR_B$, the optimization problem can be simplified as

$$\max_{L_B, R_B, R_S} \frac{1}{2}[\beta(Q - R_S) + (m - \beta)R_B][Q - (1 + \frac{m}{\beta})R_B + (\frac{2}{r_s} - 1)R_S] + L_B + w - 1,$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} (C.1)

subject to

$$R_B[\beta(Q - R_S) + (\frac{m}{2} - \beta)R_B] \geq ML_B$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} (C.2)

$$L_B + \frac{R_S}{r_s}[\beta(Q - R_S) + (m - \beta)R_B] \geq 1 - w,$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} (C.3)

$$R_B + R_S \leq Q,$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} (C.4)

$$R_B \geq 0,$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} (C.5)

$$R_S \geq 0.$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} (C.6)

Let $\lambda, \mu, \gamma$ be the multiplier of constraint (C.2), (C.3) and (C.6), respectively, one gets the Lagrangian problem as

$$\mathcal{L} = \frac{1}{2}[\beta(Q - R_S) + (m - \beta)R_B][Q - (1 + \frac{m}{\beta})R_B + (\frac{2}{r_s} - 1)R_S] + L_B + w - 1$$

$$+ \lambda \left[ R_B[\beta(Q - R_S) + (\frac{m}{2} - \beta)R_B] - ML_B \right]$$

$$+ \mu \left[ L_B + \frac{R_S}{r_s}[\beta(Q - R_S) + (m - \beta)R_B] - 1 + w \right] + \gamma R_S.$$
Taking derivative with respect to $L_B$, $R_B$ and $R_S$ respectively, yields

\[
\frac{\partial L}{\partial L_B} = 1 - \lambda M + \mu = 0, \tag{C.7}
\]

\[
\frac{\partial L}{\partial R_B} = -\beta Q - (m - \beta)(1 + \frac{m}{\beta})R_B + (\beta + \frac{m - \beta}{r_s})R_S + \lambda \left[ \beta(Q - R_S) + (m - 2\beta)R_B \right] + \mu \frac{m - \beta}{r_s}R_S = 0 \tag{C.8}
\]

\[
\frac{\partial L}{\partial R_S} = (\frac{1}{r_s} - 1)\beta Q + (\beta + \frac{m - \beta}{r_s})R_B - \beta(\frac{2}{r_s} - 1)R_S - \lambda \beta R_B + \mu \frac{\beta(Q - 2R_S) + (m - \beta)R_B}{r_s} + \gamma = 0. \tag{C.9}
\]

First, constraint (C.2) is binding, otherwise, one should have $\lambda = 0$ and equation (C.7) implies that $\mu = -1 < 0$. Thus, one should have $L^*_B = \frac{R_B}{M}[\beta(Q - R_S) + (\frac{m}{2} - \beta)R_B]$ and $\lambda > 0$. Then, assuming constraint (C.6) is binding, that is, $R_S = 0$, one should have $\gamma > 0$.

Furthermore, assuming constraint (C.3) is also binding. The optimal value of $R_B$ can be derived from (C.3) as

\[
R^*_B = R^{OB}_B = \frac{\beta Q - \sqrt{\beta^2Q^2 - 2M(2\beta - m)(1 - w)}}{2\beta - m}
\]

The binding assumption of constraint (C.3) requires that $\mu > 0$, which can be proved by replacing $R^*_B$ into (C.8) and solving an equation system composed by (C.2) and (C.3):

\[
\mu = M \left[ 1 + \frac{\beta - m + \frac{m^2}{\beta}}{2\beta - m} \left( \frac{\beta Q}{\sqrt{\beta^2Q^2 - 2M(2\beta - m)(1 - w)}} - 1 \right) \right] - 1 > 0.
\]

