Staff Working Paper/Document de travail du personnel—2024-2

Last updated: February 2, 2024

The Role of Long-Term Contracting in Business Lending

by Phoebe Tian

Financial Stability Department
Bank of Canada
XTian@bankofcanada.ca

DOI: https://doi.org/10.34989/swp-2024-2 | **ISSN** 1701-9397

Acknowledgements

I am indebted to Jean-François Houde and Ken Hendricks for their guidance and continuous support throughout the research process. I also benefit from comments and suggestions from Jason Allen, Jason Choi, Dean Corbae, Hanming Fang, Ohyun Kwon, Lorenzo Magnolfi, Nicola Pavanini, Mark Rempel, José-Víctor Ríos-Rull, Yasuo Terajima, Ikuo Takei, Joaquín Saldain, Alan Sorensen, Anson Zhou, and seminar participants at University of Wisconsin-Madison, Bank of Canada, and the International Industrial Organization Conference. The views expressed are those of the author and do not necessarily reflect the views of the Bank of Canada or any of its policy committees.

Abstract

This paper examines inefficiencies arising from a lack of long-term contracting in small business lending in China. I develop and estimate a dynamic model where firms repeatedly interact with the same lender. All loans are short-term. Collateral can be used to deter a strategic default by a firm, but the lender cannot recover the full value of the collateral in the case of a default. The endogenous contract terms—including interest rates, loan size and collateral—reflect a firm's probability of default in equilibrium. Learning drives the dynamics of contract terms because a firm's profitability type is unknown. Long-term contracts improve welfare mainly by mitigating the incentives for a firm to default.

Topics: Financial Institutions
JEL codes: D83, D86, G21, L14, L26

Résumé

Cette étude examine les inefficacités résultant d'un manque de contrats à long terme dans le secteur des prêts aux petites entreprises en Chine. J'élabore et estime un modèle dynamique dans lequel des entreprises font affaire de manière répétée avec le même prêteur. Tous les prêts consentis sont des prêts à court terme. Les garanties peuvent servir à prévenir des défauts de paiement stratégiques des entreprises, mais les prêteurs ne peuvent récupérer la valeur totale des garanties en cas de défaillance. Les modalités endogènes des contrats – taux d'intérêt, montants des prêts et garanties – tiennent compte de la probabilité de défaillance des entreprises en situation d'équilibre. L'apprentissage est au cœur de la dynamique des modalités des contrats, car le type de rentabilité des entreprises n'est pas connu au départ. Les contrats à long terme engendrent des gains de bien-être, essentiellement en réduisant les incitations à faire défaut pour les entreprises emprunteuses.

Sujets : Institutions financières
Codes JEL : D83, D86, G21, L14, L26

1 Introduction

Studies that focus on business loan markets in developed countries have shown the ubiquity of long-term contractual relationships between firms and banks.[1] Most commonly, such long-term relationships take the form of line-of-credit contracts, which typically have contract horizons of over one year. However, in loan markets in developing countries, long-term contracting can be too costly to implement. The development of these countries' financial systems and legal environments may not have reached the necessary level of maturity to effectively handle the complexities inherent in long-term contracting. How does the lack of long-term contracting affect the flow of credit from banks to firms in these markets? This is the research question of this paper.

I document the nature of short-term contracts using an eight-year sample of small- and medium-sized business (SMB) loans from a Chinese bank. All of the observed lending contracts are short term, with maturities of one year or less, and firms often repeatedly borrow on a yearly basis. I develop and estimate a model where the bank offers a short-term loan contract to the firm in each period. The loan contract is subject to limited enforceability, wherein the loan repayment cannot be enforced; instead, it is an ex-post decision of the firm. This creates agency friction and leads to inefficient loan defaults. In a counterfactual scenario with long-term contracting, the bank is able to credibly commit to all future contract terms over the entire contracting horizon, and not just for the current period. This allows for a pattern of a declining path of borrowing costs and an increased flow of capital over time. This intertemporal pattern strengthens the firm's incentive to repay their loan instead of defaulting, thereby lessening the agency friction and enhancing the firm's value. This paper provides the first quantification exercise for the value of long-term contracting in the context of a business.

Two empirical features are indicative of the sources of the inefficiencies in my setting. First, default is not uncommon: During the sample period of eight years, about one-third of firms default, after which they forego the collateral they had posted with the bank and drop out of the borrowing pool. From the bank's perspective, only about 50 percent of the collateral value can be recovered to compensate for the default loss, due to costly and lengthy legal procedures and market frictions in the resale of collateral. Therefore, even though collateral is used, default incurs considerable costs to both the firm and the bank.

Second, a large fraction of the firms are young (less than five years since establishment), which makes their profitability prospects rather opaque. New information on a firm's financial situ-

[1] See, for example, Petersen and Rajan (1994, 1995); Berlin and Mester (1998); Elsas and Krahnen (1998); Berlin and Mester (1999); Behr et al. (2011)

ation is often linked with changes in their contract terms (interest rate, loan size, collateral) in the next period. Specifically, negative news is associated with less-favourable contract terms in the next period (higher interest rates, smaller loans, and higher collateral).

Inspired by the two features of the data, I develop a dynamic model of lending with limited enforceability and uncertainty. Limited enforceability leads to the agency friction between the lender and the borrower even when collateral is used. This is because the value of the collateral that is recoverable by the bank following a loan default is less than its original value, due to the costly transfer process. Crucially, this cost is not born by the firm, leading to a lower private cost of default than the social cost. Thus, the private default incentive is always excessive from the social planner's view.[2] The presence of agency friction constrains capital flow and leads to socially wasteful defaults in equilibrium.

Uncertainty refers to a firm's productivity prospect (or type) being unknown a priori, which can be gradually learned through the arrival of new information in each period. I assume that *both* the firm and the bank do not know the firm's true type and they both observe the new information. In other words, information is incomplete but symmetric.[3] Uncertainty leads to a mismatch between the contract terms and the firm's true type. For example, a good-type firm can be given a higher interest rate and a smaller loan size than it could if its type were known. Such misprising of risk and misallocation of funds leads to lower total firm output than the complete information benchmark.

In the model with short-term (single-period) contracts, new contract terms are made at the beginning of each period, incorporating the latest information about a firm's financial situation. Since uncertainty about the firm type is gradually resolved with the arrival of new information, inefficiencies from previous mismatches can be mitigated with updates to the terms of the contract.

[2]Default can be socially optimal in my model. Think of a firm that has low productivity and a low future value. When this firm is struck by a large liquidity shock, the cost of repayment can be higher than the social cost of the default. In that case, even the social planner would find it optimal to default and exit the credit market.

[3]I make the symmetric information assumption for two main reasons. First, using positive correlation tests proposed by Chiappori and Salanie (2000), I find no evidence of asymmetric information. (See a detailed discussion in Section 3.1.) This is in contrast to findings in many previous studies, for example Crawford et al. (2018). My data contains the bank's *internal* credit assessments of borrowers, which encompass soft and hard information that the bank collected through extensive credit evaluations. When excluding this variable, the test shows a positive correlation. This suggests that the internal credit assessment is important for controlling what may otherwise be thought of as the firm's private information. Second, the literature points out that banks can accumulate considerable knowledge about the quality and prospect of a project through their experiences in funding numerous projects of similar characteristics (Manove et al., 2001). This is a unique information advantage that an individual entrepreneur does not have access to. While it is true that an individual entrepreneur could hide some private information from the bank, the bank has expertise in project appraisal and informed insights into the local and national economy, so the bank is not necessarily in a less-informed position.

Long-term (multi-period) contracts, on the other hand, have a unique feature that helps alleviate the agency friction arising from limited enforceability, as these contracts allow the lender to commit to a price schedule that declines with time. This means that the longer the lending relationship, the more a firm benefits from it, and this incentivises it to repay their loan at the end of each period. Thus, agency frictions are mitigated, which leads to lower default rates and enhanced capital flow.[4] However, if long-term contracts cannot be made contingent on new information about the borrower's type, as is often the case in reality, these contracts will not reduce the inefficiencies from this uncertainty similar to how short-term contracts do. This is because contract terms are unable to reflect the new information about the firm over time, thus, the lack of subsequent updates preserves the initial mismatch. Thus, whether long-term contracts without contingent terms are overall more efficient critically hinges on the relative degree of the uncertainty as well as the agency friction in the market.

I estimate a model of short-term contracts using the data on observed one-year loan contracts. In this model, the bank determines three contract terms at the beginning of each period, including the interest rate, loan size, and collateral.[5] A firm takes out a loan for production and chooses whether to default at the end of each period, after the realization of the productivity outcome and a liquidity shock. The productivity realization is a signal of the unknown firm type, and both the bank and the firm symmetrically update their beliefs by observing the signal. The updated beliefs are reflected in the next period's contract terms. If the firm defaults, its collateral is transferred to the bank and its liquidation value is less than firm's valuation of the collateral. The firm then exits with a salvage value.

Identification of the model relies on a novel aspect of the dataset that allows me to observe the bank's internal ratings of firms over time. This internal rating system summarizes all of what the bank knows about firms—soft and hard information—into a binary rating, updated annually.[6] I interpret the ratings as indicative of firms' productivity outcomes, with normal ratings implying good productivity outcomes and vice versa. Since productivity outcomes drive the learning process, I can use the observed ratings to pin down the belief-updating process.[7]

Inefficiencies from uncertainty come in the form of mispricing default risk and misallocating capital. From the estimation results, I find that mispricing induces a three percent increase

[4] This can only be done with the bank's commitment to these contracts (i.e., the contracts are legally binding) because the declining pricing schedule means the bank will earn negative expected profits later in the contract. In the absence of such a commitment, the bank would renege on the terms of the contract and abandon this relationship, and the equilibrium long-term contract would collapse to a sequence of short-term contracts.

[5] The observed bank is one of many homogeneous banks in a competitive lending market.

[6] Gao et al. (2019b) also utilize internal ratings to control for the soft information banks know about borrowers.

[7] A similar approach is seen in Pastorino (2019).

in the default of high-type firms, and capital misallocation results in four percent less capital borrowed by high-type firms. The effects of mispricing and misallocation are, however, almost the opposite for low-type firms and the aggregate effect is minimal. I estimate a relatively slow learning process. Over the seven-year period, learning reduces mispricing and misallocation by less than half.

Inefficiencies from agency friction result in excessive default and the associated social costs. Based on the model estimation, the socially optimal default rate is 70 percent lower than the observed default rate. In the absence of the agency friction arising from limited enforceability, total firm output would be 14 percent higher due to less-constrained capital flows. Overall, the estimation results show that the agency friction from limited enforceability is the major source of inefficiency, rather than uncertainty.

Counterfactual long-term contracts do not change the timing of financing (i.e., firms still get funded at the beginning of each year and repay at the end of the year). Instead, this contract type specifies a roadmap of interest rates (spreads), loan sizes, and collateral for all future periods within the horizons of these contracts. With contingent contract terms, long-term contracts further specify all future terms as functions of the history of the firm's financial situation. Otherwise all terms are functions of time only. I assume that the bank and the borrower can commit to the contract; that is, there are no contract abandonments at the start of each period. With these commitments, long-term contracts can achieve allocations that cannot be achieved with a sequence of short-term contracts: for example, the interest rate can be lower than the competitive market rate in later periods as long as the bank is compensated earlier by charging higher rates in the initial period and breaks even *ex ante*. This ability to intertemporally transfer profits is crucial for addressing the agency friction arising from limited enforceability. This is because from the firm's perspective, when prices are front-loaded, there is a higher continuation value and, thus, more incentive to repay. Here, private default incentives are lessened compared to the short-term case.

Quantitatively, with long-term contracts, the cumulative default rates would drop by around 17 percent.[8] And the reduced agency friction improves capital flows, which leads to an increase in total output by around two and half percent. These effects barely change when we restrict long-term contracts to being non-contingent on evolving borrower information, which disables the contract's ability to address the misallocation resulting from uncertainty. This result is foreshadowed by the estimation results that the overall effect from uncertainty in this environment is minimal. Therefore, counterfactual long-term contracts, even without contract terms con-

[8]In my counterfactual exercise I set the horizon to twenty years

tingent on new information, still improve overall welfare because the welfare gain in alleviating the agency friction arising from limited enforceability dominates the loss from misallocation.

Given the estimated welfare gain from implementing long-term contracts, why has the market not converted to such contracts?[9] There are three discernible factors contributing to the absence of long-term contracting in China's loan market. The first factor revolves around the onerous regulatory burden banks face in long-term contracting. The auditing and monitoring procedures mandated for contracts exceeding a one-year duration engender a formidable administrative burden. Such high fixed costs typically fail to rationalize the prospective revenues attainable through long-term contracts, especially those with SMEs. The second reason pertains to the lack of innovation within the banking sector. The sector is characterized by the dominance of the Big Four Banks, which have long benefited from substantial profit margins derived from the regulatory ceiling on deposit rates. This diminishes banks' incentives to innovate in terms of financial products and services. Consequently, even several decades after the inception of banking system reforms, the market remains predominantly saturated with one-year term loans, which represent the simplest and most basic form of loan contracts. Lastly, the viability of long-term contracts hinges critically on the commitment from banks. When interest rates drop below the spot rate, as prescribed in the optimal long-term contract, a bank has an incentive to abandon this contract and switch to a new one, provided that the associated legal costs of reneging are not prohibitive. Thus, in instances where the legal system is inefficient, long-term contracts are hard to implement. Notably, Demirgüç-Kunt and Maksimovic (1999) show that the origin and efficiency of a legal system affects firms' access to long-term financial agreements. Given that China is a developing country with relatively nascent legal and financial infrastructures, especially in its underdeveloped regions, this institutional environment presents a substantial impediment to ensuring the enforceability of long-term contracts for Chinese banks. Absent such commitments from banks, long-term contracts inevitably collapse into a series of short-term arrangements as observed in our setting.

Related Literature This paper is related to several strands of literature. First, there is a large literature on relationship lending.[10] This literature defines relationship lending as one where there is a close tie between the borrower and the lender when negotiating either explicit or implicit contracts (Petersen and Rajan, 1994). Many empirical studies use different measures

[9] Long-term contracting is rare not only for SMEs but also for large firms. According to Chang et al. (2014), who focus on loans made to large manufacturing firms in China, only five percent of these loans have maturities longer than one year. Recently, the SME lending market in China has witnessed the emergence of new players, such as specialized microfinance companies. However, it is noteworthy that these new entrants predominantly provide short-term loans (Dong et al., 2023).

[10] See Boot (2000) for a review of the literature

for lending relationships and show how they affect the lending terms and outcomes (Petersen and Rajan, 1994; Berger and Udell, 1995; Degryse and Van Cayseele, 2000; Chakraborty and Hu, 2006; Behr et al., 2011). A common hypothesis is that the accumulation of soft information within the relationship helps to enhance capital flow and lower borrowing costs (Berger and Udell, 1995). Here I directly observe a bank's soft information on firms, in the form of its internal ratings. My finding that these internal ratings are related to changes in contract terms is consistent with this view.

This paper deviates from this literature in several aspects. First, I do not assume the presence of long-term commitments in the empirical estimation. In the empirical setting of many studies in this literature, the observed close tie between the bank and the borrower is underpinned by some form of long-term contract. For example, the mean of the contract maturity in Chakraborty and Hu (2006) is almost two years. Thus it is natural to assume the presence of long-term commitments in these observed relationships. In my setting, however, the observed contract horizon is short term, so there is a good reason to suspect there may actually be long-term commitments underlying these seemingly close ties. Elsas and Krahnen (1998) perform an empirical test of relationship lending using loan-level data from major German banks. They find that lenders provide liquidity insurance in situations where borrowers' ratings may unexpectedly deteriorate, which is consistent with the idea of the long-term commitment. I find the opposite would occur in my empirical setting, so I assume the absence of long-term commitments in the estimated model.

Another notable deviation is that I assume symmetric information. This departs from the view, in this literature, that a borrower has private information that the inside bank does not know of and that the inside bank gathers more information about the borrower than outside banks do. The first part of the assumption—the information asymmetry between borrowers and lenders, as discussed in the previous section—is not well supported by the empirical evidence in my setting. The second part of the assumption—the asymmetric information between inside and outside banks—cannot be tested empirically using my data due to the limitation that outside offers to firms are not included in the data, posing constraints on the information assumptions I could make about outside banks. However, given the institutional setting that a bank has to report its internal ratings to the regulatory body in China and that other banks are able to assess them at a low cost, I am not uncomfortable with this assumption.

Second, there is an extensive literature, not all cited here, that studies optimal long-term financing contracts in which credit constraints emerge endogenously as a feature of the optimal contract design (Gertler, 1992; Thomas and Worrall, 1994; Albuquerque and Hopenhayn, 2004;

Quadrini, 2004; Clementi and Hopenhayn, 2006; DeMarzo and Sannikov, 2006; DeMarzo and Fishman, 2007; Biais et al., 2010; Boualam, 2018). My counterfactual long-term contracts are in line with the type of long-term contracts discussed in this literature. The closest papers are Albuquerque and Hopenhayn (2004) and Boualam (2018). Both papers focus on the friction caused by limited enforceability. Since collateral is not permitted as a contracting tool, the lender obtains zero value if the firm avoids repayment and walks away. The optimal long-term contract characterized in these authors' papers is *default free*.[11]

In my paper, collateral can be used as a contracting tool to mitigate the friction that arises from limited enforceability, as is commonly seen in practice. I also provide empirical evidence that the use of collateral is not a perfect solution because of the asymmetric valuation of collateral.[12] Another modelling difference is the addition of idiosyncratic liquidity shocks that affects the firm's default choice ex post. The equilibrium contract admits the probability of default no matter whether short or long term. In other words, default arises *endogenously* in this model.

Third, this paper is related to the discussion of limited commitment. In empirical industrial organization literature, the role of the limited long-term commitment has been studied in the context of life or health insurance (Cochrane, 1995; Hendel and Lizzeri, 2003; Atal, 2019). In Hendel and Lizzeri (2003), the supply side (i.e., insurance companies) is able to commit but the consumers cannot (i.e., they can switch), which leads to front-loading premiums in long-term contracts. This paper finds that when the supply side (i.e., banks) can commit, optimal long-term contracts are also front-loaded. The reason, however, is completely different from Hendel and Lizzeri (2003): Here front-loading is employed to alleviate the agency friction arising from the borrower's incentive to avoid repayment, while in Hendel and Lizzeri (2003) front-loading results from of a lack of consumer commitment. This paper thus complements discussions of the effect of a lack of long-term economic arrangements, as it examines a market characterized by agency friction. On the theory side, Kobayashi et al. (2022) study the implication of the lack of the lender's commitment in debt restructuring. Specifically, when the amount of debt exceeds the total repayable amount, the lender cannot credibly commit to a repayment plan where the repayment in each period is strictly smaller than the total output. Thus the borrower is discouraged from making first-best efforts. My paper focuses on loan origination as opposed to restructuring. The lack of lender commitment in this setting refers to the lender's inability

[11] In Albuquerque and Hopenhayn (2004), default is separate from liquidation. Through the contract, the lender can specify a liquidation policy, in which case the lender retains most of the liquidation value. In the optimal contract, default would not occur but liquidation endogenously arises. This context better describes the corporate bond market than the loan market.

[12] The cost of transferring collateral is one case of the asymmetric valuation of collateral between the borrower and lender.

to commit to future contract terms that deliver a lower payoff than its deviation payoff. This is similar to the type of lender commitment studied in Kovrijnykh (2013).[13]

This paper is also related to the stream of literature on collateralized debt. Chan and Kanatas (1985) find that in a competitive credit market with risk-neutral lenders, there must exist asymmetric valuation for collateral to be a meaningful contractual variable. In their paper, asymmetric valuation is modelled as the lender and the borrower having divergent beliefs on the distribution of the payoff in the case of default. Asymmetric valuation of collateral in my paper shows up as a constant governing how much the collateral is worth to the firm relative to the value recoverable by the bank. I further explore the implication of asymmetric valuation in the credit market, showing theoretically and quantitatively how this leads to agency friction and excessive defaults. On the empirical side, Ioannidou et al. (2022) use a structural approach to analyze the costs and benefits of collateral in the credit market. Their paper models a situation where the value of the collateral that a firm could pledge to secure a loan is taken as given and the bank chooses an interest rate deemed commensurate with its secured and (or) unsecured loan offer. This paper is different in two aspects. First, I consider the role of collateral in the absence of asymmetric information and I focus on the role of collateral as a disincentive for a firm to default. I do not model screening through collateral as in Ioannidou et al. (2022). Second, I take a different approach to modelling the determination of collateral. In my model, the interest rate, loan size, and collateral are jointly determined as the terms of a contract offer. In other words, banks compete on three terms simultaneously rather than on the interest rate alone. Third, I allow firms to be forward-looking in their default decisions. A firm faces a dynamic trade-off when making a default decision. Choosing to default rather than to repay the loan costs the firm its continuation value in the credit market. This approach allows me to explore the role of the long-term contractual relationship in determining the default probability and how collateral usage might be lowered as a result.

