Banking Relationships and Loan Pricing Disconnect

Francesco Beraldi*

Job Market Paper

October 30, 2024

Click here for the latest version

Abstract

How do long-term relationships between banks and firms shape loan pricing and capital allocation? Using administrative data from Mexico's credit registry, I provide stark evidence for an insurance view of relationship lending. When firms repeatedly borrow from the same bank, the pass-through of changes in their default risk to loan rates is nearly zero, and past risk assessments persistently influence credit terms. In contrast, switching to a new bank results in full risk pass-through, consistent with competitive market predictions. I rationalize this evidence in a structural model where banks compete for borrowers by offering optimal long-term contracts. Switching costs sustain commitment to banking relationships, enabling insurance. The estimated model replicates the observed pricing patterns and generates new predictions on when firms receive cheap funding and when they are tempted to switch, which I validate in the data. At the macro level, switching costs enhance capital allocation by strengthening relationships, recovering over 10 percent of welfare losses from financial frictions. However, when embedded in a New Keynesian framework, relationships dampen monetary and fiscal policy pass-through, as banks optimally absorb part of these policy shocks.

*Yale University. Email: francesco.beraldi@yale.edu

I am deeply indebted to my advisors Giuseppe Moscarini, Eduardo Dávila, Stefano Giglio and Zhen Huo. I received precious feedback from Alp Simsek, Joel P. Flynn, Paul Goldsmith-Pinkham, Michael Peters, and Martin Souchier, and from Edward Vytlacil, Kelly Shue, Fabrizio Zilibotti, John Geanakoplos, Christopher Clayton, Theis Ingerslev Jensen, Aleh Tsyvinski, David Argente, Kaushik Vasudevan, Keshav Dogra, Jake Weber, Marco Del Negro, Mary Amiti, Kinda Hachem, Miguel Faria-e-Castro, Juan M. Sánchez, and B. Ravikumar, and seminar participants at Yale University, AFA poster session, Banco de México, New York Fed, St. Louis Fed, and the Carey Finance Conference. I benefited from significant insights from Cedomir Malgieri, and from Alessandro Lavia, Chenping Yang, and Alvaro Cox. For their help in accessing the data, I am indebted to Santiago Bazdresch and the Econlab team: Claudia Velázquez Villegas, David Lara, Victor Morales, and Osvaldo Veloz.

This paper is part of the project "Bank-Firm Relationship and Lending Conditions in the Cross-Section". The data request for the project is sponsored by Santiago Bazdresch at Banco de México's Research Directorate. The EconLab collected and processed the data as part of its effort to promote evidence-based research and foster ties between Banco de México's research staff and the academic community. Inquiries regarding the terms under which the data can be accessed should be directed to: econlab@banxico.org.mx.

1 Introduction

Long-term relationships between banks and firms are a pervasive feature of credit markets, and the finance literature has long recognized their importance for screening, monitoring, and generally improving credit access for firms (Stiglitz and Weiss (1981), Diamond (1984), Petersen and Rajan (1995)). Despite this, the standard framework for assessing loan pricing is that of competitive credit markets, where long-term relationships play no role because of perfect mobility of firms across lenders. The competitive framework provides tight predictions for lending rates, which should move one-for-one with both the firm's default risk and the risk-free rate. In the presence of mobility frictions these predictions may not hold, and we still know relatively little about how banks actually incorporate firm default risk into loan rates, primarily because such risk is typically not measured. An important question then is: how do banking relationships shape loan pricing, and how does this ultimately affect the allocation of capital and the transmission of macroeconomic policy?

In this paper, I address these questions using rich administrative data from Mexico's credit registry, which contains bank assessments of firm default risk. I combine empirical evidence and a structural model to support an insurance view of relationships, whereby banks may not raise rates to firms facing adverse shocks, but charge firms higher rates when they are performing well. Therefore, I argue that the presence of relationships yields pricing and allocations that are different from those in competitive markets.

Empirically, I find that within bank-firm relationships the pass-through of changes in firm default risk—as assessed by the bank—to loan rates is nearly zero, but it is much larger and close to one when firms switch banks. Loan pricing also exhibits strong history-dependence, a typical feature of long-term insurance arrangements: the default risk at the onset of a relationship has a large influence on loan pricing even conditional on the current risk profile. My model formalizes relationships as optimal long-term contracts, through which banks provide insurance to firms, which firms value because bankruptcy costs make them act as risk averse. The model rationalizes both limited pass-through and history-dependence. It also generates new predictions, supported by the data, about when firms receive cheap loans and when they are tempted to switch to a new bank. In the model, the presence of switching costs enables partial commitment to relationships and thus risk-sharing. As a result, switching costs improve allocation and welfare in equilibrium, restoring over 10 percent of the welfare costs of financial frictions. However, they also weaken the pass-through of monetary policy and dampen the transmission of fiscal policy by nearly 20 percent relative to a model with competitive lending markets.