In addition, $\gamma$ is derived from (C.9),

\[
\gamma = \beta(Q - R^{OB}_B) - \frac{[\beta Q - (\beta - m)(1 + \frac{m}{\beta})R^{OB}_B][\frac{M}{r_s}\beta Q - (\beta + \frac{M}{r_s}(\beta - m))R^{OB}_B]}{\beta Q - (2\beta - m)R^{OB}_B}. \tag{C.10}
\]
\( \gamma > 0 \) is equivalent to

\[
(Q - R_B^{OB})(Q - (2 - \frac{m}{\beta})R_B^{OB}) - M[Q + (\frac{m^2}{\beta^2} - 1)R_B^{OB}][\frac{Q}{r_s} + (\frac{m}{\beta} - 1 - \frac{1}{M})R_B^{OB}] > 0,
\]

which is a quadratic function of \( R_B^{OB} \). Note that if \( r_s \geq \frac{M}{3} \left(1 + \frac{m}{\beta} + 2\sqrt{1 - \frac{m}{\beta} + \frac{m^2}{\beta^2}} \right) \), the determinant of the quadratic equation is negative, so the inequality always holds. When \( r_s < \frac{M}{3} \left(1 + \frac{m}{\beta} + 2\sqrt{1 - \frac{m}{\beta} + \frac{m^2}{\beta^2}} \right) \), the inequality reduces to either

\[
R_B^{OB} > R_1 = \frac{Q}{2} \frac{(m^2 + \beta^2 - m\beta)M + \beta(2\beta - m)r_s + \beta(2m - 3\beta) + \sqrt{\Delta}}{(m^2 + \beta^2 - m\beta)((1 - \frac{m}{\beta})M + r_s) - (\beta - m)(2\beta - m)}
\]

or

\[
R_B^{OB} < R_2 = \frac{Q}{2} \frac{(m^2 + \beta^2 - m\beta)M + \beta(2\beta - m)r_s + \beta(2m - 3\beta) - \sqrt{\Delta}}{(m^2 + \beta^2 - m\beta)((1 - \frac{m}{\beta})M + r_s) - (\beta - m)(2\beta - m)}.
\]

\( R_1 \) and \( R_2 \) are two roots to the equation

\[
(Q - R)[Q - (2 - \frac{m}{\beta})R] - M[Q + (\frac{m^2}{\beta^2} - 1)R][\frac{Q}{r_s} + (\frac{m}{\beta} - 1 - \frac{1}{M})R] = 0
\]

and \( \Delta = (m^2 - m\beta + \beta^2)^2M^2 - 2\beta(2\beta - m)(m^2 - m\beta + \beta^2)M + \beta^4 + m\beta^2[2(2m - \beta) - 3mr_s]r_s. \)

Replacing \( R_B^{OB} \) and solving \( w \) yields

\[
w < 1 - \frac{2\beta QR_1 - (2\beta - m)R_1^2}{2M} \equiv w_1,
\]

or

\[
w > 1 - \frac{2\beta QR_2 - (2\beta - m)R_2^2}{2M} \equiv w_2.
\]

One needs \( w_1 \leq w_2 \), which reduces to \( R_1 + R_2 \leq \frac{2\beta Q}{2\beta - m} \), that is, \( r_s \leq \frac{\frac{m}{\beta}(\frac{m^2}{\beta^2} - \frac{m}{\beta} + 1)M + \frac{m}{\beta} - 2}{\frac{m^2}{\beta^2} + 2\frac{m}{\beta} - 2}. \)
If \( w_1 \leq w \leq w_2 \), \( \gamma > 0 \) is violated, so constraint (C.6) is not binding, which means that \( R_S^* > 0 \) and \( \gamma = 0 \). Then the optimal value of \( R_B^* \), \( R_S^* \) and value of \( \lambda \) are solved from equations (C.3), (C.8) and (C.9):

\[
\begin{align*}
\frac{\partial L}{\partial L_B} &= 1 - \lambda M + \mu = 0 \\
\frac{R_B}{M} \left[ \beta(Q - R_S) + \left( \frac{m}{\beta} - \beta \right) R_B \right] + \frac{R_S}{r_s} \left[ \beta(Q - R_S) + (m - 2\beta) R_B \right] &= 1 - w \\
\frac{\partial L}{\partial R_B} &= -\beta Q - (m - \beta)(1 + \frac{m}{\beta}) R_B + (\beta + \frac{m - \beta}{r_s}) R_S + \lambda \left[ \beta(Q - R_S) + (m - 2\beta) R_B \right] + \mu \frac{m - \beta}{r_s} R_S = 0 \\
\frac{\partial L}{\partial R_S} &= \left( \frac{1}{r_s} - 1 \right) \beta Q + (\beta + \frac{m - \beta}{r_s}) R_B - \beta \left( \frac{2}{r_s} - 1 \right) R_S - \lambda \beta R_B + \mu \frac{\beta(Q - 2R_S) + (m - \beta)R_B}{r_s} = 0.
\end{align*}
\]

which is equivalent to equation system 4.12.