In addition, this paper is related to the structural estimation of learning models (Ackerberg, 2003; Crawford and Shum, 2005; Pastorino, 2019). Ackerberg (2003) and Crawford and Shum (2005) assume that the "signals" through which learning takes place are unobserved, while Pastorino (2019) treats the observed data on performance as direct evidence of these signals. My approach is similar to Pastorino (2019) in that I also utilize observed ratings as signals to pin down the belief-updating process. These papers are in the context of either the single-agent dynamic problem or the dynamic problem with no agency friction, whereas I contribute to this

[13]In Kovrijnykh (2013), the lender's commitment power lies in a continuum, with the degree of the lender's commitment modelled as the probability of contract enforcement in each period, which is an exogenous number from zero (no commitment) to one (full commitment).

literature by estimating the learning process in an environment with agency friction.

Finally, there is a vast literature on the relationship between financial friction and entrepreneurship (Cooley and Quadrini, 2001; Huynh and Petrunia, 2010; Buera et al., 2011; Arellano et al., 2012; Midrigan and Xu, 2014; Buera et al., 2015), and this paper contributes by providing micro-level evidence on how the friction in writing long-term contracts can affect the growth of small firms through inefficient default choices and the inefficient use of inputs.

The rest of the paper is organized as follows. Section 2 provides the institutional background and describes key features and patterns in the data. Section 3 describes the model. Section 4 discusses identification concerns, specifies the empirical model, and estimates the model's primitives. Section 5 conducts the counterfactual analysis. Section 6 concludes.

2 Data and Institutional Background

2.1 China's Banking Industry

The banking sector in China originated from a centralized system in 1949 when the People's Bank of China (PBC), as China's central bank, governed both commercial bank businesses (e.g., deposits, lending, and foreign exchange) and central bank functions. Along with economic opening policies being instituted by Deng Xiaoping in 1978, the banking system entered a period of reform. In 1983, the PBC began to focus on national macroeconomic policy, monetary stability, and economic development. At the same time, the Big Four commercial banks (i.e., the ICBC, ABC, BOC and, CCB) started to take over commercial bank businesses and each was specialized in a specific area. The Bank of Communications' (BOC) experience in reform and development paved the way for the development of shareholding commercial banks and exemplified banking reforms in China (Gao et al., 2019a). Chang et al. (2014) suggest that loan decisions any one of the Big Four banks make are predominantly based on commercial principles instead of government policies.

Between 1988 and 2005, twelve joint equity banks were established, mostly as state-owned enterprises (SOEs) or institutions transformed from local financial companies. Although joint equity banks are also national banks, unlike the current Big Five commercial banks, they usually focus on local business and operate on a much smaller scale. By the end of 2013, as reported by the China Banking Regulatory Commission (CBRC), the Big Five commercial banks dominated the market with a 43.3 percent share in the deposit market. The much smaller joint equity banks held a 17.8 percent deposit market share, with municipal commercial banks being among the remaining financial institutions by share of the deposit market size.

2.2 Deregulation of Credit Controls and Interest Rates

The deregulation of credit controls started from 1998. Until then, the PBC had controlled bank lending through binding credit quotas. This binding credit plan system was formally abolished in 1998 and replaced with an indicative non-binding credit target. In other words, commercial banks in China are no longer required to provide loans in compliance with state directives or policy targets. Instead, they are encouraged to allocate funds "on the basis of proper credit assessments" and to lend based on economic and commercial considerations. This change in policy has been hailed by Chinese monetary authorities as an important initial step in transforming the credit culture of Chinese banks (Xu et al., 2016).[14]

Lending rates in China became substantially more liberalized than deposit rates throughout the path of interest rate deregulation, starting with the interbank offered rate in the capital market becoming fully market priced in 1996. From 1998 onward, the PBC started to widen the floating band on banks' interest rates. In 2004, the deposit rate floor and the lending rate ceiling were eliminated for the major banks. The lending rate floor was gradually widened and eventually completely removed.[15] In practice, the lending rate floor was largely non-binding even before it was removed (Xu et al., 2016). On the other hand, the deposit ceiling was binding (He and Wang, 2012) and was not removed until October 23, 2015.

The day following October 23, 2015, marked the last change in the benchmark lending rate until this day. The benchmark lending rate refers to the official reference for the lending rate published by the PBC. It served as a non-binding "guidance" on lending rates in the market and had been an active policy instrument. Prior to October 2015, the PBC made changes in the benchmark lending rate on randomly chosen dates, typically seven or eight times a year.[16] After October 24, 2015, however, it seems that the benchmark lending rate ceased to function as an active policy instrument since no adjustment has been made so far; instead, the focus of the PBC is increasingly on short-term money-market rates, namely the seven-day interbank pledged repo rate (McMahon et al., 2018). This move is generally seen as a part of the interest

[14]There are still signs of quantitative controls on bank credit, as the central bank employs an array of quantitative instruments aimed at controlling credit growth, such as yearly aggregate target levels for new loans and the use of so-called window guidance, which can be described as a form of moral persuasion aimed at controlling the sectoral direction of lending (Okazaki, 2007).

[15]The lending rate floor was reduced to 0.9 times the benchmark official lending rate in October 2004, to 0.8 times the benchmark lending rate in June 2012, and to 0.7 times in July 2012, followed by a complete removal in July 2013.

[16]These changes had to be approved by the State Council (McMahon et al., 2018).

rate liberalization that had allowed the PBC to improve its policy framework (McMahon et al., 2018).

2.3 Government Policies for Small- and Medium-Sized Loans

The Chinese Government has long recognized SMEs' difficulties in obtaining external financing and included this concern in the national development agenda. Over the years, China has promoted SME financing mainly through interventions in financial markets, especially credit markets. Compared to many developed countries (such as the U.S., Japan, and Canada), China has not established specialized publicly supported agencies. Instead, policy priorities have centred on incentivizing existing financial intermediaries to increase SME lending. There are two broad types of policies on this front. The first type aimed to lower bank's financing costs to SMEs directly or indirectly. The central bank has set lower required reserve requirements for qualified banks that support SME and agricultural lending. Regulatory authorities also relaxed the risk weights applied for SME loans of less than RMB 5 million, applied simplified write-off procedures for SME loans and had greater regulatory tolerance for SMEs' nonperforming loans. In addition, local loan guarantee agencies were created under the SME Promotion Law 2003 in an attempt to provide assurance to banks in case of default.

The other type of policy explicitly sets yearly growth goals for SME lending. This practice of goal-setting started from 2008, when the China Banking Regulatory Commission (CBRC) implemented the "double no-lower-than" policy. This policy required that the growth rate of SME loans should not be lower than the growth rate of total loans and that the increase in SME loans should not be lower than in the previous year. The requirements on growth rates were not consistent over time: these were relaxed in 2011 and tightened again in 2015. From 2016 to 2021, the CBRC established even tighter goals for the Big Five national banks, requiring that the growth rate of SME loans be no lower than 30 percent in 2019 and 2020.

Banks face pressure to meet short-term goals and they are unsure about what the policy on SME lending will be in the following year or in the near future. Banks are thus reluctant to commit to future prices because of the short-term regulatory pressure and future policy uncertainty. For example, suppose the optimal long-term pricing structure is high-low, where banks commit to future price discounts in exchange for higher premiums in the beginning. Such a structure is unlikely to occur in the current environment: First, the pressure to meet imminent growth goals incentives banks to lower their loan prices in the first year and not to raise prices; second, the uncertainty about future policy makes it hard for banks to determine their future profit function and to find a price point to commit to. Therefore, banks are reluctant to establish long-term relationships by committing to future prices and quantities, given such pressure to achieve short-term growth targets and the uncertainty about future government policy on SME lending.

2.4 Data Description

This paper utilizes three sources of data from a Chinese bank over the period January 2010 to December 2017: (1) loan contract and loan outcome data; (2) (anonymized) corporate borrower data; and (3) internal rating data. I use loan and firm identifiers to merge different sources.[17] The loan contract data contains detailed contract information on corporate loans, including contract interest rates, loan sizes, loan maturities, repayment frequencies, dates of origination, types of collateral and the values of each type, whether they were guaranteed by a third-party guarantor, and branch and sub-branch identifiers. The loan outcome is an indicator of whether the loan is classified as a nonperforming asset (NPA) as of June 2018. NPAs are listed on the bank's balance sheet after a prolonged period of nonpayment and evidence of an extremely low repayment probability. These are typically viewed as loans that are in default.

Corporate borrower data includes the borrowing firm's industry code, location, size category, ownership type, date of incorporation, initial capital, and an internal credit assessment. The bank conducts an internal credit assessment when the firm borrows for the first time. Although the data do not have balance sheet information about firms, Chang et al. (2014) show that banks' internal credit assessments subsume firm-specific hard information and encapsulate the soft information about firm quality collected by the bank.

I also draw on data from the bank's internal loan rating system. For each loan, banks in China are required to report monthly ratings to the CBRC until the loan is off its balance sheet. These ratings are assigned according to a five-category loan classification system, where 1 is the highest rating for "normal" loans, 2 is for "special mention" loans, 3 is for "substandard" loans, 4 is for "doubtful" loans,and 5 is for loan "losses." This method of classification is mainly based on the borrower's *repayment ability*; that is, their *actual* ability to repay the principal and interest. Assessing this ability entails, for example, monitoring and analyzing changes in the borrower's revenue and profits, cash flow, financial position, and management efficiency. For a small fee, one bank can request records of a borrowing firm's past ratings history from the CBRC. Although banks report ratings every month until a loan is paid off, I only observe one snapshot taken in December of the corresponding year. Following Gao et al. (2019b), I interpret the five categories

[17] I can observe data on contract terms up to the middle of 2018, but the outcomes on loans that originated in 2018 were not recorded, so my final sample does not include those loans.

in a binary way, with category 1 indicating a good rating and anything between categories 2 and 5 indicating a poor rating. The dummy variable $PoorRating$ contains the binary loan rating information for each firm in every year of its borrowing activity.

All loans in the dataset are secured, either through collateral or a third-party guarantor, or both. Collateral can be either liquid assets such as securities, or fixed assets such as real estate. A third-party guarantor assumes joint liability for repayment if the borrower defaults. There are two stylized facts about collateral: (1) Higher-value collateral and having a guarantor are associated with lower interest rates, and the value of liquid assets has the largest negative effect on interest rates, which is consistent with the consensus that liquid assets are easiest to recover in case of default. (2) The usage of fixed-asset collateral and a third-party guarantor is sticky over time, but liquid-asset collateral is subject to substantial variation over time. In addition, the pledged value of liquid assets is sensitive to previous loan performance. This suggests that pledging fixed-assets as collateral and having a third-party guarantor are mostly constrained by the firm side, while pledging liquid-assets as collateral is subject to bank requirements.

Sample Restriction For the purpose of this paper I restrict the sample to only small- and medium-sized young firms that first borrow from the bank only after the beginning of the sample period, January 2, 2010. Specifically, the firm-level data has a categorical variable indicating firm size—small, medium, large—which is defined according to a national criteria.[18] I define a firm as being young if they are five years old or less at the first time of borrowing. In total, there are 6,358 such firms in my sample.

Feature 1: Loan contract terms are negotiated annually. This is based on three observations. First, 80.6 percent of loans have one-year maturities while the remaining loans have maturities under one year (typically six months). Second, 78.6 percent of firms only borrow once in a given year, and for those that negotiate multiple loans, the average within-year variation in interest rates is negligible relative to the between-year variation.[19] Finally, nearly all loan contracts feature a one-time loan repayment at maturity.

Based on this feature, I aggregate the loan-level data to the firm-year level. Specifically, for each firm, I take the average interest rate and value of the collateral, the sum of the loan size, and the outcome of the last loans within each year and obtain firm-level annual panel data on these

[18] See the administrative rule at `http://www.gov.cn/zwgk/2011-07/04/content_1898747.htm`. This classification rule defines the range of the number of employees and the annual revenues required to qualify as a small-sized company, which varies by industry. For example, in the retail industry, a company with from five to 20 employees and annual revenues of from 10 to 50 million RMB (1.4 to 7.2 million USD) is classified as a small sized.

[19] On average the coefficient of the variation of interest rates on multiple loans within the same year for the same firm is 0.07.

loans. In total, there are 17,518 observations at the firm-year level. Summary statistics of years one to seven are shown on Table 1.

Table 1: Summary Statistics

Year in Relationship t	Interest Rate r	Lending Spread	Loan Size k	Collateral Coverage z	PoorRating	Default
1	7.043	2.59	4.482	0.573	0.081	0.040
2	7.051	2.54	5.009	0.573	0.085	0.066
3	6.826	2.69	5.371	0.570	0.084	0.081
4	6.577	2.57	5.839	0.568	0.082	0.051
5	6.420	2.35	6.485	0.565	0.089	0.079
6	6.165	2.32	7.120	0.560	0.083	0.050
7	6.004	2.18	7.572	0.535	0.089	0.080
Average	6.863	2.56	5.195	0.567	0.085	0.058

This table presents summary statistics of the contract terms for the sample firms over the course of their lending relationships. Interest rates and lending spreads are shown in annual percentage rates and loan sizes are in units of RMB 1 million. The lending spread is the difference between the interest rate and and the Shanghai interbank offered rate (SHIBOR), which is a measurement of the cost of funds for Chinese banks. The formula for collateral coverage is the recoverable value of the collateral divided by the total loan amount. *PoorRating* is a binary variable that equals one if the bank's internal rating for the firm at year-end is classified as poor.

Feature 2: Repeated lending is common. To determine the longitudinal properties of the panel data, I define the duration of a relationship as the number of years during which the firm maintains at least one outstanding loan with the bank. For example, the duration is one year for firms that borrowed only once. Figure 1 shows the distribution of the relationship durations. Three quarters of firms borrow more than once, and the median duration is three years. For firms that repay their last year's loan, 70.09 percent return to the bank and take out a new loan. Firms that default exit the sample. Among the firms that never default, 90.6 percent borrow annually without a gap, 7.45 percent have only one gap year in their borrowing history, and only 1.95 percent of firms have a two-year gap or longer.[20] Table 1 shows the distribution of the relationship duration in my sample.

Feature 3: Default is a prominent source of risk despite common usage of collateral. There are three types of collateral: securities, fixed assets, and third-party guarantees, and these are ordered according to their perceived recoverability. In the sample, all of the loans are secured by at least one type of collateral: 35.5 percent of loans have only third-party guarantees, 28.4 percent only pledge fixed assets, 23.97 percent pledge both fixed assets and third-party guarantees,

[20]Due to data limitations, I cannot identify why a firm that had not previously defaulted would leave.

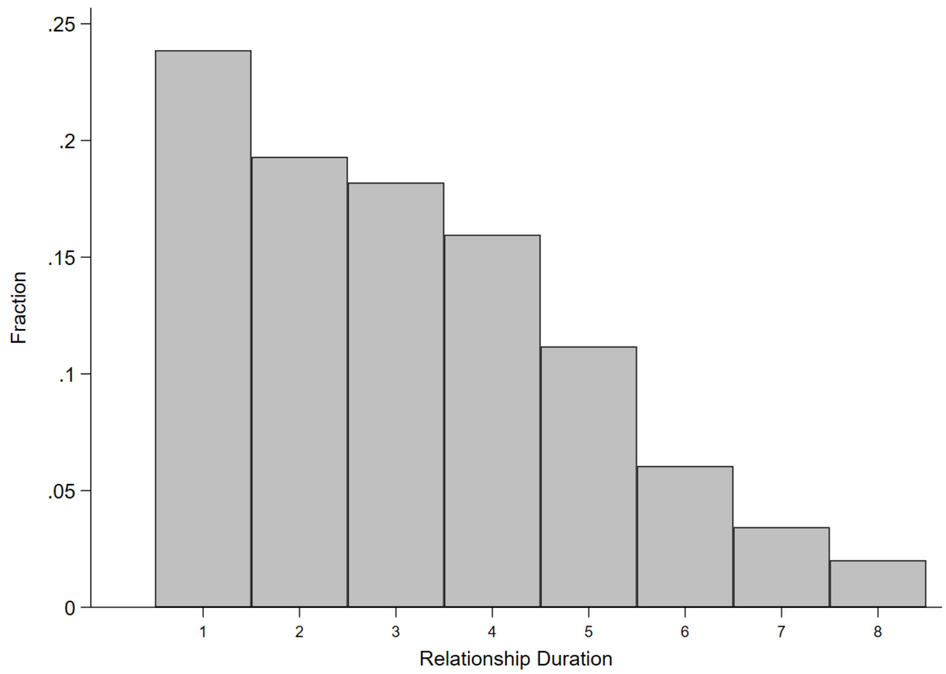


Figure 1: Distribution of Relationship Duration

This figure shows the distribution of the duration of the lending relationships among the sample firms. The duration of a relationship is defined as the number of years during which the firm maintains at least one outstanding loan with the bank. Firms whose first borrowing occurred before the sample period are not included.

and the remaining 12 percent use securities as collateral. In spite of the ubiquitous collateral usage, the bank still faces significant losses in the event of default. This is seen from the set of recovery rates, on each type of collateral, the bank uses to project its risks. The recovery rate on third-party guarantees is around 16 percent, which can vary with the credibility of the guarantor. The recovery rate on fixed assets is around 50 percent, varying with the specific type of asset. Securities have the highest recovery rate, at around 80 to 90 percent, due to their high liquidability. I calculate the recoverable value of collateral using the set of recovery rates provided by the bank and calculate the *collateral coverage* ratio as the ratio between the recoverable value of the collateral and the loan amount. The collateral coverage ratio is on average around 55 percent, which means that the bank's expected loss from default is around 55 cents on the dollar value of the loan. There are also considerable default risks: Around 33 percent of the firms in my sample end up in default over the period 2010 to 2017.

Feature 4: For firms with a history of good ratings, the contract terms improve over the course of the relationship. I define firms with good ratings histories as those firms without a single poor rating in their entire ratings history. In my sample, 76 percent of all firms have good ratings histories. Among those firms, I first regress the contract terms (lending spread, loan size, collateral coverage) on the observed firm characteristics (industry code, location, size category, ownership type, date of incorporation, initial capital, and an internal credit assessment), the cohort, the branch, and the monthly time fixed effects.[21] I then plot the residualized contract terms over year of the relationship, as shown in Figure 3. Overall, the residualized lending spread and collateral coverage decrease with the tenure, while the residualized loan size increases with the tenure. Alternatively, we can check the intertemporal trend of the contract terms within each firm. Utilizing the panel structure of the data, I run a regression of the contract terms on the year of the relationship, t, controlling for both the firm (α_i) and time fixed effects (ξ_t):

$$y_{it} = \beta_0 + \beta_1 t + \alpha_i + \xi_t + e_{it} \tag{1}$$

The results are reported in the columns (1) to (3) of Table 2 (where t stands for the year of the relationship). On average, as the relationship proceeds, the lending spread drops by 2.83 bps per year, the loan size increases by RMB 162,000 (≈ USD 24,000) per year, and the collateral coverage ratio drops by 0.35 percent per year. This suggests that the contract terms in general become more favourable for firms that are never rated poorly over the course of their borrowing relationships.

[21] The lending spread is the difference between the interest rate and the deposit rate.

Table 2: Regression of Contract Terms on Past Ratings and Year of Relationship

	Firms w. good rating history			Firms w. poor rating history		
	(1) Lending Spread	(2) Loan Size	(3) Collateral Coverage	(4) Lending Spread	(5) Loan Size	(6) Collateral Coverage
t	-0.0282	0.162	-0.00350	-0.0120	0.0387	0.00836
	(-3.72)	(8.81)	(-1.99)	(-1.53)	(0.84)	(1.07)
$PoorRating_{t-1}$				0.0399	-0.144	0.0162
				(3.15)	(-2.38)	(1.68)
$t \times PoorRating_{t-1}$				-0.00655	0.0267	-0.00797
				(-2.53)	(2.51)	(-2.17)
Observations	8328	8328	8328	2344	2344	2344

Columns (1)-(3) show the regression results from estimating Eq. (1) for firms with good ratings histories, with the lending spread, loan size, and collateral coverage, respectively, as the dependent variables. A firm with a good ratings history is one with no poor rating throughout the lending relationship. Columns (4)-(6) show the regression results from estimating Eq. (2) for firms that had at least one poor rating in the past. The units of the dependent variables are the same as on Table 1. t statistics are in parentheses. Other controls include firm and quarterly time fixed effects.