Empirics. I begin by documenting some novel empirical facts in the credit registry data. The data cover the near universe of corporate loans originated in Mexico since 2004. They offer detailed information on loan terms, along with self-reported data on sales and employment, which I complement with Orbis data. The typical firm is small and interacts with one or few banks over their observed lifespan, in sharp contrast with publicly available data on syndicated loans often used in banking research.[1] A unique feature of the credit registry data is that it contains an assessment of the default probability for each firm, which is provided by its lender under strict guidelines from the regulator. I provide evidence validating these assessments by showing that they strongly predict subsequently realized loan losses. Combining such default risk with the interbank rate in Mexico, I construct the loan rate implied by the competitive pricing rule for each firm. This allows me to test the predictions of competitive pricing, and study the empirical behavior of the loan pricing wedge—how actual loan rates differ from the competitive rule.

I obtain three main results supporting an insurance view of relationship lending. First and foremost, when a firm stays with the same bank, the *pass-through* from changes in default risk to the lending rates they get charged on new loans is close to zero: firms whose credit quality deteriorates do not experience substantial increases in their borrowing rate while improving firms do not enjoy rate reductions. Instead, when a firm switches to a new bank, this pass-through is close to one, as predicted by competitive pricing. This is a striking result that, as I will discuss below, is consistent with the insurance view of the relationship. An immediate concern with this result is whether this difference is driven by the relationship, as I argue, or by a selection of those firms that switch banks when taking new loans. I address this concern using two identification approaches, both of which confirm a sharp difference between stayers and switchers. One approach relies on a subset of firms that borrow from a new bank while simultaneously taking out a loan from their old bank, effectively circumventing the selection problem. These firms show the same differential pass-through across the two loans that I document for the broader population of firms. The other is an instrumental variable approach. I use two instruments for separation—the decision to switch banks—based on identifying shocks to the old bank that are plausibly unrelated to the firm yet induce the firm to switch, thus producing plausibly exogenous variation in whether a firm switches banks. For the first instrument, I collect data on bank branches at the municipality-level, and compute the change in the branch market share of a firm's lender in their municipality, with the intuition that firms are more likely to switch when their bank is closing branches, or if new banks are entering. For the second instrument, I estimate bank idiosyncratic credit supply shocks, using the methodology developed in Amiti and Weinstein (2018).

Second, I document strong *history-dependence* in credit conditions: firm default risk at the onset of a relationship strongly predicts its borrowing rate several years later, even

[1] The median firm in my sample has 5 employees and over 60 percent of firms borrow only from one bank over their observed lifespan. In contrast, LPC DealScan, SNC, and Compustat GMI, major syndicated loan data sources, primarily cover large-scale financing extended to sizable, often publicly listed firms, with each loan typically backed by a consortium of multiple lenders.

when controlling for the contemporaneous default risk. This result is crucial to distinguish among theories that could generate limited pass-through. Indeed, history-dependence is a typical feature of long-term implicit insurance contracts:[2] two identical firms may be paying a different loan spread due to differences in their initial conditions, which in turn influence the entire path of the relationship through long-term contracting. Instead, the relevance of past conditions is inconsistent not only with competitive pricing, but also with a large class of models with bank market power or loan evergreening to zombie firms. Moreover, I also find that the level of competition among banks is important for these results. When firms have limited outside banking options, as measured by the branch market share of their bank in the municipality where they are headquartered, then the pass-through of default risk is weaker and there is stronger history-dependence. This evidence is consistent with the mechanism in my model whereby larger switching costs, here captured by the scarcity of alternative banks in the municipality, enable relational pricing.

Third, I find strong evidence of pricing *reconnect* upon switching: the gap between loan spreads and the probability of default—which I refer to as the pricing wedge—largely closes upon switching. For instance, firms with a positive pricing wedge, meaning they pay higher rates than their fundamentals would suggest, tend to receive large rate reductions after switching. This pattern ties back to the insurance mechanism, as switchers tend to be firms whose default risk has improved since the start of their prior banking relationship, and thus find themselves disadvantaged by staying in their current match. Indeed, when a firm's default risk decreases, the insurance arrangement can leave them overcharged compared to their improved risk profile, thus making them more likely to benefit from discounts if they switch to a new lender.