\( \mu > 0 \) should always hold, since the constraint (C.3) is always binding. If not, one should have \( \mu = 0 \). Replacing \( \mu = 0 \), equation (C.9) reduces to \( (1 - r_s)(Q - R_B - R_S) = R_S + \left( \frac{r_s}{M} - \frac{m}{\beta} \right) R_B \), but the left hand side is non-positive since \( r_s \geq M \geq 1 \), and the right hand side is non-negative since \( m < \frac{\beta}{2M} \), which is contradiction. Therefore, constraint (C.3) is binding.

\section*{D Proof of Proposition 4}

The existence of shadow banking is equivalent to the existence of a solution \( r_s \in (M, \bar{r}_s) \) which satisfies the market clearing condition: \( TD = TS \). Since \( \bar{r}_s \) is a function of \( M \), one first needs that \( M < \bar{r}_s \). That is, \( M < \frac{1}{1 - \frac{m}{\beta^2}} \). If \( M \geq \frac{1}{1 - \frac{m}{\beta^2}} \), then \( \bar{r}_s \leq M \), there can not exist a shadow banking sector.

Secondly, both \( TD \) and \( TS \) are functions of the expected shadow banking rate \( r_s \), one can derive some properties of \( TS \) and \( TD \) with respect to \( r_s \). First, total supply \( TS \) is a monotonically increasing function of \( r_s \), since the bank loan demand \( L_B^* \) of lender is an
increasing function of \( r_s \), so

\[
\frac{\partial TS}{\partial r_s} = \int_{1-L_B^S}^{1} \frac{\partial L_B^S}{\partial r_s} g(w)dw + (L_B^S + w - 1)g(1 - L_B^S) \frac{\partial L_B^S}{\partial r_s} > 0.
\]

Intuitively, when the expected return from shadow bank lending is high, lenders prefer to borrow more from banks so that they can lend more through shadow banking, which drives up the total supply.

Total demand \( TD \) is a monotonically decreasing function of \( r_s \), since \( w_1 \) is an increasing function of \( r_s \), \( w_2 \) and \( L_M^S \) are both decreasing function of \( r_s \), so one gets

\[
\frac{\partial TD}{\partial r_s} = \int_{w_1}^{w_2} \frac{\partial L_M^S}{\partial r_s} g(w)dw + L_M^S g(w_2) \frac{\partial w_2}{\partial r_s} - L_M^S g(w_1) \frac{\partial w_1}{\partial r_s} < 0.
\]

Second, one can compute the values of \( TS \) and \( TD \) at two boundaries, respectively:

\[
\lim_{r_s \to M} TS = 0
\]

\[
\lim_{r_s \to \bar{r}_s} TS > 0
\]

\[
\lim_{r_s \to M} TD = \int_{w_1}^{1-L_B^S} L_M^S g(w)dw > 0
\]

\[
\lim_{r_s \to \bar{r}_s} TD = 0.
\]

Since the total supply is a monotonically increasing function, the total demand \( TD \) is a monotonically decreasing function, and values of both functions are positive on \( r_s \in (M, \bar{r}_s) \), by the fixed point theorem, there should exist a solution \( r_* \in (M, \bar{r}_s) \), which solves \( TS = TD \).

As have been proved that the total supple function is monotonically increasing with respect to \( r_s \) and total demand function is monotonically decreasing w.r.t \( r_s \). The two functions are illustrated in figure (8). When \( r_s = M \), total demand \( TD > 0 \), and when \( r_s = \bar{r}_s \), \( TD = 0 \). When \( r_s = M \), total supply \( TS = 0 \) and \( TS \) increases in \( r_s \).
Now, I prove that the intersection of the two curves is in between $M$ and $\bar{r}_s$. Suppose $r_s^* \geq \bar{r}_s$, then one should have $TD(r_s^*) < TD(\bar{r}_s) = 0$ and $TS(r_s^*) > TS(\bar{r}_s) > 0$, but one function is negative and another function is positive, there is no way to have intersection of the two curves. Thus, $r_s^* < \bar{r}_s$.