Feature 5: A poor rating is associated with an adverse change in the contract terms; such changes attenuate over the course of the relationship. Let period t denote the year of the relationship. Suppose in period t a firm is rated poorly for the first time. How will the contract terms respond in the next period? And how will such responses change with t? To see this, I regress the contract terms on a dummy variable, $PoorRating_{t-1}$, which equals one if the last period's rating is poor, the number of years in the relationship, t, and their interaction term, $t \times PoorRating_{t-1}$, while controlling for the firm fixed effects, (α_i), and time fixed effects, (ξ_t). The regression is run on firms with at least one poor rating in their history, and for each firm, i, I keep observation $t <= t_i^* + 1$, where t_i^* is the first time firm i is rated poorly.[22]

$$y_{it} = \beta_0 + \beta_1 t + \beta_2 PoorRating_{i,t-1} + \beta_3 t \times PoorRating_{i,t-1} + \alpha_i + \xi_t + e_{it} \qquad (2)$$

The results of the regressions are shown in columns (4) to (6) of Table 2. If a firm is rated poorly in the first period, then on average its contract terms the next period become tougher: the lending spread increases by around three bps, the loan size is reduced by around RMB 120,000, and the collateral coverage ratio increases by around one percent. However, if a firm's first poor rat-

[22]Firms are more likely to exit the sample after a poor rating due to default. Firms captured in this regression sample are those with poor ratings but that do not default. Thus, this is an underestimate of the responses, given this survivor bias.

ing happens only in period $t > 1$, then the adverse changes in its contract terms in the following period are not as large in magnitude. In other words, news that arrives later in the relationship worsens the contract terms to a lesser degree.

In sum, Feature 1 suggests that the observed loan contracts are short term, with annual negotiations. Feature 2 suggests that there can be meaningful dynamics in repeated lending contracts. Feature 3 suggests that the default risk is substantial. Features 4 and 5 together are consistent with a Bayesian learning process, where poor ratings lead to more-pessimistic beliefs about the firm's profitability and higher expected default risks, and this is reflected in tougher contract terms. A feature of Bayesian learning is that new information occurring later in the process updates the beliefs to a lesser extent, which can explain why the contractual responses to poor ratings attenuate over time.

3 Model

Before describing the full-fledged dynamic lending model, it is useful to show the intuition behind the main friction in the model. Appendix B provides a simple example in a static environment. In this example, collateral can be used to mitigate the agency friction from the limited enforceability of repayment. However, in the case of a default, the value that the lender can recover from the pledged collateral is less than its original value, which incurs a social cost of default. This cost is not internalized by the firm, thus, leading to a wedge between the private default and the socially optimal default. However, this wedge can be minimized when the forward-looking firm faces a high future value. To endogenize the formation of future value, I specify a dynamic lending model, as shown in the next section.

3.1 A Dynamic Lending Model: Overview

I develop a dynamic structural model of firms (borrowers) and banks (lenders) in an environment featuring the aforementioned agency friction and incomplete information. Information is incomplete in the sense that a firm's profitability prospect is unknown a priori. Beliefs are updated through a symmetric learning process. Symmetric learning means the banks and the firm are equally uninformed about the unobserved firm type and they share the same initial prior beliefs. In subsequent periods, they all update their beliefs based on publicly observed signals, so they share the same beliefs at any point in time.

My model of symmetric learning is consistent with the empirical results of this study. I conduct a test of the presence of asymmetric information in a contractual relationship, as proposed by Chiappori and Salanie (2000). Specifically, the presence of asymmetric information implies positive correlations in two places: The first place is between the decision to take up a loan and the decision to default. The second place is between the decision on the loan size and on whether to default, conditional on the observables. Details of the positive correlation test are provided in Appendix A. I do not find such positive correlations. In other words, firms that take out larger loans or borrow more frequently are not more likely to default, conditional on the observables. This is in contrast to the findings in Crawford et al. (2018) but is in line with the literature on entrepreneurship and project evaluation. For example, Manove et al. (2001) emphasize the role of banks' in project evaluation. Given their experience in funding large numbers of projects in specific sectors of the local economy, banks are well resourced to appraise the profitability prospect of those projects in an objective manner, while entrepreneurs often overestimate the profitability of their own projects (Cooper et al., 1988). In my setting, firms are young with little experience in their market and are more like the entrepreneurs discussed in this literature.[23] The bank in my setting invests considerable human resources in project evaluation, especially at the early stages, with experts who are well-trained in screening and monitoring projects across different sectors of the local economy. Given these experts' experience and the bank's investment in project-evaluation technology, it is reasonable to assume that the learning is symmetric in this environment. It is also tractable and allows us to quantify how informational frictions arising from incomplete information affect firm dynamics.

3.2 Model Setup

Time is discrete with an infinite horizon. There are two types of infinitely lived agents: firms and banks. Agents share the same discount factor $\beta \in (0,1)$.

Firms

Banks are risk neutral. Firms, or entrepreneurs, on the other hand, are risk averse. They operate their firms to maximize their expected utility from consumption within each period, where consumption is the total output net any payment. Firms are assumed to be hand-to-mouth without access to a storage technology, although this model can be easily generalized to allow capital accumulation. Flow utility $u : \mathbb{R} \to \mathbb{R}$ satisfies the following standard regularity conditions: $u' > 0$, $u'' < 0$, $\lim_{x\to 0} u'(x) = \infty$, and $\lim_{x\to\infty} u'(x) = 0$.[24]

[23]In Crawford et al. (2018), the borrowing firm's mean age is about 13 years, while I restrict my sample to firms less than five years old.

[24]Although many papers on dynamic debt contracting assume risk neutrality (with exceptions like Thomas and Worrall (1994) and Boualam (2018)), this paper adopts risk aversion to better represent young or distressed firms that lack the ability to diversify their firm-specific risks.

Each firm has access to a production technology but it is cashless initially and has to seek out external financing in order to start production. When funded with capital k, a firm of *quality type* $\theta \in \Theta$ and *characteristics* $s \in \mathscr{S}$ can generate output

$$y = af(k), \quad a \sim G(\cdot|\theta, s),$$

where the function f is differentiable, strictly increasing and strictly concave, and satisfies $f'(k) > 0$, $f''(k) < 0$, $f(0) = 0$, $\lim_{k\to 0} f_k(k) = +\infty$, and $\lim_{k\to\infty} f_k(k) = 0$.[25] Productivity a of a firm with type θ and characteristics s in each period t is a random draw from the probability distribution $G(\cdot|\theta, s)$, with the associated probability density function (or probability mass function) denoted as $g(\cdot|\theta, s)$.

The firm's quality type and characteristics are both related to the distribution of its productivity.[26] Specifically, conditional on the same s, a higher value of θ means higher productivity in the first-order stochastic dominance sense. Formally, if $\theta' > \theta$, then $G(a|\theta', s) \geq G(a|\theta, s), \forall a$. In other words, a higher θ indicates higher quality.

Firms have a choice of whether to repay the bank loan in each period. Following production, an i.i.d utility shock ϵ associated with repaying the bank loan is realized. The firm's repayment decision is based on its observation of the repayment shock. Repayment shocks are assumed to be independent of the firm's quality type. The mean of ϵ is zero, with the variance denoted by σ. Repayment shocks can be interpreted, for example, as temporary liquidity shocks or the unobserved demand shocks the firm faces.

Note that this setup nests the standard optimal long-term debt contract model with endogenous borrowing constraints (Albuquerque and Hopenhayn, 2004; Boualam, 2018) by setting σ to zero. In this case there is no randomness in the default behaviour and it is possible to completely prevent default by designing the contract subject to a default-free constraint. In general, a positive σ introduces some randomness in the firm's default behaviour, which allows positive default rates in equilibrium.

[25] The production function abstracts from labour and can be viewed as a profit function that already accounts for the optional choice of labour input and the associated wages.

[26] In the empirical specification, I let quality type θ and characteristics s play different roles in productivity distribution G. Conditional on the same θ, I assume s only moves G horizontally without changing its shape. Formally, that means for $s' \neq s$ and there exists a scalar b such that $G(a;\theta, s') = G(a + b;\theta, s'), \forall a$. On the other hand, θ can change the shape of the distribution G conditional on the same s. A high θ means a more right-skewed distribution G and, thus, a higher expectation of productivity.

Loan Supply and Information Structure

There are two banks in a competitive bank loan market. Each bank acts as an intermediary channelling funds from depositors to firms at funding cost c, and the funding cost is the same for every bank. Abstracting from the bank heterogeneity allows me to focus on how agency friction and learning shape the dynamics of contract terms. Recall that funding (deposits) is highly regulated in China; thus, I assume that funding cost $c \in \mathbb{C}$ follows a stochastic process with transition probability $\Gamma_c : \mathbb{C} \times \mathbb{C} \to [0,1]$.

Information is symmetric among the agents in this model. In the beginning, these agents do not know the firm's quality type θ but they can observe its characteristics s. The agents' initial prior belief is given by the distribution of θ conditional on the firm characteristics; that is, the agents have rational expectations. In subsequent periods, the realizations of the firm's output, or productivity, are commonly observed, based on which the agents update their beliefs via Bayes rule.

Formally, denote $p_t(\theta)$ as the belief that a given firm is of type θ at the beginning of period t (before production takes place in this period). Suppose that $\Lambda(\theta|s)$ is the true distribution of the quality types conditional on firms' characteristics. The initial prior $p_1 : \Theta \to [0,1]$ is

$$p_1(\theta) = \Lambda(\theta|s).$$

Given period t, belief p_t, and this period's productivity realization, a_t, we can find the updated beliefs in $t+1$, $p_{t+1} : \Theta \to [0,1]$ by using Bayes rule:

$$p_{t+1}(\theta) = \frac{g(a_t|\theta, s)p_t(\theta)}{\int_{\theta'} g(a_t|\theta', s)p_t(\theta')d\theta'} \tag{3}$$

Timing

Figure 2 describes the timeline of the model, which consists of two stages. Period zero is the origination stage where banks compete by offering short-term contracts and the firm chooses one bank. The firm-bank pair proceeds to the dynamic contracting stage (period $t \geq 1$). In each period, the order of events is as follows: First, capital is advanced and production takes place. Then productivity is observed and used to update the beliefs. The repayment shock is then realized, after which the firm makes its repayment decision. The contract continues only when the firm repays; otherwise, the contract ends.

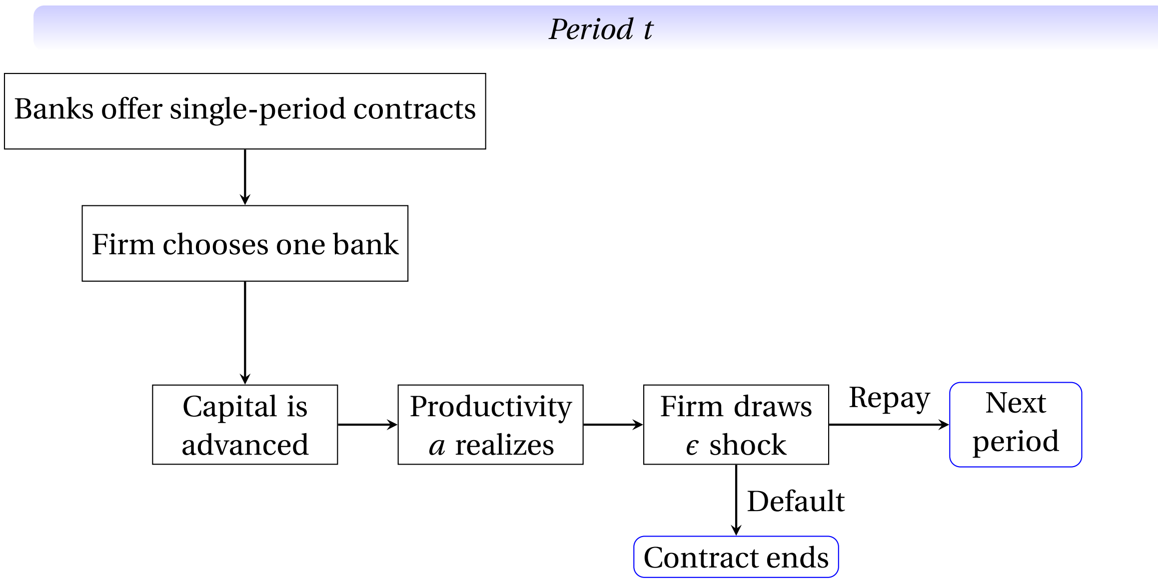


Figure 2: Timing

Lending Contracts

Contract offers are built from standard single-period lending contracts. A single-period lending contract in period t consists of the interest rate (plus one) $r_t \in \mathbb{R}^+$, loan size $k_t \in \mathbb{R}^+$, and collateral coverage $z_t \in [0,1]$. The collateral value that is *recoverable* by the bank in the event of default is given by zk. However, there are two reasons why the collateral is worth more to the firm than to the bank. First, consider specialized production equipment as collateral and the firm that specializes in production has a better use for this equipment than the non-specialist bank. Second, the process of transferring this collateral from the firm to the bank is lengthy and costly, and the process of reselling it to recover its value faces market frictions that lead to significant discounts. Therefore, I assume the value of the collateral is $(1+\eta)zk$ to the firm, which is forfeited in the event of default.[27] Note that the asymmetry in the collateral valuation is notationally different from the previous example in Section B. This is because the data contains the bank's assessment of the recoverable value of collateral zk but not the firm's valuation. By putting parameter η on the firm's valuation, I can ensure that all contract terms $r,\ k,\ z$ are observables.[28]

[27]Here we think of collateral as coming from the firm's illiquid assets, which is different from working capital. I do not consider the constraints on the total pledgeable assets that a firm might face or how past bank loans might lead to the accumulation of total pledgeable assets. These are interesting theory extensions, but extra data on the firm's assets is needed to identify the related parameters.

[28]In a previous version of this paper I also include a cost the firm bears for pledging this collateral *regardless of* whether it defaults. This is to capture the implicit costs the borrower incurs in being forced to relinquish discretionary use of the asset as well as upfront costs of the legal documentation and appraisal. The estimated magnitude of such costs is economically insignificant.

What then happens to a defaulting firm depends on the larger economic institutional systems in place. Here I summarize the various costs related with default in a reduced-form way by assuming that a firm would lose a fraction $1-\delta$ of its firm value. Put differently, a defaulted firm has salvage value that equals δ times the future firm value if the firm were to not default. An extreme case of $\delta = 1$ is where default does not affect a firm's franchise value but allows it to seamlessly continue accessing the credit market. The case of $\delta = 0$ is the opposite situation where defaulting firms are excluded from credit markets forever and completely cease to produce.

In general, an environment with a higher δ gives a firm a higher incentive to default, since the salvage value is higher. Thus, the parameter δ determines the level of the agency friction in place.[29] Another important parameter determines that the agency friction is η. As the illustrative example shows, the differential valuation of the collateral determines the wedge between the firm's default-decision region and the social planner's solution.

3.3 Equilibrium Short-Term Contracts

I consider an environment where long-term contracting is inaccessible. Alternatively, we can think of this case as lacking long-term commitments; that is, a bank would be free to renege on a promise it had made in a previous contract that is related to future contract terms, and the firm can unilaterally leave the relationship. In other words, if a bank were to make credible promises related to future contract terms in a present contract, these terms cannot be reinforced (i.e., promises on future terms cannot be *enforced* in this institutional environment)[30] and the firm cannot make credible promises related to future borrowing. It can only choose contract terms for the current period.

An equilibrium short-term contract is a single-period contract that maximizes the firm's value while also being subject to zero expected bank profits. This is a direct result from bank competition on three dimensions: interest rates, loan size, and collateral coverage. To see this, first consider the case where a contract brings positive expected profits to a bank. Another bank can undercut this by lowering its prices by a small amount and winning contracts from every firm (which is the standard Bertrand competition logic), so this cannot be the equilibrium contract. The contract cannot bring negative expected bank profits either, by the bank rationality requirement. The equilibrium contract, therefore, must bring zero expected bank profits. There can be many contracts that satisfy a zero-profit constraint and it turns out that the contract that

[29] Parameters δ and σ can be different for firms with different characteristics s, as in the estimated specification.

[30] It is worth noting that the lack of enforcement is on *future* transactions as well as default behaviour. But conditional on default, the transfer of collateral can be enforced.

maximizes the firm's value is the only one that can be the equilibrium contract. If a bank offers a zero-profit contract that does not bring the highest possible firm value, then another bank could optimize the contract structure and achieve higher firm value while still earning infinitesimal positive profits per contract, which attracts all firms. Thus, the equilibrium contract must bring the highest possible firm value while being subject to zero expected bank profits.

Let $W(p_t, c_t)$ and $\Pi(p_t, c_t)$ denote the expected firm value and bank profits when the belief of the firm's unobserved type is p_t and the funding cost is c_t in period t. To express the bank profits and firm value, we can first find the firm's value in choosing to repay or default as well as the associated repayment probability (or, equivalently, the firm's default probability) on a given short-term contract (r_t, k_t, z_t).

Given period t's realization of productivity a_t and repayment shock ϵ_t, the value of choosing to repay for a firm with characteristics s is

$$u\left(y_t - r_t k_t\right) - \epsilon_t + \beta E_t W(p_{t+1}, c_{t+1})$$

where $E_t W(p_{t+1}, c_{t+1})$ is short for $E_{c_{t+1}}\left[W(p_{t+1}, c_{t+1}) | p_t, c_t, a_t\right]$ with the next period's belief p_{t+1} completely determined by p_t and a_t through Equation (3).

If the firm chooses to default, it looses the value of collateral $(1+\eta) z_t k_t$, so the firm's value of choosing to default is

$$u\left(y_t - (1+\eta) z_t k_t\right) + \beta\delta E_t W(p_{t+1}, c_{t+1})$$

where $\delta E_t W(p_{t+1}, c_{t+1})$ is the salvage value for the defaulted firm.

Thus, the probability that the firm repays the loan viewed at the beginning of period t can be expressed as

$$\Phi_t(a_t) \equiv \Pr\left(\epsilon_t \leq u\left(y_t - r_t k_t\right) - u\left(y_t - (1+\eta) z_t k_t\right) + \beta(1-\delta) E_t W(p_{t+1}, c_{t+1}) \middle| a_t\right) \tag{4}$$

Note that the repayment probability is contingent on the realized a_t, since the default decision is made after the productivity realization. Specifically it determines the next period's belief p_{t+1}.

Using this notation for the repayment probability, we can express the bank's expected profits at the beginning of period t as

$$\Pi(p_t, c_t) = -c_t k_t + E_{a_t}\left[\Phi_t(a_t) r_t k_t + (1 - \Phi_t(a_t))\, z_t k_t \middle| p_t\right] \tag{5}$$

Finally, we can define the equilibrium firm value as $W(\cdot)$, which satisfies

$$W(p_t, c_t) = \max_{r,k,z} E_{a_t,\epsilon_t}\Big[\max\big\{u\big(y_t - r_t k_t\big) - \epsilon_t + \beta E_t W(p_{t+1}, c_{t+1}),$$

$$u\big(y_t - (1+\eta) z_t k_t\big) + \beta\delta E_t W(p_{t+1}, c_{t+1})\big\}\Big| p_t\Big] \tag{6}$$

$$\text{s.t. } 0 = \Pi(p_t, c_t) \tag{7}$$

and the associated policy function $r(p_t, c_t), k(p_t, c_t), z(p_t, c_t)$ constitutes the equilibrium contract terms. Note that the expectation is over a_t and ϵ_t, both of which are realized after the loan origination. Productivity realization a_t determines the next period's belief p_{t+1} and ϵ_t determines the repayment/default choice (both are inside the maximization operator).

4 Empirical Analysis

We now move to the empirical aspects of the model, starting with a more detailed discussion of the assumptions, identification, and specification. Since I only observe firms borrowing from *the* sample bank, from now on, I denote by $t = 1$ the first year that a firm borrows from the bank. Thus, t denotes the year of the firm-bank lending relationship.

4.1 Preliminaries

I assume that there are two quality types: high type (θ^H) and low type (θ^L).[31] This means that the beliefs can be summarized only by a scalar, p, which is the perceived probability that a firm is, say, a high type. This greatly simplifies the model while still maintaining the essence of learning. The true fraction of the high types conditional on s is given by λ_s, which by rational expectations also constitutes the initial prior belief for firms with characteristic s, by assuming rational expectations.

The distribution of productivity conditional on s and θ is assumed to be a two-point distribution: $a \in \{\overline{a}_s, \underline{a}_s\}$ with $\overline{a}_s$ strictly larger than $\underline{a}_s$. Note that the support of a does not change with the unobserved type θ. Otherwise, agents would immediately infer the firm's type after one period. Conditional on the same s, different types of firms, in terms of quality type θ, have different probabilities of realizing $\overline{a}_s$ (and thus $\underline{a}_s$). Specifically, the high type θ^H has a strictly lower probability of realizating low productivity: $g(\underline{a}_s|\theta^H, s) < g(\underline{a}_s|\theta^L, s)$.[32]

[31] The binary assumptions on the unobserved types are also seen in Pastorino (2019), Xin (2020).

[32] This is implied by the first-order stochastic dominance assumption (mentioned in the model section) in the context of a binary type and two-point distribution.

Whether the productivity realization is low ($\underline{a}_s$) or high ($\overline{a}_s$) constitutes the signal that drives the learning process. I use the bank's internal ratings as *direct* measurements of such signals, which are observed by the bank, firm, and econometrician. As mentioned in Section 2.4, the internal ratings, by definition, directly reflect the bank's dynamic assessments about the firm's repayment ability. Formally, I assume that an observed poor rating for firm i in period t reveals a realization of low productivity: $PoorRating_{it} = \mathbb{1}\{a_{it} = \underline{a}\}$ where $\mathbb{1}\{\cdot\}$ is the indicator function.

Note that internal ratings are different from the actual loan outcomes: close monitoring results in banks often downgrading firms' internal ratings before the actual delinquency happens (Gao et al., 2019b). The data used in this paper contain one snapshot of the monthly internal ratings on every loan. Specifically, this snapshot is taken in December, which means that for loans that are due, say, next March, this rating is recorded three months ahead of the due date. I use observations with due dates that are at least three months ahead of December. This ensures that the internal rating does not merely reflect the loan outcome; it contains valuable information that shapes the bank's belief about the firm's quality.

I allow firm characteristics s, which are known to both the banks and the firms, correlated with a set of firm-level variables X that is observed by the econometrician. Specifically, I assume that these firms' characteristics fall into S discrete groups: $s = 1, 2, ..., S$ (where S is known) and the distribution of s conditional on X is given by a known function $H(X;\omega)$. So ω is the set of parameters that determines the distribution of s.

In addition, the firm-level production function $f(k)$ is assumed to be of a Cob-Douglas form $f(k) = k^\alpha$, with decreasing returns to scale parameter α. The period utility for entrepreneurs is CRRA with coefficient of relative risk aversion ρ. I fix the risk-aversion parameter and discount factor so that the remaining structural parameters in the econometric model include the following:

1. the parameters of the conditional distribution of θ types, $\{\lambda_s\}_{s=1}^S$, and the parameters of the marginal distribution of the characteristic groups s: ω;

2. the parameters of the productivity distribution conditional on s groups and θ types, $\left\{\overline{a}_s, \underline{a}_s, g(\underline{a}_s|\theta^H, s), g(\underline{a}_s|\theta^L, s)\right\}_{s=1}^S$;

3. the parameters of the variance of the repayment shocks, $\{\sigma_s\}_{s=1}^S$, the parameter of the default costs, δ, and the parameter of the firms' valuations of their collateral relative to the banks' valuations of these collateral: η; and

4. the parameters of the returns to scale, α;

4.2 Identification

I start the discussion by considering how the structural parameters are identified *conditional on s*. Then I discuss how the distribution of s is identified conditional on the variables X that are observed in the data. Here I mostly provide heuristic identification arguments.

1. The observed response of the contract terms to past ratings helps determine how different the high and low types are in terms of their productivity distributions, or, more specifically, the ratio of $g(\underline{a}|\theta^L)$ to $g(\underline{a}|\theta^H)$.[33] Consider an extreme case where a high type will never have a low realization and a low type always has a low realization (i.e., the ratio $g(\underline{a}|\theta^L)/g(\underline{a}|\theta^H)$ approaches ∞). In this case, firm types are completely learned after the first loan, which means that we would see the contract terms change from period one to period two but stay constant afterwards (conditional on the funding costs). At the other extreme, the high type has almost the same probability of realizing a low productivity as the low type does (i.e., the ratio is near 1). In this case, each observation of the productivity realization does not lead agents to update their beliefs very much and learning takes place slowly, over a long period of time.[34] In the data, this would imply a very slow rate of change to contracts in response to changes in firm ratings.

The rate of change of contractual responses to ratings is captured by the coefficient of the interaction term, β_3, of regression (2). Consider a group of firms that has at least one poor rating in their history, with some firms' poor ratings occurring early in the relationship and others late. If learning is fast, beliefs about firm type in later periods are sticky. A poor rating that arrives late has a smaller effect on the contract terms than a poor rating that happens early. This means that the sign of β_3 is the opposite of β_2, and the magnitude of β_3 is large compared to that of β_2. On the contrary, when learning is slow, new information arriving in later periods can be almost as important as in the early periods, so we would expect β_3 to be small in magnitude. From Table 2, we find that β_3 does have the opposite sign of β_2 in all three regressions, suggesting that learning does seem to take place. The magnitudes of β_3 are relatively small compared to those of β_2, indicating that learning is still happening during the latter half the relationship.

2. Once the ratio of $g(\underline{a}|\theta^L)$ to $g(\underline{a}|\theta^H)$ is determined, the separate values of $g(\underline{a}|\theta^H)$ and $g(\underline{a}|\theta^L)$, as well as the parameter of θ-type distribution, λ, can be identified from the first two

[33] s is omitted here since we are conditional on the same s throughout this part. The ratio $g(\underline{a}|\theta^L)/g(\underline{a}|\theta^H)$ lies in $(1,\infty)$ by definition.

[34] This can be seen mathematically from the Bayesian rule:

$$p_t = \frac{p_{t-1}}{p_{t-1} + (1-p_{t-1})g(\underline{a}|\theta^L)/g(\underline{a}|\theta^H)}$$

When $g(\underline{a}|\theta^L)/g(\underline{a}|\theta^H) \to \infty$, p_t shrinks to zero immediately; when $g(\underline{a}|\theta^L)/g(\underline{a}|\theta^H) \to 1$, the belief hardly updates.

periods' rating data. Intuitively, this is because when there is only one firm type, the average firm rating in period two should be similar to that in period one, assuming that the sample attrition for reasons other than defaults are exogenous. When the type distribution is, say, half high and half low, since low types are more likely to have low repayment realizations and thus poor ratings and the default probability conditional on a poor rating is higher, more low types drop out of the cohort than high types, resulting in positive *selection*. As a consequence, the second period's average firm rating is better than that of the first period and the magnitude of the improvement is informative of the type distribution. In addition, the level of the average firm's rating helps pin down the productivity distributions for each type.

We can see this in a more formal way. Note that the mean of the poor rating in period one is (conditional on s)

$$E[PoorRating_{i1}] = \lambda g(\underline{a}|\theta^H) + (1-\lambda) g(\underline{a}|\theta^L) \tag{8}$$

Suppose D_1 is the fraction of the defaulted firms at the end of the first period, which can be observed from the data. Then the proportion of the low types within these defaulted firms is given by

$$\eta_1 = \frac{(1-\lambda) g(\underline{a}|\theta^L)\underline{d}_1}{\lambda g(\underline{a}|\theta^H)\underline{d}_1 + (1-\lambda) g(\underline{a}|\theta^L)\underline{d}_1} \tag{9}$$

where $\underline{d}_1$ is the probability of default conditional on a poor rating (or a low realization of a) in the first period. Importantly, this probability is only determined by the *belief* the firm holds after its first loan and not by the firm's true type, since we assume that firms do not observe their true type. In other words, if a high-type firm and a low-type firm both have low realizations (and thus poor ratings) in the first period, they will have the same beliefs when they make their default decisions on their first loans and they will have the same default probability (conditional on s). Thus, the same period-one default probability conditional on a poor rating appears on both the numerator and the denominator of Equation (9), so it can be cancelled out, leaving η_1 a function only of $g(\underline{a}|\theta^H), g(\underline{a}|\theta^L), \lambda$.

Given η_1 and D_1, we can find the proportion of low types at the beginning of period two, $(1-\lambda-D_1\eta_1)/(1-D_1)$, and the proportion of high types $(\lambda - D_1(1-\eta_1))/(1-D_1)$. So the mean of the poor ratings in period two is (conditional on s):

$$E[PoorRating_{i2}] = \frac{\lambda - D_1(1-\eta_1)}{1-D_1} g(\underline{a}|\theta^H) + \frac{1-\lambda-D_1\eta_1}{1-D_1} g(\underline{a}|\theta^L) \tag{10}$$

Given the data on firm ratings, we can measure the right-hand side of Equations (8) and (10) using the data, and the left-hand side of the two equations can be re-written in terms of $g(\underline{a}|\theta^L)/g(\underline{a}|\theta^H)$, $g(\underline{a}|\theta^H)$, and λ, in which only the latter two objects are unknown. So they can be jointly determined by the two equations.

3. The parameter of default costs δ and the variance of repayment shock σ can be pinned down by the level of the default rate and its response to the funding costs. Intuitively, both parameters can control the observed level of the default rate, but σ also determines how "exogenous" the default events are.

Consider an extreme case of σ close to infinity. In this case default becomes almost irrelevant with comparison of the firms' values associated with whether they repay or default: that is, it is close to a random event that happens exogenously to firms. On the other extreme, when σ is close to zero, then default is almost deterministic and it almost surely will not happen if the values these firms assign to repaying are strictly larger than the values they assign to defaulting. It follows that σ should determine how sensitive defaults are with respect to exogenous variations in the values these firms assign to repaying, say, compared to variations in the funding costs, which affect the interest rates and repayment. Thus, co-movements of the default rates and the funding costs can identify σ. Once σ is pinned down, the observed level of the default rates can then pin down the parameters of the default costs, δ, since it controls these firms' default incentives.

4. The parameters of the collateral costs can be identified by the observed usage of collateral after pinning down δ and σ. The role of collateral in this model is to provide a disincentive for the firms to default. Once the levels of the default incentives are pinned down by δ and σ, the most important factor affecting the usage of collateral is the parameter determining the firms' valuations of their collateral, given by η.

5. The returns to scale parameter α can be identified from the joint distribution of the initial loan size and the funding costs. This is because conditional on the initial beliefs (which are given by the true distribution of firm types), a higher α means a higher sensitivity to the loan size with respect to the funding costs.

6. The high and low realizations of productivity $\{\overline{a}, \underline{a}\}$ can be identified from the joint distribution of the loan sizes and the ratings history at the end of the firms' and bank's relationship. Once we have identified the learning parameters, we can pin down the difference in the beliefs for firms with different ratings histories at the end of the sample period and how this difference translates into different loan sizes depending largely on the levels of productivity $\{\overline{a}, \underline{a}\}$ and the

returns to scale parameter α. Once α is identified from the previous step, we can identify $\{\overline{a}, \underline{a}\}$.

So far, the identification argument is conditional on firm characteristic group s. The (parametric) distribution of group s conditional on the observed firm variables X can be pinned down by the distribution of the initial contract terms conditional on firms' X, since the initial contract terms are completely determined by each firm's group, s. In other words, the pattern of how firms with certain values of X tend to share similar initial contract terms helps identify the link between group s and firm variables X, which are summarized in the set of parameters ω.

4.3 Functional Forms

Repayment Shocks. I assume that the repayment shock ϵ_t has a normal distribution with mean zero and variance σ_s.[35]

Firm Characteristics. In the estimation, I let the number of firm characteristics groups $S = 3$, which means that a firm belongs to one of the three characteristics groups $s = 1,2,3$. I use an ordered probit model to model the distribution of s conditional on X. Specifically, $s = j \Leftrightarrow \omega_j^c \leq X'\omega^f + e \leq \omega_{j+1}^c$, for $j = 1,2,3$. Here e is a standard normal random variable, ω^f is a set of linear coefficients associated with X, and $\{\omega_j^c\}_{j=1}^4$ is the set of cut points with $\omega_1^c = -\infty$ and $\omega_4^c = +\infty$. In other words, the distribution of the characteristics groups conditional on X is given by $P(s = j|X) = \Phi(\omega_{j+1}^c - X'\omega^f) - \Phi(\omega_j^c - X'\omega^f)$, where $\Phi(\cdot)$ is the standard normal distribution function.

To construct X, I consider the following six firm-level attributes: industry, firm size, region, registered capital, cohort, and internal credit assessment. All of the variables except for the registered capital are categorical, and the registered capital is continuous. Thus, I generate dummy variables for each categorical variable, for a total of 20 dummy variables. In all, there are 20+1 variables for X.

Distribution of Productivity Conditional on Firm Type. Conditional on the same θ, firms with different characteristics s have different supports of productivity distributions. I make two simplifying assumptions with regard to how s affects the productivity distribution: First, firms of the same quality type all share the same probability of success, regardless of their characteristics. In other words, $g(\overline{a}_s|\theta, s)$ is the same across s; thus, I drop s from the parenthesis and denote it as $g(\overline{a}_s|\theta)$. Second, I assume that the difference between $\overline{a}_s$ and $\underline{a}_s$ does not vary with

[35]The normal distribution assumption aids in the calculation of the term $E[\epsilon|repay]$, which is a truncated mean. Since the expectation of a truncated normal distribution has a simple analytical form, I assume a normal distribution for ϵ.

s, and I define this difference as $\Delta a \equiv \overline{a}_s - \underline{a}_s$. In other words, by fixing θ, the characteristics, s, only horizontally shift the productivity distribution and do not change its shape.

Lending Cost I use the benchmark one-year deposit rate as the empirical counterpart of c. Using data on the deposit rates from 2009 to 2018, I discretize the data into four bins: 1.75 percent, 2.25 percent, 2.75 percent, and 3.25 percent, and I estimate the transition matrix, which is then transformed to a transition matrix for the yearly frequencies.

4.4 Estimation

The estimation is based on a simulated method of moments. The set of parameters to be estimated are collected in vector Ξ and listed on Table 3. Following Boualam (2018), the discount factor β is set to 0.9542 and the relative risk-aversion parameter, ρ, to 0.6.

Table 3: List of Parameters

Parameters to be Estimated	Definition
$\{\bar{a}_s\}_{s=1}^3$	High productivity realization for each characteristics group s
Δa	Difference between high and low productivity realization
$\{\lambda_s\}_{s=1}^3$	Fraction of high type within each characteristics group s
$\omega^f, \omega_2^c, \omega_3^c$	Coefficients and cut points in determining s
$g(\bar{a}\|\theta^h)$	Prob. of high type realizing high productivity
$g(\bar{a}\|\theta^l)$	Prob. of low type realizing high productivity
$\{\delta_s\}_{s=1}^3$	Default costs as a fraction of firm value for each s
η	Firm's collateral valuation relative to bank's
$\{\sigma\}_{s=1}^3$	Variance of liquidity shock for firms in each s
α	Returns to scale parameter
mc	Bank's marginal cost of lending (in addition to funding cost)

Given a vector of primitives Ξ, for each $s \in \{1,2,3\}$, I solve for the value function $W^s(p,c)$, the associated policy functions $r^s(p,c)$, $k^s(p,c)$, $z^s(p,c)$, as well as the repayment probabilities conditional on each productivity outcome.

Then for each firm $i = 1$ to $N = 6,358$,

1. I draw its quality type θ_i, based on which I draw a panel of productivity realizations. Since high/low productivity realizations directly translate into good/poor ratings, I can obtain the binary ratings for each firm: $\{rating_{it}\}$ for $t = 1$ to $T = 7$.

2. Based on the initial state of the funding cost c_{i1}, I simulate a path of the funding cost for T periods ahead: $\{c_{it}\}$ for $t = 1, ..., T$.

3. I also draw the characteristics, s, it belongs to based on its X. The s characteristics determine the initial prior beliefs p_{i1}, which, combined with the panel of productivity realizations, completely determine the path of the beliefs $\{p_{it}\}$.

4. Now we have the funding costs and the beliefs (p_{it}, c_{it}) for $t = 1,...,T$, so we can apply the policy functions $r^s(p,c)$, $k^s(p,c)$, $z^s(p,c)$ to find the contract terms $\{r_{it}, k_{it}, z_{it}\}$ in each period.

5. The default probabilities can also be found using both (p_t, c_t), and the simulated firm productivity outcomes, based on which the default outcomes $\{d_{it}\}$, are drawn. Defaulted firms are then removed from the simulated data, following their default.

I estimate the primitive, using the method of moments, by minimizing the differences between the model prediction and the data counterpart of the following moments, conditional on the exogenous firm variables X: (1) the contract terms and defaults $o_{it} = \{r_{it}, k_{it}, z_{it}, d_{it}\}$; (2) the firm ratings $rating_{it}$; (3) the covariance between the firm ratings, contract terms and default $rating_{it} \times o_{it}$; and (4) the covariance between the funding costs, contract terms and default $c_i \times o_{it}$. Formally, let $m_{it}(\Xi) = \left\{o_{it}, rating_{it}, rating_{it} \times o_{it}, c_i \times o_{it}\right\}$ and its data counterpart be $\hat{m} = \left\{\hat{o}_{it}, rat\hat{i}ng_{it}, rat\hat{i}ng_{it} \times \hat{o}_{it}, \hat{c}_{it} \times \hat{o}_{it}\right\}$, then the moment restrictions used in the estimation (detailed lists are on Table 16) are

$$g_t(\Xi) = \mathbb{E}_{\{i\}}\left[m_{it}(\Xi) - \hat{m}_{it}|X_i\right], \forall t = 1,...,T \tag{11}$$

I estimate the following, where $\hat{\Gamma}$ is the optimal weighting matrix obtained from an initial estimation with the identity weighting matrix and $g(\Xi) = \{g_t(\Xi)\}_{t=1}^{T}$.

$$\hat{\Xi} = \arg\min_{\Xi} g(\Xi)\hat{\Gamma} g(\Xi)$$

In total, I have 21 covariate vectors and $21 \times (28 + 7 + 28 + 28) = 1,911$ moment restrictions to estimate 35 parameters.

5 Estimation Results

In this section, I first discuss my parameter estimates. I then provide several pieces of evidence to evaluate the model fit. I report the parameter estimates on Table 4. Panel A of Table 4 shows estimates for the parameters that vary across characteristics groups, s, which include the high productivity realization, $\bar{a}_s$, the fraction of the firm value that can be salvaged after a firm de-

Table 4: Parameter Estimates

Panel A: Parameters varying across characteristics s

	$\bar{a}_s$	δ_s	λ_s	σ_s
$s=1$	2.807	0.937	0.298	3.3923
$s=2$	3.405	0.8946	0.301	3.1539
$s=3$	4.155	0.939	0.411	1.089

Panel B: Parameters that do not vary with s

Δa	$g(\theta^h)$	$g(\theta^l)$	α	η	mc
0.532	0.963	0.801	0.594	1.146	0.025

faults, δ_s, the fraction of high-type firms conditional on s, λ_s, and the variance of the liquidity shock, σ_s. This implies that there is considerable firm heterogeneity. The relatively large range in $\bar{a}_s$ translates into a large range for the initial loan sizes: ranging from RMB 2.43 million to RMB 8.93 million. The equilibrium default probabilities also have a wide range (1.3 percent to 12.03 percent), largely due to the relatively large variation in δ_s and σ_s across s.

Panel B of Table 4 shows estimates for the parameters that do not vary across characteristics groups, including the difference between high and low productivity realizations δa, the probability of high-quality types realizing high productivity, $g(\theta^h)$, the probability of low types realizing high productivity, $g(\theta^l)$, the return-to-scale parameter, α, the firms' valuations of own collation relative to the bank's η, and the constant marginal cost of lending mc. Reading from $g(\theta^h)$ and $g(\theta^l)$, the high type has a 3.7 percent chance of realizing low productivity , while low types have about a 20 percent chance. To interpret η, recall that the mean level of z is about 0.55, which is the bank's recovery rate on a defaulted loan. The parameter $\eta = 1.146$ implies that the firm's valuation of its collateral is, on average, $(1+1.146)*0.55 = 1.18$ times the loan size, which implies a haircut of about 18 percent. The marginal cost of lending is estimated to be 2.5 percent. This is close to the estimated profit rate of 2.2 percent reported in Porter et al. (2009).

5.1 Model Fit

The model predicts five observed variables (interest rates, loan sizes, collateral requirement, firm ratings, and default) over a period of seven years. I first check how these variables track the ones observed in the data over the seven years across all firms, then I break this down by different firm characteristics. The results are shown on Table 5.

Another important aspect of the model is the co-movement of firm ratings and contract terms.