E Proof of Corollary 1

Suppose $r_s^*$ is not an increasing function of $M$, then for any $1 \leq M_2 < M_1 < \frac{1}{1 - \frac{m^2}{M^2}}$, one should have $r_s^*(M_1) \leq r_s^*(M_2)$. Because the total supple function $TS$ is an increasing function of $r_s$, and the total demand function $TD$ is a decreasing function of $r_s$, as derived from section D. One should have $TS(r_s^*(M_1)) < TS(r_s^*(M_2))$ and $TD(r_s^*(M_1)) > TD(r_s^*(M_2))$. But the definition of $r_s^*$ shows that $TD(r_s^*(M_1)) = TS(r_s^*(M_1)) < TS(r_s^*(M_2)) = TD(r_s^*(M_2))$, which is a contradiction. Therefore, $r_s^*$ is an increasing function of $M$. 

Figure 8: Plots of total demand and total supple functions
Proof of Lemma 2

As $w^N$ and $R_{OB}^B$ are both function of $M$, taking derivative of $TW^{OB}$ with respect to $M$ yields

$$\frac{\partial TW^{OB}}{\partial M} = \frac{\beta}{2} \int_{w^N}^{1} 2 \left( \frac{\beta^2 - m^2}{\beta^2} R_{OB}^B - Q \right) \frac{\partial R_{OB}^B}{\partial M} g(w) dw$$

$$- \frac{\beta}{2} \left[ Q^2 - 2Q \frac{\beta Q}{2\beta - m} + \frac{\beta^2 - m^2}{\beta^2} \left( \frac{\beta Q}{2\beta - m} \right)^2 \right] g(w^N) \frac{\partial w^N}{\partial M}$$

$$= \beta \int_{w^N}^{1} \left( \frac{\beta^2 - m^2}{\beta^2} R_{OB}^B - Q \right) \frac{\partial R_{OB}^B}{\partial M} g(w) dw - \frac{\beta^2 (\beta - 2m)}{2(2\beta - m)^2} Q g(w^N) \frac{\partial w^N}{\partial M}.$$ 

By equations (3.6) and (3.4), we can derive that $\frac{\partial R_{OB}^B}{\partial M} > 0$ and $\frac{\partial w^N}{\partial M} > 0$. Since $R_{OB}^B < Q$, one has $\frac{\beta^2 - m^2}{\beta^2} R_{OB}^B - Q < 0$. Moreover, $\frac{\beta^2 (\beta - 2m)}{2(2\beta - m)^2} > 0$, so overall, $\frac{\partial TW^{OB}}{\partial M} < 0$.

Proof of Proposition 5

For the first sector with banks and lending firms, one can calculate the total welfare as

$$TW_1^S = \int_{1-L_B^S}^{1} \left[ p(Q - R_B^S) - \frac{\beta^2}{2\beta} + r_s(L_B^S + w - 1) \right] g(w) dw$$

$$= \frac{\beta}{2} \int_{1-L_B^S}^{1} \left[ Q^2 - 2QR_B^S + \frac{\beta^2 - m^2}{\beta^2} (R_B^S)^2 \right] g(w) dw + \int_{1-L_B^S}^{1} r_s(L_B^S + w - 1) g(w) dw.$$ 

The total welfare in the case with no shadow banking sector is

$$TW_1^{OB} = \frac{\beta}{2} \int_{1-L_B^B}^{1} \left[ Q^2 - 2QR_{OB}^B + \frac{\beta^2 - m^2}{\beta^2} (R_{OB}^B)^2 \right] g(w) dw.$$
The difference between the two welfare is given as
\[
TW_1^S - TW_1^{OB} = \frac{\beta}{2} \int_{1-L_B^S}^{1} (R_B^S - R_B^{OB}) \left[ \frac{\beta^2 - m^2}{\beta^2} (R_B^{OB} + R_B^S) - 2Q \right] g(w)dw \\
+ \int_{1-L_B^S}^{1} r_s(L_B^S + w - 1)g(w)dw.
\]