Table 5: Model Fit: Contract Terms, Ratings, and Loan Outcomes

t	r		k		z		PoorRating		Default	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
1	7.043	7.351	4.482	5.726	0.573	0.592	0.081	0.140	0.040	0.035
2	7.051	7.036	5.009	5.868	0.573	0.569	0.085	0.140	0.066	0.032
3	6.826	6.725	5.371	6.006	0.570	0.561	0.084	0.139	0.081	0.030
4	6.577	6.467	5.839	6.134	0.568	0.557	0.082	0.138	0.051	0.027
5	6.420	6.255	6.485	6.250	0.565	0.555	0.089	0.139	0.079	0.026
6	6.165	6.077	7.120	6.358	0.560	0.552	0.083	0.138	0.050	0.024
7	6.004	5.923	7.572	6.458	0.535	0.551	0.089	0.138	0.080	0.022

This table compares the summary statistics of five observed variables from the data with their predicted values from the estimated model. The data consists of firms that take out one-year loans repeatedly over a seven-year period (t 1 to 7). The five variables are the interest rate r (in %), loan size k (in 1 million RMB), collateral coverage z, an indicator of a poor rating $PoorRating$, and an indicator of a loan default. Each row shows the means of the variables in the corresponding relationship year t.

Specifically, how contract terms respond to new information on firms' ratings, and how these responses change over time, speak directly to the speed of learning. To summarize this information succinctly, I use a firm fixed effects model to regress the contract terms on the relationship years, t, an indicator of these firms' poor ratings in the previous period, $PoorRating_{t-1}$, and the interaction term $t \times PoorRating_{t-1}$, as in Equation (2). I conduct this regression for firms with no poor ratings in their history (in which case the variable $PoorRating t-1$ and its interaction term with t are omitted, and we only compare the rate of change of the contract terms over time) for firms with at least one poor rating in their history (again for these firms I keep observation $t <= t_i^* + 1$, where t_i^* is the first time firm i is rated poorly). The results are shown on Table 6.

The signs of the coefficients in the regression using the simulated data are in line with the empirical results shown on Table 2, where we see a decline of the lending spreads and the collateral coverage as well as a gradual growth in the loan sizes over the course of the relationship for firms with no poor ratings (which are overwhelmingly high-type firms). As for firms with poor ratings in period t, the interest rate and the collateral coverage are likely to increase and the loan size is likely to decrease in the next period.

5.2 Inefficiencies from Agency Friction and Incomplete Information

In the estimated model, there are two sources of inefficiency: incomplete information and agency friction. Incomplete information refers to the fact that neither the banks nor the firms

Table 6: Regressions of Contract Terms on Past Ratings: Using Simulated Data

	Firms w. Good Rating History			Firms w. Poor Rating History		
	(1) Lending Spread	(2) Loan Size	(3) Collateral Coverage	(4) Lending Spread	(5) Loan Size	(6) Collateral Coverage
t	-0.00610	0.0190	-0.00277	-0.00750	0.0194	-0.00174
	(-33.59)	(118.22)	(-25.40)	(-15.17)	(59.62)	(-5.97)
$PoorRating_{t-1}$				0.0605	-0.109	0.0162
				(13.92)	(-38.05)	(6.32)
$t \times PoorRating_{t-1}$				-0.000478	-0.0107	-0.000995
				(-0.53)	(-18.13)	(-1.88)
Observations	18195	18195	18195	7613	7613	7613

The t statistics are in parentheses. Other controls include firm fixed effects and funding costs fixed effects.

know the firms' true types θ. This leads to a mismatch between the contract terms and the firms: Firms are assigned contract terms that are not optimal for their true types. Although learning can reduce uncertainty and alleviate the implications from having incomplete information in subsequent periods, the reduction in the uncertainty might not be complete. To quantify the inefficiencies caused by incomplete information, I consider the counterfactual case with full information, where firms' types are fully known to both the banks and the firms from the beginning. In this case, the firms' value functions are type-specific and the dynamics in the contract terms are not driven by learning.

The agency friction comes from the fact that firms cannot fully internalize the social costs of default; thus, their private default decision is not aligned with the social planner's problem. The social costs of default include the value lost when collateral is surrendered and liquidated to recover the banks' losses. Such costs are not internalized by the firms, which leads to excessive defaults compared to the socially optimal level. This is also referred to as strategic default or opportunistic default Kang and Jang (2019).

To see how contract terms and default rates are different in the absence of agency friction, I consider the planner's problem, where the contract can enforce the repayment behaviour. This means that, in the planner's contracting problem, the choice of variables not only includes (r, k, z) but also the repayment policy (to repay or default). For example, in the case with full information and no agency friction, the contract, as written, could solve the problem of maximizing the joint value between the firm and the bank, subject to the bank earning zero profits:

$$V_\theta = \max_{r,k,z} E \max\left\{ \underbrace{u\left(y - rk\right) + rk + \beta V_\theta}_{v_{\text{repay}}} - \epsilon, \underbrace{u\left(y - (1+\eta)zk\right) + zk + \beta\delta V_\theta}_{v_{\text{default}}} \right\} - ck$$

$$\text{s.t. } 0 = -ck + rk - D(r-z)k$$

where default probability $D \equiv 1 - \Phi\left(\frac{v_{\text{repay}} - v_{\text{default}}}{\sigma}\right)$.

I calculate (counterfactual) short-term contracts in four scenarios: (1) no agency friction and full information; (2) no agency friction and incomplete information; (3) with agency friction and full information; and (4) with agency friction and incomplete information (the estimated baseline model). For each firm, I find the associated cumulative default probability over the seven-year period, the total amount of capital and output over the seven-year period, the firm size measured by its output in the last period, and the initial firm value in the four equilibria. A summary of these statistics by firm quality type is show on Table 7.

Comparing column (3) with column (4), we can see that incomplete information leads to a misallocation of capital: the high types in the baseline case are assigned lower capital than the full-information case, as in column (3), whereas the low types in the baseline case are assigned higher capital than the full-information case. This mismatch of capital and firm productivity leads to lower overall output. In addition, due to incomplete information, there is an issue of mispricing, where the high types in the baseline case are charged higher interest rates than are the those in the full-information case, and the low types are charged lower interest rates than the those in the full-information case. Such mispricing of the default risk and capital misallocation leads to negative impacts on the outcomes of high-type firms, as shown in Figure 4. Learning, to some degree, can reduce the negative impacts over the course of the relationship. As highlighted in Figures 5 and 6, however, there is still substantial mispricing and misallocation.[36]

Comparing column (2) with column (4), we can see that in the benchmark scenario with no agency friction (i.e., in the planner's problem), default would be extremely rare and capital flow would be significantly enhanced. Specifically, the number of defaulted firms would be reduced by 71.63 percent compared to the baseline case and total output would increase by 14.83 per-

[36] Here I measure mispricing by $(r_i^{CI}/\hat{r}_i - 1) \times 100\%$, where r^{CI} is the interest rate in the counterfactual scenario of complete information and $\hat{r}$ is the interest rate in the baseline model. I measure capital misallocation by $(k_i^{CI}/\hat{k}_i - 1) \times 100\%$, where k^{CI} is the loan size in the counterfactual scenario of complete information and $\hat{k}$ is the loan size in the baseline model. The mispricing (of capital) in the initial period is ≈ 1.7 percent, and the misallocation of capital is ≈ 4.8 percent.

Table 7: Comparisons Across the Baseline Model and Models with Either One or No Source of Inefficiency

Scenario	Benchmarks			Baseline
	(1) No AF Full Info.	(2) No AF Inc. Info.	(3) AF Full Info.	(4) AF Inc. Info.
	High-Type Firms			
Cumulative Default Prob.	.0877069	.0968651	.3237055	.3325114
Total Capital	37.72948	36.20978	32.05138	30.71789
Total Output	65.62923	63.9358	56.94183	55.4055
Final Size	5.571345	5.380186	5.261407	5.071709
Firm Value	88.68119	87.42284	87.55135	86.39585
	Low-Type Firms			
Cumulative Default Prob.	.0934925	.0863882	.316283	.3117363
Total Capital	33.13112	34.22402	28.31473	29.27848
Total Output	58.12727	59.32193	50.78448	51.87864
Final Size	4.898396	5.039846	4.612279	4.751198
Firm Value	84.28747	85.30146	83.33981	84.25609
	All Firms			
Cumulative Default Prob.	.0909091	.0910664	.3195974	.3210129
Total Capital	35.1844	35.11071	29.98323	29.92121
Total Output	61.47708	61.38213	53.53389	53.45347
Final Size	5.198884	5.191816	4.90213	4.894314
Firm Value	86.24937	86.24871	85.22037	85.21154

This table compares outcome variables under four different scenarios. The scenarios presented in columns (1)-(3) are counterfactual scenarios: (1) no agency friction and full information; (2) no agency friction and incomplete information; and (3) with agency friction and full information. Column (4) is the baseline scenario used in the estimated model. The cumulative default probability is the probability that a firm defaults in any one of the eight periods. Total capital and total output are the sum of the borrowed capital k and the sum of the firm's output y. The final size is the firm's output y produced in the last period. Firm value is its initial expected value in period one. All monetary values are in units of 1 million RMB.

cent. In other words, the borrowing constraints in the baseline case are largely due to agency friction and less due to incomplete information. In total, agency friction accounts for the vast majority of the total welfare loss from both agency friction and incomplete information.

6 Long-Term Financing Contracts

In this section, I consider a counterfactual scenario where banks compete by offering long-term contracts. An illustrative example of a two-period contract is included in Section 6.1. This example shows features of the optimal long-term contract and explains my solution method. Building on the intuition from this example, I then describe an example of a multi-period contingent contract and present the results from my counterfactual simulation exercises.

6.1 Two-Period Contingent Contracts

To better illustrate the how long-term contracts work, I begin with the simplest case: the two-period contract. Suppose banks offer two-period contingent contracts to *new firms* at the beginning of period one and that such contracts stipulate the lending terms in the first period as well as the state-contingent lending terms for the second period. After the two-period contract ends, firms enter an infinite-horizon short-term lending market, where competition takes place as described in the previous section. In other words, the thought experiment I engage in here is about providing long-term contracts at the *beginning* of a firm's life and seeing how it affects the firm's outcomes, compared to the baseline model where firms are born into a short-term lending market.

In this model, banks compete at the beginning of period one by offering two-period contingent contracts. Let l_t denote the lending terms in period t, that is, $l_t = (r_t, k_t, z_t)$. Then the two-period contract can be written as $\{l_1, l_2(p_2, c_2)\}$, where the second-period terms l_2 are a function of state variable (p_2, c_2).

I consider two-period contracts that are *binding* conditional on repayment, meaning that as long as the borrower repays the first-period loan, the second-period borrowing activity must take place according to the lending terms pre-determined in the contract. Specifically, if the first-period loan is repaid then this implies that the lender cannot renege on the second-period terms and the borrower cannot switch to default. If the borrower defaults on the first-period loan, then the contractual relationship ends and the firm receives scrap value that equals a fraction δ of its equilibrium firm value in the short-term lending market.

By the same logic as before, a Bertrand equilibrium in this model results in maximizing the firm's value subject to the zero profit condition. However, the zero profit condition in this case means that a bank expects that its overall profit *over the course of the long-term contract* will be zero, as apposed to anticipating zero profits within each period. In other words, it is possible that the first-period expected profits are positive (negative) and the ex-ante second-period profits are negative (positive). This is because the binding long-term contract implies that there is no competition from outside banks at the beginning of the second period; thus, the second-period expected profit in each possible state is no longer constrained to zero.

To succinctly state the problem solving for the equilibrium long-term contract, I first define a few objects. Let $U(l_t; a_t)$ and $U^F(l_t; a_t)$ denote the firm's flow utility in period t for repayment and default, respectively. These are functions of this period's lending terms l_t as well as the realization of productivity a_t, which determines the level of output.

$$U(l_t; a_t) = u\left(y_t(a_t) - r_t k_t\right)$$
$$U^F(l_t; a_t) = u\left(y_t(a_t) - (1+\eta) z_t k_t\right)$$

Using $U(l_t; a_t)$ and $U^F(l_t; a_t)$, we can express the firm's utility at the beginning of $t = 2$ as a function of the second-period lending terms l_2 and state variables p_2, c_2:

$$\begin{aligned} w_2(l_2, p_2, c_2) = E_{a_2,\epsilon_2}\Big[\max\{ & U(l_2; a_2) - \epsilon_2 + \beta E\left[W(p_3, c_3)|a_2\right], \\ & U^F(l_2; a_2) + \beta\delta E\left[W(p_3, c_3)|a_2\right]\}\Big| p_2, c_2\Big] \end{aligned} \tag{12}$$

where $W(\cdot,\cdot)$ is the firm's value function in the short-term lending equilibrium. Here we maintain the same assumption as for the baseline model with regard to default: defaulting firms receive scrap value that equals a fraction δ of the firm's value W.

It follows that the default probability in the second period, conditional on the realization of a_2, can be written as

$$\Phi_2(l_2, p_2, c_2; a_2) = \Pr\Big(\epsilon_2 \le U(l_2, a_2) - U^F(l_2, a_2) + \beta(1-\delta) E\left[W(p_3, c_3)|a_2, p_2, c_2\right]\Big) \tag{13}$$

Using the default probability function, we can find the expected bank profits for the second period:

$$\pi_2(l_2, p_2, c_2) = -(c_2 - z_2)k_2 + E_{a_2}\left[\Phi_2(l_2, a_2, p_2, c_2)(r_2 - z_2)k_2 \middle| p_2\right] \tag{14}$$

At the beginning of the first period, for any given lending policy in periods 1 and 2, (l_1, l_2), we can express the firm's expected utility as a function of l_1 and l_2 as well as state variables (p_1, c_1):

$$w_1\left(l_1, l_2, p_1, c_1\right) = E_{a_1,\epsilon_1}\Big[\max\{U(l_1; a_1) - \epsilon_1 + \beta E\left[w_2\left(l_2, p_2, c_2\right)|a_1\right], \\ U^F(l_1; a_1) + \beta\delta E\left[W(p_2, c_2)|a_1\right]\}\Big|p_1, c_1\Big] \tag{15}$$

The difference in the future values for the choice of repayment and default comes from the assumption that default breaks the long-term contract and leaves the firm with a scrap value that equals a fraction δ of firm value W. This leads to the following expression for the default probability in the first period, conditional on a_1:

$$\Phi_1(l_1, l_2, p_1, c_1; a_1) \equiv \Pr\Big(\epsilon_1 \leq U(l_1, a_1) - U^F(l_1, a_1) + \beta E\left[w_2\left(l_2, p_2, c_2\right) - \delta W(p_2, c_2)|a_1, p_1, c_1\right]\Big) \tag{16}$$

Thus the bank's expected profit over the two periods is

$$\pi_1(l_1, l_2, p_1, c_1) = -(c_1 - z_1)k_1 + E_{a_1}\left[\Phi_1(l_1, l_2, p_1, c_1; a_1)(r_1 - z_1)k_1 + \beta E\left[\pi_2(l_2, p_2, c_2)|a_1\right]\big|p_1, c_1\right] \tag{17}$$

6.1.1 Original Problem

The original problem a bank solves to obtain the lending terms of the two-period contract in equilibrium is as follows:

$$\begin{aligned} &\max_{l_1, l_2(p_2,c_2)} w_1\left(l_1, l_2(p_2, c_2), p_1, c_1\right) \\ &\text{s.t.} \quad 0 = \pi_1\left(l_1, l_2(p_2, c_2), p_1, c_1\right) \end{aligned} \tag{18}$$

In other words, the bank chooses its first-period lending terms, l_1, as well as its second-period terms contingent on state (p_2, c_2), which maximize the firm's utility subject to zero expected profits over the course of the two periods.

In the case where productivity a_t has a two-point distribution, then, conditional on p_1, the next period's belief p_2 has two possible values:$p_2\left(p_1|a_1 = \bar{a}\right)$ and $p_2\left(p_1|a_1 = \underline{a}\right)$. Continuing the assumption that c_t takes four values, there are $2 \times 4 = 8$ states for the second period. Since the lending terms, l, include three variables, r, k, z, in total there will be $3 + 8 \times 3 = 27$ choice variables in this problem. And if we later extend the two-period contracts to multiple periods, the number of choice variables will explode. Thus, in this paper, I consider an alternative approach to solve for equilibrium long-term contracts.

6.1.2 Transformed Problem

Inspired by Albuquerque and Hopenhayn (2004), I transform the problem into one where the lender chooses the summary statistics for the future contract instead of fully spelling out each lending term. This also helps form a recursive representation of the original problem (18).

Note that given each state in period two, (p_2, c_2), for a certain level of expected bank profits, $\bar{\pi}$, there is a unique solution to the constrained maximization problem of firm value, subject to the bank's expected profits equal to $\bar{\pi}$, and the maximum firm value is defined as $W_2(\bar{\pi}, p_2, c_2)$:

$$W_2(\bar{\pi}, p_2, c_2) \equiv \max_{l_2} \ w_2(l_2, p_2, c_2) \ \text{ s.t. } \ \bar{\pi} = \pi_2(l_2, p_2, c_2) \tag{19}$$

Using $\bar{\pi}$ as an "index" for the second-period contract contract contingent on the state (p, c), I re-write the period-one objects (15)-(17) using $\bar{\pi}$ instead of l_2: (I suppress the parenthesis of $\bar{\pi}$ in (15') and (16'))

$$\begin{aligned} w_1(l_1, \bar{\pi}, p_1, c_1) = E\Big[\max\big\{ & U(l_1; a_1) - \epsilon_1 + \beta E\left[W_2\left(\bar{\pi}(p_2), p_2, c_2\right) | a_1 \right], \\ & U^F(l_1; a_1) + \beta\delta E\left[W(p_2, c_2) | a_1 \right] \big\} \Big| p_1, c_1 \Big] \end{aligned} \tag{15'}$$

$$\Phi_1\left(l_1, \bar{\pi}, p_1, c_1; a_1\right) = \Pr\Big(\epsilon_1 \leq U(l_1, a_1) - U^F(l_1, a_1) + \beta E\left[W_2\left(\bar{\pi}, p_2, c_2\right) - \delta W(p_2, c_2) | a_1, p_1, c_1 \right] \Big) \tag{16'}$$

$$\pi_1(l_1, \bar{\pi}, p_1, c_1) = -(c_1 - z_1) k_1 + E_{a_1}\left[\Phi_1\left(l_1, \bar{\pi}, p_1, c_1; a_1\right)(r_1 - z_1) k_1 + \beta E\left[\bar{\pi}(p_2, c_2) | a_1 \right] \Big| p_1, c_1 \right] \tag{17'}$$

Thus, at the beginning of period one, when the banks prepare their contracts, instead of choosing three lending terms for each state of period two, they choose an index $\bar{\pi}$:

$$\begin{aligned} \max_{l_1, \bar{\pi}(p_2, c_2)} \ & w_1\left(l_1, \bar{\pi}(p_2, c_2), p_1, c_1\right) \\ \text{s.t. } \ & 0 = \pi_1\left(l_1, \bar{\pi}(p_2, c_2), p_1, c_1\right) \end{aligned} \tag{20}$$

The optimal choice of $\bar{\pi}$ corresponds to a set of lending terms in the second period, as described in (19). In this problem, the number of choice variables is reduced to $3 + 2 \times 4 = 11$, which is less than half of those in the original problem (18).

Still, this is not a small number of choice variables by any measure, so to further simplify the problem, I restrict myself to the case where the choice of $\bar{\pi}$ is a function of only p_2, and not c_2. This is *not* to say that the next period's interest rates do not vary with its funding costs; rather, this is akin to keeping the bank's markups the same across the next period's funding costs. To put it another way, the second-period interest rates still vary with the cost realizations but in a way that is not about insuring firms against those shocks and not treating costs as merely an anchor. By doing so, I focus on whether and how the contract terms ensure against individual firm risks and not against aggregate cost shocks. In the end, the version of the problem I solve in the two-period setting is the following:

$$\begin{aligned} \max_{l_1,\bar{\pi}(p_2)} \; & w_1\left(l_1,\bar{\pi}(p_2),p_1,c_1\right) \\ \text{s.t. } \; & 0=\pi_1\left(l_1,\bar{\pi}(p_2),p_1,c_1\right) \end{aligned} \tag{21}$$

The number of choice variables in this problem is $3+2=5$, making the extension to multiple periods computationally feasible. In the numerical exercise, I first solve the second-period problem (19) under each combination of $\bar{\pi}$, p and c, where $\bar{\pi}$ and p are discretized on a grid and c takes four discrete values.[37] I interpolate in between the grids of $\bar{\pi}$ to obtain the function $W_2(\cdot,p,c)$ for each combination of p and c.[38] In the first period, for any p_1 on the grid of p and for each of the four values of c_1, I find the two possible next-period beliefs, p_2, conditional on high and low realizations of a_1, denoted as $p_2^H=p_2\left(p_1|a_1=\bar{a}\right)$ and $p_2^L=p_2\left(p_1|a_1=\underline{a}\right)$. Then I plug functions $W_2(\cdot,p_2^H,c_2)$ and $W_2(\cdot,p_2^L,c_2)$ for each of the four values of c_2 into the first-period problem (21), which is solved to obtain the optimal first-period lending terms, l_1, as well as indices $\bar{\pi}$ following the high and low realizations: $\bar{\pi}(p_2^H)$ and $\bar{\pi}(p_2^L)$. Finally, using the chosen indices $\bar{\pi}(p_2^H)$ and $\bar{\pi}(p_2^L)$, I can back out the associated optimal second-period lending terms under each value of c_2, that is, $l_2(p_2^H,c_2)$ and $l_2(p_2^L,c_2)$. Thus, for each initial state (p_1,c_1), I obtain period-one terms $l_1(p_1,c_1)$ and state-contingent period-two terms $l_2\left(p_2(p_1|a_1),c_2\right)$ for eight possible states, and I also derive the equilibrium default probabilities in each period as functions of (p_1,c_1).

6.2 Multi-period Contingent Contracts

The model with two-period contracts can be naturally extended to multiple periods using the transformation approach introduced in Section 6.1.2. Suppose the duration of the long-term

[37] The grid of p is the same as in the baseline estimation, from 0 to 1 with step length 0.01. The grid of $\bar{\pi}$ is -1 to 1 with step length 0.01, and this grid is large enough in the sense that the chosen values of $\bar{\pi}$ in equilibrium do not hit the bounds of the grid.

[38] I use cubic spline interpolation to ensure smoothness and facilitate the first-period maximization problem.

contingent contract is $T \geq 2$. Note that the horizon of the model is still infinite, similar to the short-term model, and this makes firms' outcomes comparable between the long- and short-term models. Specifically, newborn firms take T-period contingent contracts, and after T periods they enter the infinite-horizon short-term lending market. Again we are looking at the effects of long-term contracts at the beginning of a firm's life.

Assumptions on defaults are same as in Section 6.1, where default triggers the severance of the lending relationship between the firm and the current lending bank, with the defaulted firm's salvage value equals a fraction $\delta \in [0,1]$ of the equilibrium firm value on the spot market. Another way to interpret this is that defaulted firms only regain access to the spot market with constant probability $\delta \in [0,1]$. The spot market is filled with short-term contracts only, which are equivalent to the contracts in the no-commitment case, so the firm value in this market is given by $W(p,c)$, where W is defined in Equation (6).

Following the notations introduced in Section 6.1.2, I start with the last period's problem. As in (19), I define the last period's firm value as a function of index $\bar{\pi}$, p_T and c_T

$$W_T(\bar{\pi}, p_T, c_T) \equiv \max_{l_T} w_T(l_T, p_T, c_T) \text{ s.t. } \bar{\pi} = \pi_T(l_T, p_T, c_T) \tag{22}$$

where w_T is the firm's utility as a function of lending terms l_T and states (p_T, c_T):

$$\begin{aligned} w_T(l_T, p_T, c_T) = E_{a_T,\epsilon_T}\Big[& \max\big\{U(l_T; a_T) - \epsilon_T + \beta E\left[W(p_{T+1}, c_{T+1})|a_T\right], \\ & U^F(l_T; a_T) + \beta\delta E\left[W(p_{T+1}, c_{T+1})|a_T\right]\big\}\Big| p_T, c_T\Big] \end{aligned} \tag{23}$$

It follows that the default probability in the second period, conditional on the realization of a_T, can be written as

$$\Phi_T(l_T, p_T, c_T; a_T) = \Pr\Big(\epsilon_T \leq U(l_T, a_T) - U^F(l_T, a_T) + \beta(1-\delta) E\left[W(p_{T+1}, c_{T+1})|a_T, p_T, c_T\right]\Big) \tag{24}$$

Using the default probability function, we can find the expected bank profit in period T:

$$\pi_T(l_T, p_T, c_T) = -(c_T - z_T)k_2 + E_{a_T}\left[\Phi_2(l_T, a_T, p_T, c_T)(r_T - z_T)k_2 \middle| p_T\right] \tag{25}$$

The only difference between the multi-period case and two-period case is the addition of interim-period t, where $t < T$ and $t > 1$. For the interim period, we need to derive the firm's value function as a function of the *current* period's index $\bar{\pi}$ and states (p_t, c_t) by maximizing over this period's lending terms l_t and *next periods*' state-contingent indices $\bar{\pi}'(p_{t+1})$.

$$W_t(\bar{\pi}, p_t, c_t) \equiv \max_{l_t, \bar{\pi}'(p_{t+1})} w_t\left(l_t, \bar{\pi}'(p_{t+1}), p_t, c_t\right)$$
$$\text{s.t. } \bar{\pi} = \pi_t\left(l_t, \bar{\pi}'(p_{t+1}), p_t, c_t\right) \tag{26}$$

where w_t is different from w_T in its future value:

$$w_t\left(l_t, \bar{\pi}', p_t, c_t\right) = E_{a_t, \epsilon_t}\Big[\max\{U(l_t; a_t) - \epsilon_t + \beta E\left[W_{t+1}\left(\bar{\pi}', p_{t+1}, c_{t+1}\right)|a_t\right],$$
$$U^F(l_t; a_t) + \beta\delta E\left[W(p_{t+1}, c_{t+1})|a_t\right]\}\Big| p_t, c_t\Big]$$

and the bank's expected profit function π_t is

$$\pi_t(l_t, \bar{\pi}', p_t, c_t) = -(c_t - z_t)k_t + E_{a_t}\left[\Phi_t(l_t, \bar{\pi}', p_t, c_t; a_t)(r_t - z_t)k_t + \beta E\left[\bar{\pi}'(p_{t+1})|a_t\right]\Big| p_t, c_t\right]$$

with the conditional default probability function in period t defined as

$$\Phi_t(l_t, \bar{\pi}', p_t, c_t; a_t) \equiv \Pr\Big(\epsilon_t \leq U(l_t, a_t) - U^F(l_t, a_t)$$
$$+ \beta E\left[W_{t+1}\left(\bar{\pi}', p_{t+1}, c_{t+1}\right) - \delta W(p_{t+1}, c_{t+1})|a_t, p_t, c_t\right]\Big)$$

Since a has two possible realizations, so does p_{t+1}, which means $\bar{\pi}'(p_{t+1})$ is a two-dimensional vector. In total, there are five choice variables in this constrained optimization problem.

The first period's problem is the same as in (21). In terms of the solution method, I use backward induction, starting from the last period. Using the estimated parameters and the firm's value function in the short-term model, I first calculate $W_T(\cdot,\cdot,\cdot)$ from (22) under each combination of $\bar{\pi}$, p and c, where $\bar{\pi}$ and p are discretized on a grid and c takes four discrete values.[39] I interpolate in between the grids of $\bar{\pi}$ to obtain the function $W_T(\cdot, p, c)$ for each combination of p and c.[40]

For any $t = T-1, ..., 1$, given the function $W_{t+1}(\cdot, p, c)$, for any $\bar{\pi}_t$ on the grid of $\bar{\pi}$, any p_t on the grid of p and each of the four values of c_t, I find the two possible next-period beliefs, p_{t+1}, conditional on high and low realizations of a_t, denoted as $p^H_{t+1} = p_{t+1}\left(p_t | a_t = \bar{a}\right)$ and $p^L_{t+1} = p_{t+1}\left(p_t | a_t = \underline{a}\right)$. Then I plug functions $W_{t+1}(\cdot, p^H_{t+1}, c_{t+1})$ and $W_{t+1}(\cdot, p^L_{t+1}, c_{t+1})$ into the period-

[39]The grid of p is the same as in the baseline estimation, from 0 to 1 with step length 0.01. The grid of $\bar{\pi}$ is -1 to 1 with step length 0.01, and this grid is large enough in the sense that the chosen values of $\bar{\pi}$ in equilibrium do not hit the bounds of the grid.

[40]I use cubic spline interpolation to ensure smoothness and facilitate the first-period maximization problem.

t problem (26), where $\bar{\pi}_t$ serves as the constraint. By solving this problem, I obtain the optimal first-period lending terms l_t, as well as indices $\bar{\pi}$ following high and low realizations $\bar{\pi}(p^H_{t+1})$ and $\bar{\pi}(p^L_{t+1})$. And I interpolate in between the grids of $\bar{\pi}$ to obtain the function $W_t(\cdot, p, c)$.

I repeat the above process until I reach the first period. The only difference between $t = 1$ and $t > 1$ is that I do not need to solve the first-period problem under various values of $\bar{\pi}_1$; instead, the equilibrium condition imposes $\bar{\pi}_1 = 0$; that is, the bank's expected profits at the beginning of the first period should be zero.

After obtaining the firm's value function $W_t(\cdot, p, c)$ and policy functions $l_t(\cdot,\cdot,\cdot)$ and $\bar{\pi}'(\cdot)$, I simulate a panel of lending terms, state variables, productivity realizations, and default outcomes for each firm in the observed data, where the parameters are based on estimates of the baseline model. The results shown in the next section are under $T = 20$. (But only the first seven years' statistics are shown because the baseline model is only for seven periods).

6.3 Simulation Results

Optimal Long-Term Contracts

Optimal long-term contracts specify a schedule of future contract terms based on both the length of the contractual relationship and the performance in each year of the relationship, where performance is measured by the loan rating. I later refer to this as a *relationship- and performance-based schedule*. This shows different dynamics than the repeated short-term contract. As a first pass, we can look at Table 8, which shows means r, k, z and the default rates from the baseline and counterfactual models in each period by different firm characteristics groups. The initial interest rates in the counterfactual long-term contracts are much higher than in the short-term contracts but, later on, interest rates in the long-term contracts drastically decline and in most cases are even lower than the interest rates in the short-term contracts for the last several periods. This *front-loading* pricing scheme locks firms in by charging high initial prices and gradually letting prices decline (becoming even lower than the spot-market price) in later periods.

Why is this front-loading pricing structure optimal? This is because the promise of lower prices in the future can incentivize firms to hold on and not default, which can alleviate the problem of agency friction to some degree. The reduced agency friction also means less-constrained capital input; that is, the loan sizes in later periods of the long-term contracts are larger than in the short-term contracts. Another consequence of reduced agency friction is the need for less collateral coverage in long-term contracts, which is beneficial because it reduces the costs associated with the transferal of collateral in the case of a default. Overall we see default rates decline in each single period, and fewer defaults mean more firms can survive.

Long-term contracts also differ from repeated short-term ones in their contractual responses to firms' performance. Recall that on Table 6 the short-term contracts in the baseline model feature a jump in interest rates for firms with poor ratings in the last period. Does the optimal long-term contract still have this feature? It turns out it is the opposite: The long-term contract actually specifies a lower price if the firm is hit by a negative productivity shock and performs worse. We can see this from running the same regression as in Equation (2), where I use a firm fixed effects model to regress the contract terms on the years of the relationship, t, an indicator of a poor rating in the previous period, $PoorRating_{t-1}$, and the interaction term $t \times PoorRating_{t-1}$. I conduct this regression for firms with no poor ratings in their histories (in which case the variable $PoorRating_{t-1}$ and its interaction term with t being omitted, and we will only compare the rate of change of the contract terms over time), and for firms with at least one poor rating in their histories (again for these firms, I keep observation $t <= t_i^* + 1$, where t_i^* is the first time firm i is rated poorly). The results are shown on Table 9.

Comparing columns (1) to (3) of Table 9 with columns (1) to (3) of Table 6, we find that the signs of the coefficients in the two tables are the same but the magnitudes of the coefficients in the counterfactual long-term contracts are much bigger, meaning that the long-term contract has a more significant time trend in declining lending spreads and collateral coverage and growing loan sizes. Focusing on column (4), we find that the coefficients of $PoorRating$ and the interaction term in the counterfactual model have the opposite signs to those in the baseline model: On Table 9, the lending spreads on average decrease by nearly four percentage points for firms with poor previous performance, whereas in Table 6 it increases by around seven basis points. This reflects the contingent "insurance" function of long-term contracts, which smooth consumption across different states and protect firms against risks from uncertainty about their productivity status.

How important is the insurance structure in terms of welfare improvement compared to the intertemporal structure? To see this, I also consider another type of long-term contract that does not have the insurance structure, that is, the relationship-based schedule.

Relationship-Based Schedule

Long-term contracts with relationship-based schedules come from solving (26) with π' not being contingent on a_t. In other words, future contract terms do not directly depend on this period's performance (measured by the loan rating). For example, firms with poor performance

Table 8: Comparison of Long-Term Counterfactual Model vs. Short-Term Baseline Model by Firm Characteristics

	Interest Rate		Loan Size		Collateral Coverage		Default	
	Baseline	C.F.	Baseline	C.F.	Baseline	C.F.	Baseline	C.F.
	Characteristic Group $s = 1$							
1	7.276	20.891	2.518	2.485	0.546	0.561	0.053	0.043
2	7.105	18.885	2.529	2.511	0.545	0.421	0.053	0.041
3	6.906	16.788	2.542	2.536	0.544	0.287	0.063	0.048
4	6.775	14.483	2.550	2.558	0.544	0.160	0.055	0.040
5	6.677	12.390	2.557	2.578	0.544	0.056	0.049	0.037
6	6.622	9.997	2.561	2.611	0.543	0.000	0.061	0.046
7	6.567	7.764	2.564	2.655	0.543	0.000	0.054	0.033
	Characteristic Group $s = 2$							
1	9.830	23.422	4.072	4.094	0.619	0.633	0.110	0.088
2	9.647	19.872	4.090	4.157	0.618	0.548	0.136	0.117
3	9.469	16.669	4.106	4.190	0.618	0.432	0.134	0.102
4	9.325	13.605	4.120	4.216	0.617	0.332	0.111	0.095
5	9.254	10.574	4.128	4.235	0.617	0.248	0.114	0.086
6	9.188	6.782	4.134	4.248	0.617	0.179	0.101	0.073
7	9.105	3.889	4.144	4.261	0.617	0.142	0.119	0.097
	Characteristic Group $s = 3$							
1	5.088	15.366	8.644	8.631	0.480	0.418	0.017	0.013
2	4.882	12.650	8.688	8.677	0.479	0.299	0.013	0.012
3	4.746	9.603	8.718	8.712	0.479	0.201	0.017	0.015
4	4.602	6.358	8.748	8.746	0.478	0.133	0.012	0.012
5	4.535	3.454	8.763	8.766	0.478	0.097	0.012	0.011
6	4.482	3.270	8.774	8.777	0.478	0.096	0.018	0.019
7	4.433	3.223	8.785	8.788	0.477	0.096	0.013	0.013

This table compares the contract terms and default probabilities in both the baseline and counterfactual scenarios (C.F.). The counterfactual scenario uses the optimal long-term contract as described in Section 6.2. Interest rates are shown in annual percentage rates and loan sizes are in units of RMB 1 million.

Table 9: Regressions of Contract Terms on Past Performance: Using Counterfactual Long-Term Contracts

	Firms w. Good Rating History			Firms w. Poor Rating History		
	(1) Lending Spread	(2) Loan Size	(3) Collateral Coverage	(4) Lending Spread	(5) Loan Size	(6) Collateral Coverage
t	-2.223	0.0225	-0.112	-2.263	0.0224	-0.130
	(-400.19)	(134.16)	(-274.70)	(-192.97)	(49.02)	(-144.29)
$PoorRating_{t-1}$				-4.152	-0.137	-0.460
				(-36.67)	(-30.93)	(-52.80)
$t \times PoorRating_{t-1}$				0.491	-0.000174	0.0824
				(21.47)	(-0.20)	(46.87)
Observations	18859	18859	18859	7899	7899	7899

The t statistics are in parentheses. Other controls include firm fixed effects and funding costs fixed effects.

in the first period do not receive lower second-period interest rates than those with good first-period performance; however, in the precious case, firms with poor performance in the first period receive interest rates that are, on average, 1.59 percentage points lower in the second period than firms with a good first-period performance.

As a first pass, in Figure 7, we can check how effective it is to reduce default rates, compared to relationship- and performance-based schedules (i.e., optimal long-term contracts). We find that among high-type firms default rates are lower in the relationship-based schedule than in the relationship- and performance-based schedule, whereas for low-type firms, default rates are lower in the relationship- and performance-based schedule, as this schedule provides price discounts for firms with poor performance, which are more likely to be the low-type firms, and such price discounts lead to lower default rates. Essentially, the relationship- and performance-based schedule employs a transfer from high- to low-type firms.

To see the overall welfare implications of the two schedules, I calculate the associated cumulative default probability over the seven-year period, the total amount of capital and output taken over the period, the firm size measured by its output in the last period, the initial firm value in the baseline model with the short-term contract, the counterfactual model with relationship-based long-term contract, and the counterfactual model with the relationship- and performance-based long-term contract. A summary of these statistics by firm quality type is shown on Table 10.

Comparing column (3) with column (1), we find that with long-term contracts with fully flexible schedules the cumulative default probability over the seven-year period drops by 17.15 percent, with high-type firms' cumulative default probability dropping by 16.86 percent and low-type firms' by 17.41 percent. Total firm output over the seven-year period increases by 2.63 percent, where high-type firms' total output increases by 2.84 percent (due to a 2.88 percent increase in total capital input) and low-type firms' total output increases by 2.45 percent (due to a 2.58 percent increase in total capital input). Comparing column (2) with column (3), we find that the relationship-based schedule does a good job of capturing the vast majority of the efficiency improvement without using complex contingency contracting.

Table 10: Comparison Across Different Contracts

Scenario	Baseline	Counterfactual Long-Term Contracts	
	(1)	(2)	(3)
	Short-Term	Relationship-Based	Relationship- and Performance-Based
		High-Type Firms	
Cumulative Default Prob.	.3325114	.2736879	.2765058
Total Capital	30.71789	31.64557	31.62901
Total Output	55.4055	57.00239	56.97651
Final Size	5.071709	5.130478	5.128197
Firm Value	86.39585	86.72219	86.7231
		Low-Type Firms	
Cumulative Default Prob.	.3117363	.2585962	.2574595
Total Capital	29.27848	30.05414	30.03602
Total Output	51.87864	53.19209	53.15497
Final Size	4.751198	4.806808	4.809732
Firm Value	84.25609	84.58102	84.58202
		All Firms	
Cumulative Default Prob.	.3210129	.265335	.2659641
Total Capital	29.92121	30.76476	30.74733
Total Output	53.45347	54.89348	54.86138
Final Size	4.894314	4.951335	4.951934
Firm Value	85.21154	85.5371	85.53806

This table compares the outcomes of counterfactual scenarios with those of the baseline scenario. Cumulative default probability is the probability that a firm defaults in any one of the eight periods. Total capital and total output are the sum of the capital k borrowed and the sum of output y produced by the firm. Final size is the output y produced in the last period of the firm. Firm value is initial expected value of the firm in period one. All monetary values are in units of 1 million RMB.

I compare firm welfare in the model with long-term contracts with fully flexible schedules with the efficient benchmark (with no agency friction and full information, as in column (1) of Table 7), and I find that long-term contracts recover 31.46 percent of the total welfare loss from both sources of inefficiency (i.e., agency friction and incomplete information). And the average compensating variation across firms is around $92,106.

7 Conclusion

This paper utilizes data from a Chinese bank where long-term contracting is hard to access and documents several stylized facts on this novel dataset. I find that observed contracts are short term, with banks and firms negotiating contract terms on a annual basis. I find evidence of learning by examining the dynamics of contract terms, the contractual response to measures of firm performance, and how these responses change over time. This paper provides a framework to analyze the welfare impacts of contract structures. I develop a dynamic model of lending markets to analyse a bank's choice of contract terms and a firm's choice of whether to repay or default. I estimate the model with short-term contracts using panel data, where identification comes mainly from variations of contract terms over time and their correlation with firm performance and funding costs.

In the model there are two sources of inefficiency: agency friction (i.e., repayment behaviour cannot be enforced by contracts) and incomplete information (i.e., the firm's quality is unknown in the beginning). Estimates show that agency friction is the more important source of inefficiency in this market. By conducting counterfactual analyses of enabling banks to commit to long-term contracts, I find that long-term contracting can effectively alleviate agency friction through its use of both intertemporal and intratemporal structures. The intertemporal structure front-loads prices, which allows contract terms to become increasingly favourable to firms over time and, thus, disincentivizes firm defaults. The intratemporal structure is like insurance as it protects firms against the risks of negative productivity shocks. I find that the majority of welfare improvements come from the intertemporal structure of long-term contracts and a long-term contract with only a relationship-based schedule is almost as effective as the more-complex schedule that is contingent on the firm's performance history.