The second part \( \int_{1-L_B^S}^{1} r_s(L_B^S + w - 1)g(w)dw \) is non-negative because \( L_B^S \geq 1 - w \). The first part \( \frac{\beta}{2} \int_{1-L_B^S}^{1} (R_B^S - R_B^{OB}) \left[ \frac{\beta^2 - m^2}{\beta^2} (R_B^{OB} + R_B^S) - 2Q \right] g(w)dw \) is positive because \( R_B^S < R_B^{OB} \) and \( 0 < \frac{\beta^2 - m^2}{\beta^2} < 1 \). Thus, \( TW_1^S > TW_1^{OB} \).

For the second sector with banks and borrowing firms, one can calculate the total welfare as
\[
TW_2^S = \int_{w_1}^{w_2} \left[ p(Q - R_B^M - R_S^M) - \frac{p^2}{2\beta} + L_B^M + L_S^M + w - 1 \right] g(w)dw \\
= \frac{\beta}{2} \int_{w_1}^{w_2} \left[ (Q - R_S^M)^2 - 2(Q - R_S^M)R_B^M + \frac{\beta^2 - m^2}{\beta^2} (R_B^M)^2 \right] g(w)dw.
\]

The total welfare in the case with no shadow banking sector is
\[
TW_2^{OB} = \frac{\beta}{2} \int_{w_1}^{w_2} \left[ Q^2 - 2QR_B^{OB} + \frac{\beta^2 - m^2}{\beta^2} (R_B^{OB})^2 \right] g(w)dw.
\]

I will show that \( TW_2^S > TW_2^{OB} \) by proving that \( TW_2^S \) is a decreasing function of \( r_s \) and
$TW^S_2(r_s) > TW^S_2(r_s = \bar{r}_s) = TW^{OB}_2$. Taking the derivative of $r_s$ on $TW^S_2$ yields

$$\frac{\partial TW^S_2}{\partial r_s} = \beta \int_{w_1}^{w_2} \left[ (R^M_B + R^M_S - Q) \frac{\partial R^M_S}{\partial r_s} + \left( \frac{\beta^2 - m^2}{\beta^2} R^M_B + R^M_S - Q \right) \frac{\partial R^M_B}{\partial r_s} \right] g(w)dw$$

$$+ \frac{\beta}{2} \left[ Q^2 - 2QR^{OB}_B + \frac{\beta^2 - m^2}{\beta^2} (R^{OB}_B)^2 \right] g(w_2) \frac{\partial w_2}{\partial r_s}$$

$$- \frac{\beta}{2} \left[ Q^2 - 2QR^{OB}_B + \frac{\beta^2 - m^2}{\beta^2} (R^{OB}_B)^2 \right] g(w_1) \frac{\partial w_1}{\partial r_s}.$$

Because $\frac{\partial R^M_S}{\partial r_s} > 0$, $\frac{\partial R^M_B}{\partial r_s} > 0$, $\frac{\partial w_2}{\partial r_s} < 0$ and $\frac{\partial w_1}{\partial r_s} > 0$, $R^M_B + R^M_S \leq Q$, one has $\frac{\partial TW^S_2}{\partial r_s} < 0$. Moreover, when $r_s \to \bar{r}_s$, $R^M_S \to 0$ and $R^M_B \to R^{OB}_B$, so $TW^S_2(r_s \to \bar{r}_s) = TW^{OB}_2$. Therefore, we proved that $TW^S_2(r_s) > TW^S_2(r_s = \bar{r}_s) = TW^{OB}_2$.

### H Proof of Corollary 2

One can split the total welfare with a shadow banking sector into two parts:

$$TW^S = TW^{OB} + \int_{w_1}^{w_2} \left[ U_F(L^M_B) - U_F(L^{OB}_B) \right] g(w)dw + \int_{1-L^S_B}^{1} \left[ U_F(L^*_B) - U_F(L^{OB}_B) \right] g(w)dw.$$

It has been proved that $\frac{\partial TW^{OB}_2}{\partial M} < 0$ in Lemma 2, if one can show that the second part is also a decreasing function of $M$, then the total welfare $TW^S$ is also a decreasing function of $M$. Define the second part as

$$\Delta TW : = \int_{w_1}^{w_2} \left[ U_F(L^M_B) - U_F(L^{OB}_B) \right] g(w)dw + \int_{1-L^S_B}^{1} \left[ U_F(L^*_B) - U_F(L^{OB}_B) \right] g(w)dw$$

$$= \int_{w_1}^{w_2} \left[ L^M_S - \frac{\beta}{2} [(R^M_S)^2 - 2QR^M_S - \left( \frac{m^2}{\beta^2} \frac{m^2 - m}{\beta^2} + 1 \right) (R^M_B)^2 - (R^{OB}_B)^2] \right] g(w)dw$$

$$+ \int_{1-L^S_B}^{1} \left[ (r_s - 1)(L^S_B + w - 1) - \frac{\beta}{2} \frac{m^2 - m}{\beta^2} + 1 \right] [(R^S_B)^2 - (R^{OB}_B)^2] g(w)dw.$$
Taking derivative on $\Delta TW$ w.r.t $M$ yields

$$\frac{\partial \Delta TW}{\partial M} = \int_{w_1}^{w_2} \left[ \left( \frac{\partial R^M_M}{\partial M} + \frac{\partial R^M_B}{\partial M} \right)(R^M_B + R^M_S - Q) - \frac{\partial R^O_B}{\partial M}(R^O_B - M) \right] g(w)dw$$

$$+ \int_{w_1}^{w_2} \left[ \frac{m^2}{\beta^2} \left( \frac{\partial R^O_B}{\partial M} R^O_B - \frac{\partial R^M_B}{\partial M} R^M_B \right) \right] g(w)dw$$

$$+ \int_{1 - l_B^S}^{1} \beta \left[ \frac{\partial R^S_B}{\partial M} \left( \frac{\beta^2 - m^2}{\beta^2} R^S_B - Q \right) - \frac{\partial R^O_B}{\partial M} \left( \frac{\beta^2 - m^2}{\beta^2} R^O_B - Q \right) \right] g(w)dw$$

$$+ \int_{1 - l_B^S}^{1} \left[ r_s \frac{\partial L^S_B}{\partial M} + \frac{\partial r_s}{\partial M} (L^S_B + w - 1) \right] g(w)dw. \quad (H.1)$$

Since $R^M_B + R^M_S > R^O_B$, one has $(\frac{\partial R^M_M}{\partial M} + \frac{\partial R^M_B}{\partial M})(R^M_B + R^M_S - Q) - \frac{\partial R^O_B}{\partial M}(R^O_B - M) < \frac{\partial R^O_B}{\partial M}(R^M_B + R^M_S + R^O_B - 2Q) < 0$, so the first part of equation (H.1) is negative. Because $R^O_B < R^M_B$, one has $\frac{\partial R^O_B}{\partial M} R^O_B - \frac{\partial R^M_B}{\partial M} R^M_B < \frac{\partial R^M_B}{\partial M}(R^O_B - R^M_B) < 0$, so the second part of equation (H.1) is also negative. Since $R^S_B < R^O_B$, one has $\frac{\partial R^S_B}{\partial M} \left( \frac{\beta^2 - m^2}{\beta^2} R^S_B - Q \right) - \frac{\partial R^O_B}{\partial M} \left( \frac{\beta^2 - m^2}{\beta^2} R^O_B - Q \right) < (\frac{\partial R^S_B}{\partial M} - \frac{\partial R^O_B}{\partial M}) \left( \frac{\beta^2 - m^2}{\beta^2} R^O_B - Q \right) < 0$, so the third part of equation (H.1) is also negative. Because $\frac{\partial L^S_B}{\partial M} < 0$, $\frac{\partial r_s}{\partial M} < 0$ and $L^S_B + w - 1 \geq 0$, so the last part of equation (H.1) is also negative. Therefore,

$$\frac{\partial TW^S}{\partial M} = \frac{\partial TW^O_B}{\partial M} + \frac{\partial \Delta TW}{\partial M} < 0$$