The main contribution of this paper is two fold. First, this paper provides an empirical framework, for the corporate lending market, that incorporates both observed and unobserved firm heterogeneity. Compared to previous models in macro-finance literature, my model can accommodate assumptions on different contract structures and utilize collateral information, firm performance and loan outcome data to form a more accurate estimate on the level of the agency frictions in place. Second, this paper quantifies the value of long-term contracting for small- and medium-sized young firms in China. There has long been discussion about long-term financing support for small- and medium-sized firms in China and this paper can provide a measure of potential benefits that long-term financing arrangements can bring about.

The model can also be used for various policy experiments. For example, we can calculate how a reduction in a bank's funding costs translates into changes in interest rates, loan sizes, and collateral coverage, and the associated default rates. As another example, we can simulate how a loan guarantee program (like the Small Business Administration loan in the U.S.) changes interest rates, loan sizes, and the associated default rates. We can further compare the effectiveness of different policy interventions in terms of default prevention and the enhancement of capital flows.

The model can be extended in the following aspects: (1) Learning by doing. We can allow firms to "learn" through past production; that is, more capital taken in the past can lead to higher productivity on average. Then part of the observed growth in capital could be attributed to learning-by-doing (In fact, the observed capital growth is larger than the predicted capital growth from the current model, and the unexplained part can be accounted for by a learning-by-doing process). This can still be identified because we do see the firm's performance history, which pins down the belief-updating process. So capital growth that cannot be explained by the belief-updating process will identify the learning-by-doing process. This is similar to the identification argument in Pastorino (2019). (2) Aggregate productivity shocks: In the current model firm's productivity shocks are i.i.d. However it is possible that macro-economic conditions can cause a firm's productivity realizations to be correlated (say firms within the same industry have correlations in their productivity). It is possible to allow correlation structures among firms with similar observed characteristics.

The model is better suited to studying young firms, since I abstract away from asymmetric information, which could be a more severe issue in established firms. Another limitation is that the model does not take into account the *ex-ante* model hazard problem where, at the beginning of each period, firms choose the "effort" levels that are unobserved by banks but that could affect their payoffs. In this case, the insurance-like structure of the long-term contract with the relationship- and performance-based schedule might be problematic but the relationship-based schedule can still use an intertemporal pricing structure as an incentive devise.

References

D. A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: an empirical examination. *International Economic Review*, 44(3):1007–1040, 2003.

R. Albuquerque and H. A. Hopenhayn. Optimal lending contracts and firm dynamics. *The Review of Economic Studies*, 71(2):285–315, 2004.

C. Arellano, Y. Bai, and J. Zhang. Firm dynamics and financial development. *Journal of Monetary Economics*, 59(6):533–549, 2012.

J. P. Atal. Lock-in in dynamic health insurance contracts: Evidence from Chile. PIER Working Paper No. 5694, 2019.

P. Behr, A. Entzian, and A. Güttler. How do lending relationships affect access to credit and loan conditions in microlending? *Journal of Banking & Finance*, 35(8):2169–2178, 2011.

A. N. Berger and G. F. Udell. Relationship lending and lines of credit in small firm finance. *Journal of Business*, pages 351–381, 1995.

M. Berlin and L. J. Mester. On the profitability and cost of relationship lending. *Journal of Banking & Finance*, 22(6-8):873–897, 1998.

M. Berlin and L. J. Mester. Deposits and relationship lending. *The Review of Financial Studies*, 12(3):579–607, 1999.

B. Biais, T. Mariotti, J.-C. Rochet, and S. Villeneuve. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78(1):73–118, 2010.

A. W. A. Boot. Relationship Banking : What Do We Know? 25:7–25, 2000.

Y. Boualam. Credit markets and relationship capital. Working paper, University of North Carolina, 2018.

F. J. Buera, J. P. Kaboski, and Y. Shin. Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002, 2011.

F. J. Buera, J. P. Kaboski, and Y. Shin. Entrepreneurship and financial frictions: A macrodevelopment perspective. *Annual Review of Economics*, 7(1):409–436, 2015.

A. Chakraborty and C. X. Hu. Lending relationships in line-of-credit and nonline-of-credit loans: Evidence from collateral use in small business. *Journal of Financial Intermediation*, 15(1):86–107, 2006.

Y.-S. Chan and G. Kanatas. Asymmetric valuations and the role of collateral in loan agreements. *Journal of Money, Credit and Banking*, 17(1):84–95, 1985.

C. Chang, G. Liao, X. Yu, and Z. Ni. Information from relationship lending: Evidence from loan defaults in China. *Journal of Money, Credit and Banking*, 46(6):1225–1257, 2014.

P.-A. Chiappori and B. Salanie. Testing for asymmetric information in insurance markets. *Journal of Political Economy*, 108(1):56–78, 2000.

G. L. Clementi and H. A. Hopenhayn. A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics*, 121(1):229–265, 2006.

J. H. Cochrane. Time-consistent health insurance. *Journal of Political Economy*, 103(3):445–473, 1995.

T. F. Cooley and V. Quadrini. Financial markets and firm dynamics. *American Economic Review*, 91(5):1286–1310, 2001.

A. C. Cooper, C. Y. Woo, and W. C. Dunkelberg. Entrepreneurs' perceived chances for success. *Journal of Business Venturing*, 3(2):97–108, 1988.

B. G. S. Crawford, N. Pavanini, and F. Schivardi. Asymmetric Information and Imperfect Competition in Lending Markets. 108(7):1659–1701, 2018.

G. S. Crawford and M. Shum. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173, 2005.

H. Degryse and P. Van Cayseele. Relationship lending within a bank-based system: Evidence from European small business data. *Journal of FinancialIntermediation*, 9(1):90–109, 2000.

P. M. DeMarzo and M. J. Fishman. Optimal long-term financial contracting. *The Review of Financial Studies*, 20(6):2079–2128, 2007.

P. M. DeMarzo and Y. Sannikov. Optimal security design and dynamic capital structure in a continuous-time agency model. *The Journal of Finance*, 61(6):2681–2724, 2006.

A. Demirgüç-Kunt and V. Maksimovic. Institutions, financial markets, and firm debt maturity. *Journal of Financialeconomics*, 54(3):295–336, 1999.

Y. Dong, Q. Gou, and H. Qiu. Big tech credit score and default risk——evidence from loan-level data of a representative microfinance company in China. *China Economic Review*, page 102010, 2023.

R. Elsas and J. P. Krahnen. Is relationship lending special? Evidence from credit-file data in germany. *Journal of Banking & Finance*, 22(10-11):1283–1316, 1998.

H. Gao, H. Ru, R. M. Townsend, and X. Yang. Rise of bank competition: Evidence from banking deregulation in china. *NBER working paper*, (w25795), 2019a.

H. Gao, H. Ru, and X. Yang. What do a billion observations say about distance and relationship lending? *Available at SSRN 3195616*, 2019b.

M. Gertler. Financial capacity and output fluctuations in an economy with multi-period financial relationships. *The Review of Economic Studies*, 59(3):455–472, 1992.

D. He and H. Wang. Dual-track interest rates and the conduct of monetary policy in China. *China Economic Review*, 23(4):928–947, 2012.

I. Hendel and A. Lizzeri. The role of commitment in dynamic contracts: Evidence from life insurance. *The Quarterly Journal of Economics*, 118(1):299–328, 2003.

K. P. Huynh and R. J. Petrunia. Age effects, leverage and firm growth. *Journal of Economic Dynamics and Control*, 34(5):1003–1013, 2010.

V. Ioannidou, N. Pavanini, and Y. Peng. Collateral and asymmetric information in lending markets. *Journal of Financial Economics*, 144(1):93–121, 2022.

K.-Y. Kang and I. Jang. Collateral risk, information sensitivity, and default frictions on collateralized debt. *Information Sensitivity, and Default Frictions on Collateralized Debt (January 24, 2019)*, 2019.

K. Kobayashi, T. Nakajima, and S. Takahashi. Debt overhang and lack of lender's commitment. *Journal of Money, Credit and Banking*, 2022.

N. Kovrijnykh. Debt contracts with partial commitment. *American Economic Review*, 103(7): 2848–74, 2013.

M. Manove, A. J. Padilla, and M. Pagano. Collateral versus project screening: A model of lazy banks. *RAND Journal of Economics*, pages 726–744, 2001.

M. McMahon, M. A. Schipke, and X. Li. *China's monetary policy communication: Frameworks,* *impact, and recommendations*. International Monetary Fund, 2018.

V. Midrigan and D. Y. Xu. Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–58, 2014.

K. Okazaki. Banking system reform in China: The challenges of moving toward a market-oriented economy. Technical report, Rand Corp Arlington Va National Security Research Div, 2007.

E. Pastorino. Careers in firms: the role of learning and human capital. Technical report, Technical report, mimeo, 2019.

M. A. Petersen and R. G. Rajan. The benefits of lending relationships: Evidence from small business data. *The Journal of Finance*, 49(1):3–37, 1994.

M. A. Petersen and R. G. Rajan. The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443, 1995.

N. Porter, E. Takáts, and T. Feyzioglu. Interest rate liberalization in China. *IMF Working Papers*, pages 1–28, 2009.

V. Quadrini. Investment and liquidation in renegotiation-proof contracts with moral hazard. *Journal of Monetary Economics*, 51(4):713–751, 2004.

J. Thomas and T. Worrall. Foreign direct investment and the risk of expropriation. *The Review of Economic Studies*, 61(1):81–108, 1994.

Y. Xin. Asymmetric information, reputation, and welfare in online credit markets. *Available at SSRN 3580468*, 2020.

B. Xu, A. van Rixtel, and M. Van Leuvensteijn. Measuring bank competition under binding interest rate regulation: the case of China. *Applied Economics*, 48(49):4699–4718, 2016.

X. Zhang, R. Ouyang, D. Liu, and L. Xu. Determinants of corporate default risk in China: The role of financial constraints. *Economic Modelling*, 92:87–98, 2020.

Appendix A Test of Asymmetric Information

A reduced-form test of adverse selection, inspired by Chiappori and Salanie (2000), looks at the correlation between the unobservables driving loan demand and the default decision. I first look at the intensive margin of the loan demand—the loan size—and set up the following seemingly unrelated regression, following Crawford et al. (2018):

$$LoanSize_{it} = X_{it}\beta + \epsilon_{it}$$
$$Default_{it} = X_{it}\beta + \nu_{it}$$

where X_{it} includes the relationship year, the firm credit levels, industry, ownership types, and age, and the branch and year fixed effects. From the estimation results, the correlation coefficient ρ between ϵ and ν is estimated to be -0.0167 and the Breusch-Pagan test of independence yields p-value = 0.0165. Thus, the loan size and default are not positively correlated conditional on the observables.

On the extensive margin, I look at the loan utilization and test for the correlation between the unobservables driving the frequency of borrowing and default. Specifically, for firms with the same duration of relationship, I check whether the total number of loans is correlated with the final default, controlling for the observables by the following seemingly unrelated regression:

$$\#ofLoan_i = X_i\beta + \epsilon_i$$
$$FinalDefault_i = X_i\beta + \nu_i$$

where X_{it} includes the fixed effects of the starting year of the lending relationship, the firm credit levels, industry, ownership type, initial age, and location, and the branch fixed effects.

Table 11 shows the correlation coefficient ρ between ϵ and ν for each group of firms with the same relationship duration. None of these factors shows a statistically significant positive correlation coefficient.

Table 11: Positive Correlation Tests

Duration (yrs)	3	4	5	6	7
Num. Obs.	891	779	601	335	197
ρ	-0.0233	-0.0452	0.0265	0.0762	-0.0524
p-values	0.487	0.2075	0.197	0.1632	0.4624

In summary, firms with higher loan demand, either in the form of larger loan sizes or more-frequent borrowing, are not more likely to default, conditional on the observables. This is not consistent with the presence of asymmetric information.

Appendix B Source of Agency Friction: An Illustration

For now I take the contract terms and the firm's future value as given, in order to focus on the firm's default decision in a given period and how this differs from the social planner's solution. In the full dynamic model, I will characterize the bank's problem and endogenize the contract terms. The firm's future value will also be endogenously determined through a dynamic learning process.

Suppose in period zero, a bank lends one dollar to a firm and the repayment in period one is $R > 1$. This loan is collateralized and the value of the collateral is R for the firm. Crucially, the bank values the same collateral at less than R. From the bank's perspective, the collateral's value is only ηR, with $\eta < 1$. This is a reasonable assumption given that collateral is often in the form of fixed assets like specialized equipment, which is more valuable for the firm that specializes in the type of production that requires that equipment than to non-specialists. Moreover, the process of forfeiting the collateral and transferring it from the firm to the bank is costly and such costs are born by the bank.

In period one, the firm is faced with a liquidity shock. This is an *ex-post* shock that realizes only after period zero. Suppose the realization is L, which means that there is an extra cost, L, if the firm repays the loan. This can be viewed as the cost of repaying the loan when the firm is cash constrained due to unexpected demand or cost shocks. In that case, the firm might resort to borrowing from a high-cost source to repay the bank loan. The benefit of repaying the loan is that the firm can continue borrowing and realize future value W.

If the firm chooses to default, in addition to losing the value of its collateral, it will also face depreciation of its future value. According to Zhang et al. (2020), most defaulted firms in China survive and continue to trade in the market. However, the default record may result in limited access to the credit market in the future.

Suppose the future value in this case is δW with $\delta < 1$. The firm's problem is to choose whether to default by comparing the payoffs. Let w^p and w^d denote the firm's payoffs from either repaying or defaulting on the loan, respectively. Based on the previous discussion,[41]

$$w^p = -R - L + W$$
$$w^d = -R + \delta W.$$

Therefore the firm defaults when $w^p < w^d \Leftrightarrow W < W^f \equiv \frac{L}{1-\delta}$.

For a social planner, the default decision is based on comparing the total surplus, which is the sum of the firm and the bank's payoffs. The bank's payoff when the firm repays the loan is simply R.[42] And it is ηR when the firm defaults. Thus, the total surplus in the case of either repayment (the first equation) or default (the second) is

$$v^p = -L + W$$
$$v^d = -(1-\eta)R + \delta W.$$

Therefore a socially optimal default occurs when $v^p < v^d \Leftrightarrow W < W^* \equiv \frac{L-(1-\eta)R}{1-\delta}$

As long as $\eta < 1$ and $L > (1-\eta)R$, the social planner's default threshold, W^*, is lower than the firm's threshold, W^f, and there exists a *wedge* between private default and the socially optimal default. When the firm's future value W is in this wedge region, $\left(W^*, W^f\right)$, its decision to default is not socially optimal. In other words, there is excessive default in this region. Intuitively, this is because the firm does not fully internalize the social cost of default, which includes the costly transfer of the collateral $(1-\eta)R$. For a smaller η, this transfer cost is higher and the wedge between the private and social default decision is larger. In addition, δ also governs the size of the wedge. When δ is small, the cost of default to a firm's future value is larger and the wedge is smaller.

Socially optimal default occurs for firms with very low future value, that is, where W is lower than W^*. In this case, the low future value cannot justify the high cost of repaying the loan, due to the large liquidity shock, so default is socially optimal. In the case of $W \geq W^f$, the high future value induces the firm to repay the loan, which aligns with the social planner's solution.

What determines the firm's future value? In the broader model, the firm's future value depends on the firm's (belief of) its profitability prospect as well as the *contractual environment* it is in, that is, whether the contractual relationship it has with the bank is long or short term. In Section 3, I focus on the case where the contractual environment is fixed to be short term, which fits the

[41] For simplicity, the discount factor is assumed to be one.
[42] I assume a competitive market and, thus, zero future value for the bank.

observed market features in my empirical setting. The structural model allows me to estimate the important parameters that govern the extent of the agency friction and the other frictions in the environment. I use the estimates in a counterfactual exercise, where I fix the contractual environment to be long term. I show that the new environment improves the firm's future value and, thus, reduces default and better restores the social planner's solution.

Appendix C Numerical Results from Two-Period Contingent Contracts

Case 1: With complete type information. To begin with, I abstract from the learning dynamics and consider the case with complete information about the firm type. In other words, my analysis first focuses on the case with $p_1 = 0$ (low type) and $p_1 = 1$ (high type). For each type (high/low), each characteristic group s and each c_1, I calculate the first-period lending terms (interest rates, loan size, and collateral coverage) and the expected second-period lending terms.[43] Table 12 shows the results for the lending terms by each characteristic group, compared with the short-term model, where the first-period funding cost is fixed at $c_1 = 1.03$. Interest rates are shown in terms of the annualized percentage rate; that is, they equal $(r-1)\times 100\%$.

The intertemporal pattern of the interest rate is similar cross all characteristics groups and productivity types: In the first period, the interest rates are higher than the short-term rates, but they are lower than the short-term rates in the second period. Note that the short-term rates show a slight decrease from the first to the second period because of the funding costs. But this decrease is more dramatic for the model with the short-term contracts, which is almost a 50 percent decrease from period one to period two. Meanwhile, the loan size increases and the collateral coverage decreases from the first to the second period. And the period-two collateral coverage in the long-term contracts is always lower than in the short-term contracts. This pattern translates into *negative* expected bank profits for the second period.[44] For the high-type firm, the average second-period expected bank profit across s group is -0.0988; for the low-type firm, it is -0.0874.

To compare firm outcomes, I find the associated default probabilities over the two periods, from which I calculate the cumulative default probability, defined as the probability that a firm defaults in either of the first two periods.[45] Lastly, I obtain the expected firm value in the beginning

[43]The expectation is taken over c_2, conditional on c_1.

[44]Note that in the complete information case $\bar{\pi}$ is no longer contingent on the realization of a_1.

[45]That is to say, one minus the cumulative default probability is the probability of surviving for the first two periods.

of the first period. I average the cumulative default probability and firm value across c_1 and s, using the empirical distribution obtained from the estimation, and I compare those statistics with the short-term model. The results are shown on Table (13). The cumulative default probabilities decrease for both firm types but their magnitudes are small. For the high type, the cumulative default probability with the two-period long-term contract is 10.665 percent, which is a 0.244 percent decrease from the short-term contract model, and the expected firm value at the beginning increases by 0.024 percent. For the low type, the cumulative default probability with the two-period long-term contract is 11.402 percent, which is a 0.018 percent decrease from the short-term contract model, and the expected firm value at the beginning increases by 0.002 percent. Overall, the high-type firm benefits more from long-term contracts than the low-type firm.

Table 12: Lending Terms in the Model with Two-Period Long-Term Contracts vs. Short-Term Contracts

Type	Period	Interest Rate (%)		Loan Size		Collateral Coverage	
		LT	ST	LT	ST	LT	ST
		Characteristic Group $s = 1$					
High	1	9.1419	7.258	2.7016	2.7019	0.86738	0.85783
	2	5.1543	7.1211	2.7115	2.7109	0.81941	0.85697
Low	1	9.2978	7.5127	2.3948	2.3992	0.87777	0.86663
	2	5.57	7.3753	2.4115	2.4072	0.83137	0.86577
		Characteristic Group $s = 2$					
High	1	12.8213	9.8274	4.2589	4.2516	1	0.98152
	2	6.4374	9.6849	4.2601	4.2659	0.92143	0.98053
Low	1	13.0551	10.169	3.8861	3.8856	0.99913	0.97816
	2	6.8497	10.0259	3.8997	3.8987	0.91933	0.97717
		Characteristic Group $s = 3$					
High	1	6.7322	5.2464	8.8922	8.8923	0.74596	0.74719
	2	3.5596	5.1143	8.9213	8.9218	0.72034	0.74634
Low	1	6.7078	5.2677	8.3683	8.372	0.76982	0.76798
	2	3.6624	5.1355	8.4019	8.3997	0.74261	0.7671

[a] LT stands for the model with two-period long-term contracts; SL stands for the model with short-term contracts.
[b] The column labelled "Interest Rate" is transformed from the variable r via $(r-1)\times 100\%$.
[c] The results are calculated under the first-period funding cost $c1 = 1.03$.