### I Proof of Proposition 6

I first prove that the lender firm’s profit is lower than that of benchmark, that is, $\pi^S_F < \pi^O_B$. The lender firm’s profit $\pi^S_F$ is a function of $r_s$, and $r_s \in (\tilde{r}_s, \bar{r}_s)$. One can get some properties
of $\pi^S_F$:

$$
\frac{\partial \pi^S_F}{\partial r_s} = \left[ \frac{m}{\beta} - 2 \right] Q + \left[ \frac{r_s}{M} \left( \frac{m}{\beta} - 2 \right) - 2 \left( \frac{m}{\beta} - 1 \right) \right] R_B^S \frac{\partial R_B^S}{\partial r_s} \\
= \left( \frac{m}{\beta} - 1 \right) \left( \frac{r_s}{M} + \frac{m^2}{\beta^2} - \frac{m}{\beta} \right) Q \frac{\partial R_B^S}{\partial r_s} \\
< 0,
$$

because $\frac{\partial R_B^S}{\partial r_s} > 0$. So $\pi^S_F$ is a decreasing function of $r_s$ on $r_s \in [\tilde{r}_s, \bar{r}_s]$, then one should have $\pi^S_F \leq \pi^S_F(r_s = \tilde{r}_s) = \pi_F^B$.

Next, I prove that the borrower firm’s profit is also lower than that of benchmark, that is, $\pi^M_F < \pi^O_B$. Note that the profit of benchmark $\pi^O_B = \pi^M_F(R_S = 0)$, so the proof of $\pi^M_F < \pi^O_B$ is equivalent to the proof that $\frac{\partial \pi^M_F}{\partial R_S} < 0$. As the profit of borrower firm is $\pi^M_F = p(Q - R_B - R_S)$, taking the derivative w.r.t $R_S$ gives $\frac{\partial \pi^M_F}{\partial R_S} = -p + (Q - R_B - R_S) \frac{\partial p}{\partial R_S}$.

In the equilibrium, one has $\frac{\partial p}{\partial R_S} < 0$, since shadow bank loan replaces bank loan, but bank provides monitoring which increases $p$ but shadow bank does not monitoring, so the increase of shadow bank loan reduces bank monitoring and results in lower success probability $p$.

Overall, $\frac{\partial \pi^M_F}{\partial R_S} < 0$, so $\pi^O_B = \pi^M_F(R_S = 0) > \pi^M_F$.

J Proof of Lemma 4

Replacing the optimal entrepreneur effort $p^*$ by equation (6.1), one can derive the Lagrangian problem as

$$
\mathcal{L} = \beta(Q - R_S)R_S - L_S + \lambda \left[ \frac{\beta(Q - R_S)^2}{2} + L_S + w - 1 \right] + \mu(L_S + w - 1),
$$
where $\lambda$ is the multiplier of constraint (6.3) and $\mu$ is the multiplier of constraint (6.4). Taking derivative with respect to $L_S$ and $R_S$ respectively, yields

\[
\frac{\partial L}{\partial L_S} = -1 + \lambda + \mu, \quad (J.1)
\]
\[
\frac{\partial L}{\partial R_S} = \beta(Q - 2R_S) - \lambda\beta(Q - R_S). \quad (J.2)
\]

Assuming constraint (6.4) binds, so $L^*_S = 1 - w$ and one needs $\mu > 0$. Further, assuming (6.3) also binds, then one should have $R^*_S = Q$, which contradicts with the requirement that $R_S < Q$. So (6.3) does not bind, then one gets $\lambda = 0$. Equation (J.2) gives the optimal shadow bank return $R^*_S = \frac{Q}{2}$. Substituting $\lambda = 0$ into equation (J.1), one gets $\mu = 1 > 0$, which confirms the assumption that constraint (6.4) indeed binds. At last, replacing the optimal values of $L^*_S$ and $R^*_S$ into shadow bank constraint (6.5), one gets $w \geq w^S$.

**K Proof of Lemma 5**

First proving that the equilibrium expected rate of return $\bar{r}^*_s$ with restricted industry is an increasing function of $q$. Taking $q_1 > q_2 > 0$, suppose $\bar{r}^*_s$ is not increasing function of $q$, then one should have either $\bar{r}^*_s(q_2) > \bar{r}^*_s(q_1)$, or $\bar{r}^*_s(q_2) = \bar{r}^*_s(q_1)$. In the first case, assuming $\bar{r}^*_s(q_2) > \bar{r}^*_s(q_1)$. Since $\bar{TD}$ is a decreasing function of $r_s$, one gets $\bar{TD}(\bar{r}^*_s(q_2)) < \bar{TD}(\bar{r}^*_s(q_1))$. One the other hand, $\bar{TS}$ is an increasing function of $r_s$, so one gets $\bar{TS}(\bar{r}^*_s(q_2)) > \bar{TS}(\bar{r}^*_s(q_1))$. By the definition of $\bar{r}^*_s$, one should have $\bar{TD}(\bar{r}^*_s(q_1)) = \bar{TS}(\bar{r}^*_s(q_1))$, and $\bar{TD}(\bar{r}^*_s(q_2)) = \bar{TS}(\bar{r}^*_s(q_2))$. But we have $\bar{TS}(\bar{r}^*_s(q_2)) > \bar{TS}(\bar{r}^*_s(q_1)) = \bar{TD}(\bar{r}^*_s(q_1)) > \bar{TD}(\bar{r}^*_s(q_2))$, which is a contradiction. In another case, assuming $\bar{r}^*_s(q_2) = \bar{r}^*_s(q_1)$, one should have $\bar{TS}(\bar{r}^*_s(q_2)) = \bar{TS}(\bar{r}^*_s(q_1))$. But $\bar{TS}$ is an decreasing function of $r^*_s$, so $\bar{TS}(r^*_s(q_1)) < \bar{TS}(r^*_s(q_2))$, which is again a contradiction. Therefore, $\bar{r}^*_s$ is an increasing function of $q$. So for any $q > 0$, $\bar{r}^*_s(q) > \bar{r}^*_s(q = 0) = r^*_s$.

When $q = 0$, there exists an upper bound $\bar{r}_s$ such that for any $r_s \geq \bar{r}_s$, the total
demand from unrestricted industries shrinks to zero, that is $TD(r_s) = 0$. But when $q > 0$, the total demand expands by the restricted industries. When $r_s \geq \bar{r}_s$, the demand from the unrestricted industries shrinks to zero, but the demand from the restricted industries is still positive. Hence, when $q$ is large enough, the equilibrium $\bar{r}_s$ could go above $\bar{r}_s$. In this case, only restricted industries require shadow bank loan, which is plotted in figure 9.

![Figure 9: Total demand and total supply function with restricted industries](image-url)
L Proof of Proposition 7

Note that when \( q = 0, \tilde{r}_s^* = r_s^* \), and \( TW_1 = TW_2 \). One can prove that \( TW_1 \) is a decreasing function of \( q \), if so, then \( TW_1 \leq TW_1(q = 0) = TW_2 \). Taking derivative of \( TW_1 \) with respect to \( q \) yields

\[
\frac{\partial TW_1(\tilde{r}_s^*)}{\partial q} = -TW_1(\tilde{r}_s^*) + (1 - q) \frac{\partial TW_1(\tilde{r}_s^*)}{\partial \tilde{r}_s^*} \frac{\partial \tilde{r}_s^*}{\partial q} + \int_{u^S}^{1} \frac{\beta Q^2}{8} g(w)dw - q \frac{\beta Q^2}{8} g(w^S) \frac{\partial w^S}{\partial \tilde{r}_s^*} \frac{\partial \tilde{r}_s^*}{\partial q}.
\]

One has \( \frac{\partial TW_1(\tilde{r}_s^*)}{\partial \tilde{r}_s^*} < 0, \frac{\partial \tilde{r}_s^*}{\partial q} > 0, \frac{\partial w^S}{\partial \tilde{r}_s^*} > 0 \), moreover,

\[
\int_{u^S}^{1} \frac{\beta Q^2}{8} g(w)dw < \int_{u^N}^{1} U^{OB} g(w)dw < TW_1(\tilde{r}_s^*).
\]

So \( \frac{\partial TW_1(\tilde{r}_s^*)}{\partial q} < 0 \), that is, the total welfare \( TW_1(\tilde{r}_s^*) \) is a decreasing function of \( q \). Thus, \( TW_1(\tilde{r}_s^*) < TW_1(q = 0) = TW_2(r_s^*) \).