Case 2: With incomplete type information. Now I examine the case where the initial prior belief is $p_1 \in (0,1)$. For each s, I fix p_1 at the estimated value of λ_s, which is the proportion of high-type firms in the population, conditional on s. And c_1 is fixed at 1.03, as in the previous

Table 13: Firm Outcomes in the Model with Two-Period Long-Term Contracts vs. Short-Term Contracts

	High Type			Low Type		
	LT	ST	$\Delta \times 100\%$	LT	ST	$\Delta \times 100\%$
Cum. Def. Prob. (%)	10.6647	10.6791	-0.24454	11.4016	11.4028	-0.018184
Firm Value	86.4002	86.3805	0.023763	84.2922	84.2905	0.0020495

[a] LT stands for the model with two-period long-term contracts; SL stands for the model with short-term contracts; $\Delta \times 100\%$ is $(LT - ST)/ST \times 100\%$.
[b] Cum. Def. Prob. stands for the cumulative default probability, calculated as the probability of default in either period one or period two, represented in percentage terms.
[c] The items in this table are averaged across c_1 and s, using empirical distributions.

case. I find that both the first-period and the expected second-period lending terms follow a high and low realization of a_1. (As in the previous case, I take the expectation over c_2, using the transition matrix of c). The results are shown on Table 14. First, the intertemporal pricing pattern is still a front-loading one, where prices in the first period are much higher than in the second period. This echoes the findings in the complete-information case. Second, by comparing the second-period terms in the good state, where the first-period productive realization is high ($a_1 = \bar{a}$)), and in the bleak state, where the first-period productive realization is low ($a_1 = \underline{a}$), we find that the interest rates and the collateral requirements are lower in the bleak state than in the good state. The loan size, though, is smaller in the bleak state than in the good state. This is because the firm's expected productivity turns down according to the updated belief, which translates into a smaller loan size. Still, the loan size for the bleak state under long-term contracts is slightly larger than that in short-term contracts, so, conditional on a bad realization of a_1, firms are strictly better off under long-term rather than short-term contracts. This speaks to the insurance role of the long-term contract.

Another angle to look at this is through $\bar{\pi}(p_2)$, which is the bank's expected profits in the second period when the updated belief is p_2. For each group s, I tabulate $\bar{\pi}(p_2)$ for $p_2 = p_2^H$ and $p_2 = p_2^L$ in the last two rows of Table 15. I also find the expected profit for the first-period *only*, which is shown in the first row of Table 15. The weighted sum of the three items in each column should be zero, where the weight for the second-period profits includes the discount factor β, the probability of each realization of a_1, and the probability of repaying the first loan conditional on a_1. The results show a clear pattern of front-loading, where banks expect to earn positive profits on the first-period loans and "give back" to firms in the second period in the form of low interest rates. Furthermore, banks give back more when firms are hit by a negative cost shock and, thus, hold a more pessimistic outlook for their future prospects. This echoes the insurance role of long-term contracts, as shown in Table 14, which benefits risk-averse firms.

Table 14: Lending Terms in the Model with Two-Period Long-Term Contracts vs. Short-Term Contracts with Incomplete Firm Type Information

Period	Contingency	Interest Rate (%)		Loan Size		Collateral Coverage	
		LT	ST	LT	ST	LT	ST
		Characteristic Group $s = 1$					
1		9.2046	7.4128	2.5057	2.5092	0.87225	0.86272
2	$a_1 = \bar{a}$	5.5453	7.268	2.5296	2.5264	0.82875	0.8616
	$a_1 = \underline{a}$	2.9986	7.3481	2.4479	2.4389	0.78225	0.865
		Characteristic Group $s = 2$					
1		12.9132	10.004	4.0586	4.0553	0.99983	0.9795
2	$a_1 = \bar{a}$	6.8118	9.8511	4.0785	4.0799	0.9233	0.97862
	$a_1 = \underline{a}$	4.6683	9.9744	3.9526	3.9524	0.88098	0.97758
		Characteristic Group $s = 3$					
1		12.9132	10.004	4.0586	4.0553	0.99983	0.9795
2	$a_1 = \bar{a}$	6.8118	9.8511	4.0785	4.0799	0.9233	0.97862
	$a_1 = \underline{a}$	4.6683	9.9744	3.9526	3.9524	0.88098	0.97758

[a] LT stands for the model with two-period long-term contracts; SL stands for the model with short-term contracts.
[b] The column labelled "Interest Rate" is transformed from the variable r via $(r-1) \times 100\%$.
[c] The results are calculated under first-period funding cost $c1 = 1.03$.
[d] The initial prior beliefs p_1 for $s = 1,2,3$ are set at 0.399, 0.486, and 0.471, respectively (from Table 4).

Table 15: Expected Bank Profits in Each Period

Period	Contingency	Charateristic Group s		
		$s=1$	$s=2$	$s=3$
1		0.046506	0.12135	0.13054
2	$a_1 = \bar{a}$	-0.043113	-0.12937	-0.11749
	$a_1 = \underline{a}$	-0.11142	-0.22231	-0.29787

[a] The items in this table show the expected bank profits for the given period only. In period one, bank profits are calculated under first-period funding cost $c1 = 1.03$ and initial prior beliefs $p_1 = 0.399, 0.486, 0.471$ for $s = 1,2,3$, respectively. In period two, they are calculated under updated beliefs following a high/low realization of a_1 and averaged across different funding costs c_2.

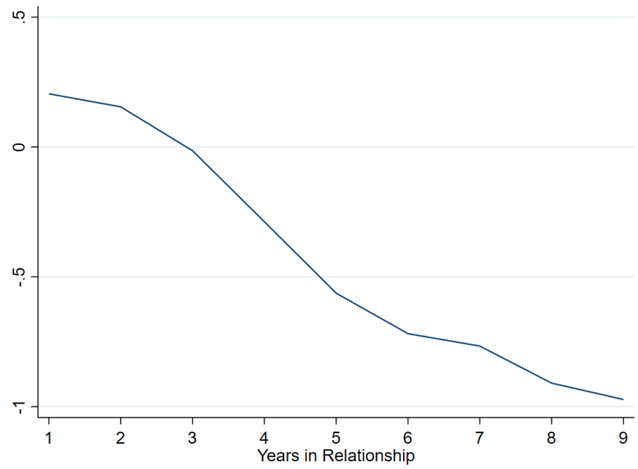
[b] Profits are in units of millions RMB (approximately US$150 thousand).

Appendix D Additional Tables

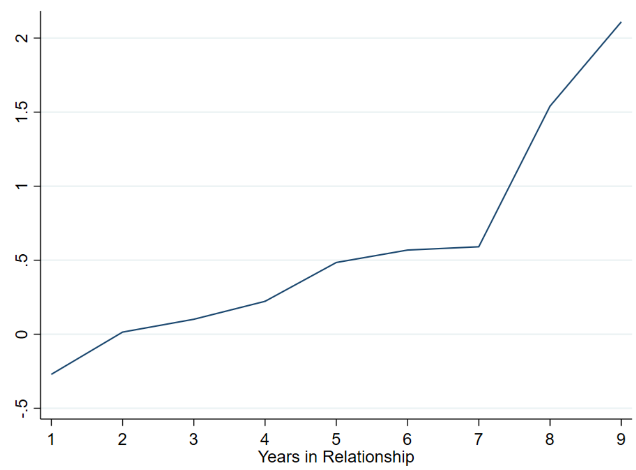
Table 16: Moment Restrictions

Meaning	Moment Restrictions	Dimension
Decisions on contract terms and default	$\mathbb{E}_{\{i\}}\left[O_i - \hat{O}_i \vert X_i\right]$	21×28
Firm performance	$\mathbb{E}_{\{i\}}\left[P_i - \hat{P}_i \vert X_i\right]$	21×7
Covariance between firm performance and decisions	$\mathbb{E}_{\{i\}}\left[P_i O_i - \hat{P}_i \hat{O}_i \vert X_i\right]$	21×28
Covariance between cost of funds and decisions	$\mathbb{E}_{\{i\}}\left[C_i O_i - \hat{C}_i \hat{O}_i \vert X_i\right]$	21×28

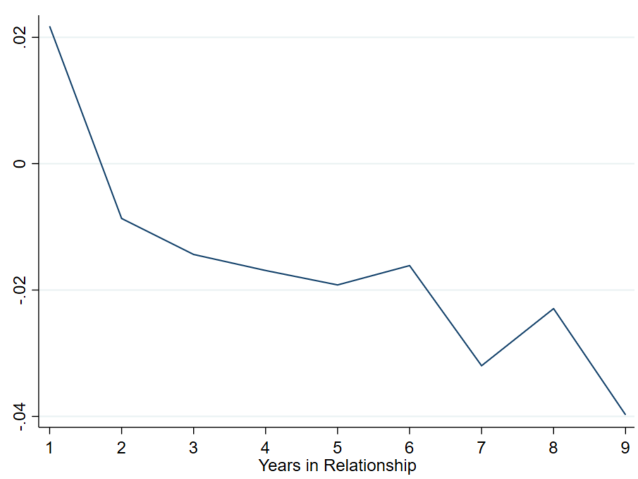
Appendix E Additional Graphs



(a) Residualized Lending Spread

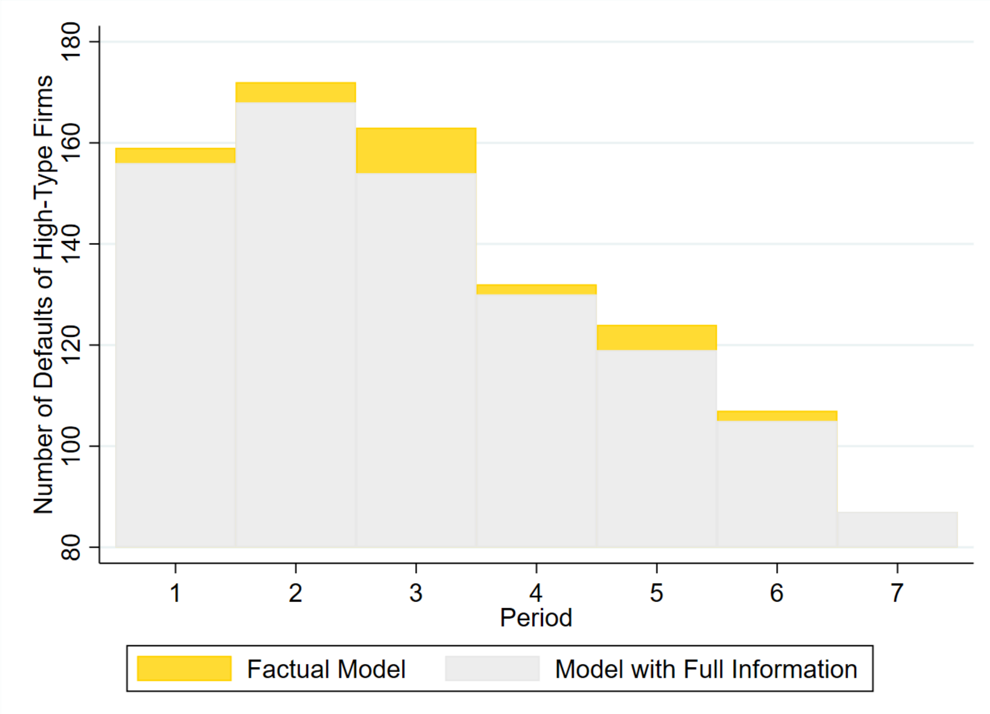


(b) Residualized Loan Size

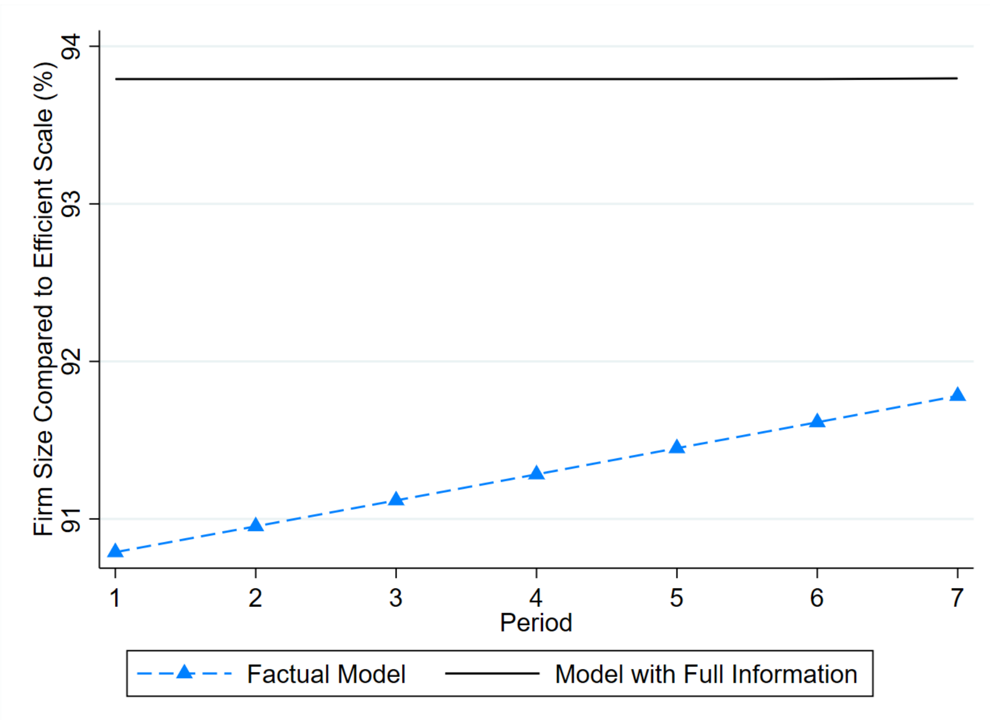


(c) Residualized Collateral Coverage

Figure 3: Time Profile of Residualized Contract Terms

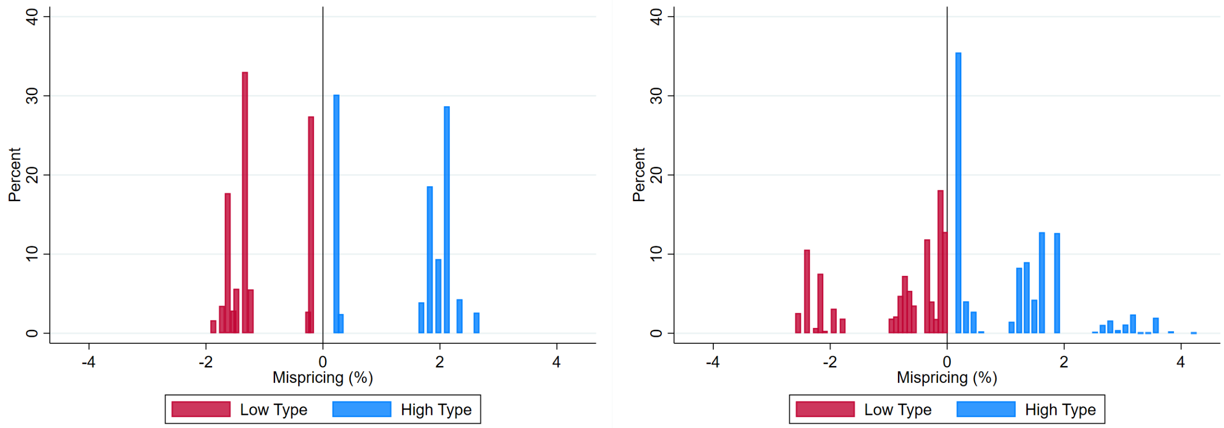


(a) Defaults over the Course of the Relationship



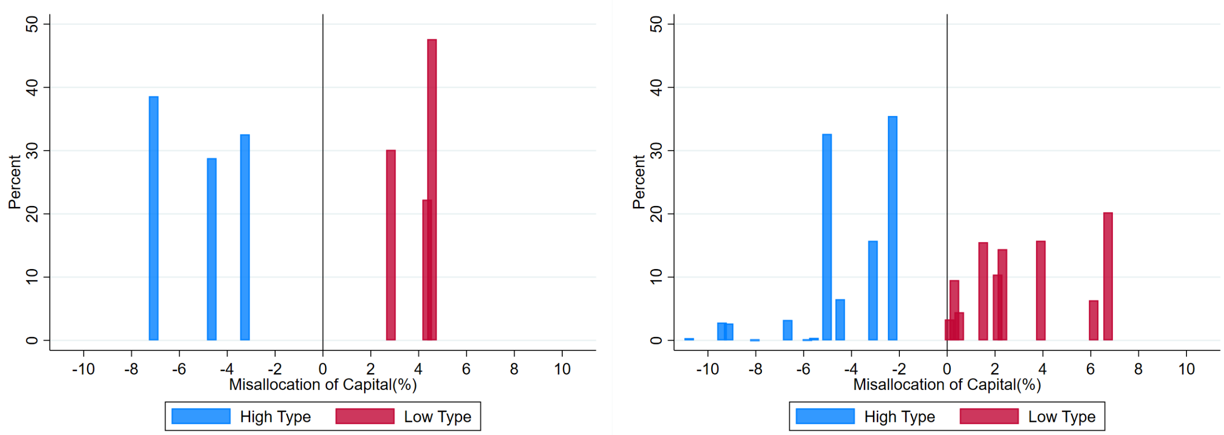
(b) Firm Size Compared to Efficient Size

Figure 4: Implications of Incomplete Information on Outcomes of High-Type Firms



(a) Mispricing in the First Period (b) Mispricing in the Last Period

Figure 5: Mispricing over the Course of the Relationship



(a) Misallocation of Capital in the First Period (b) Misallocation of Capital in the Last Period

Figure 6: Misallocation of Capital over the Course of the Relationship

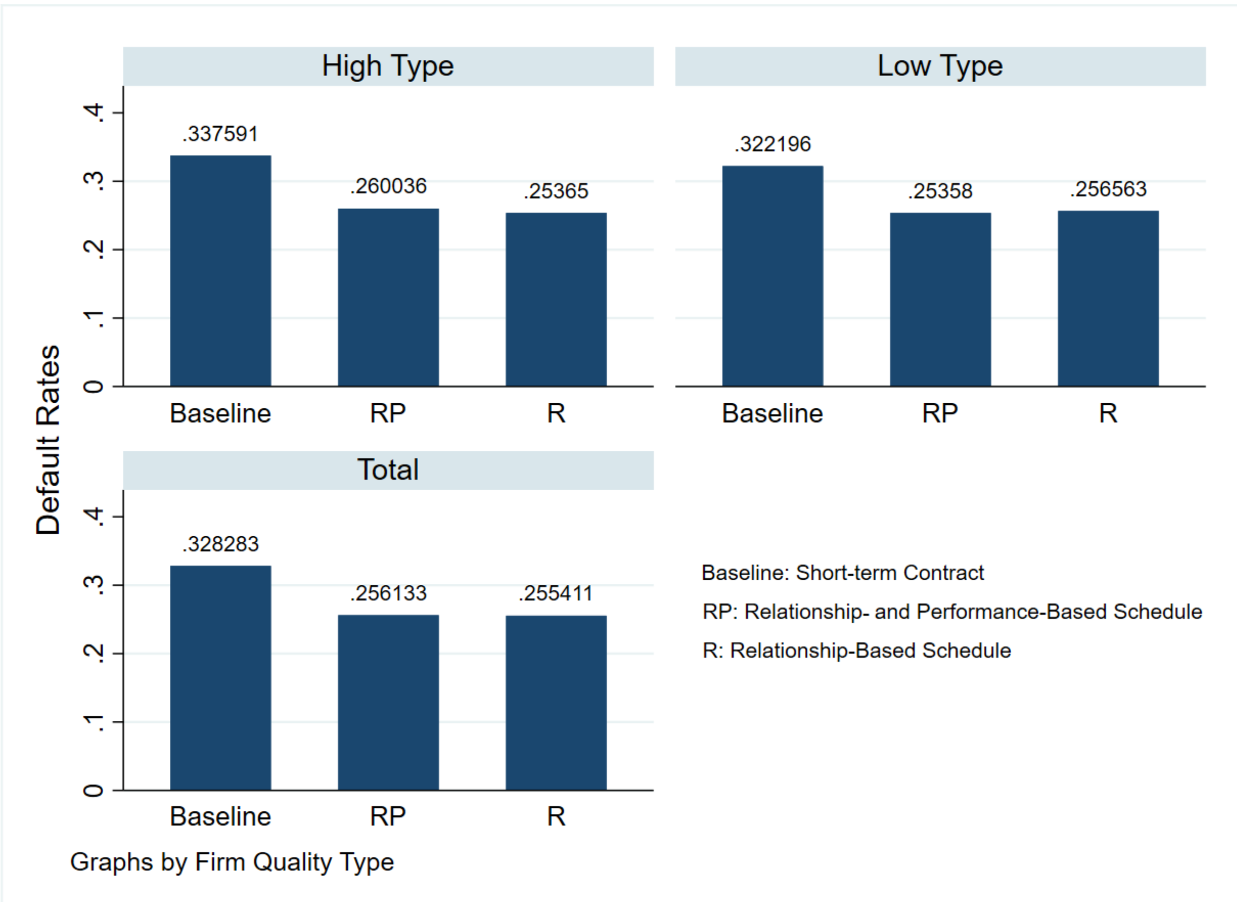


Figure 7: Comparison of Default Rates Across Different Contracts