Theory. To rationalize these empirical loan pricing patterns and to study their aggregate implications, I build a general equilibrium model with optimal contracts between banks and firms. On the real side, the model has standard firm dynamics à la Hopenhayn and a representative household. On the financial side, it can accommodate deviations from competitive pricing through two new elements compared to existing models with optimal contracts (as Albuquerque and Hopenhayn (2004) and Cooley, Marimon, and Quadrini (2004)). First, one-period debt is enforceable, meaning it must be repaid even if a firm switches to a new bank. This allows the model to speak to the notions of lending rate and default, typically absent in contracting models, making it possible to connect to my empirical findings. The limited commitment is no longer about the decision to repay, as in these models, but about the decision to stick to the agreed credit policy, including the lending rate, or to switch to a new lending partner. Second, banks and firms face costs when switching to a new relationship. When switching costs go to zero, the model reverts to a competitive credit market, with shocks fully passed through to the lending rate. This special case is akin to the frameworks

[2] See Beaudry and DiNardo (1991) for the case of labor markets.

with one-period debt and endogenous default in Cooley and Quadrini (2001), Khan, Senga, and Thomas (2014), and Ottonello and Winberry (2020). Conversely, positive switching costs create a commitment to the relationship, enabling risk-sharing: firms pay higher rates in good times, in exchange for cheap credit in bad times, when they are more constrained.

The demand for insurance arises from financial frictions, in the form of a bankruptcy cost, which make firms act as if they were risk-averse.[3] Due to these frictions, firms may be constrained and operate below their optimal size—as in the misallocation literature (Moll (2014), Khan, Senga, and Thomas (2014)). This makes the value of an extra dollar within the firm greater than that of a dollar worth of consumption. More importantly for the insurance mechanism, the value of a dollar is not constant, but is higher in some states than in others, responding to productivity shocks and the firm's ability to adjust its stock of debt and capital. Through long-term contracts, banks provide insurance by offering more favorable financing conditions to firms when they are more constrained, allowing them to free up resources to invest and hire workers when the marginal value of dollars is high, in return for higher rates in states when the firm is close to its optimal size. Effectively, deviations from competitive loan pricing arise to cross-subsidize across firms, channeling resources to more constrained firms. I analytically characterize the pass-through of default risk and show that limited pass-through for stayers emerges as a result of banks subsidizing firms when their default risk increases.

Quantification. I estimate the model to match some key moments of the equilibrium firm distribution and, crucially, to replicate the observed pass-through of default risk to loan rates within relationships, which disciplines the magnitude of the switching costs. The estimated model delivers the targeted limited pass-through within relationships while matching the large untargeted pass-through for bank switchers. Furthermore, it matches the other two documented facts, which are untargeted. As in the data, it exhibits pricing reconnect upon switching and a substantial degree of history-dependence, one of the quintessential features of long-term insurance contracts, with past default risk almost as important for loan pricing as the contemporaneous one.

The model also generates new predictions regarding when firms receive cheap credit and when they are tempted to switch to a new bank, for which I find strong empirical support. Firms receive cheap loans, compared to their default risk, when they are not distributing dividends and when their sales decline, which reflects the arrival of shocks that make them more constrained. Also, banks deliver cheap loans early in their relationship with the firm, and charge higher rates later, which is an expression of inter-temporal risk-sharing. This

[3] In the model, both banks and firms are owned by the representative household. In the steady-state, this implies both agents are effectively risk-neutral, except for financial frictions. When studying aggregate shocks, instead, this implies that firms and banks cash-flows are both priced according to the same stochastic discount factor, and insurance against aggregate shocks is only valuable if such shock tighten the financial constraints on firms.

arrangement is optimal because firms are typically less financially constrained as they age and retain earnings. Just like in the data, firms are tempted to switch when their default risk has declined since the onset of the bank-firm relationship, which makes them disadvantaged from the insurance arrangement and possibly ready to benefit from starting a new relationship.[4]

Using the estimated model, I quantify the welfare gains from banking relationships. In an economy with positive switching costs, the model recovers over 10 percent of the welfare losses caused by financial frictions, compared to a world of perfect mobility. These switching costs enhance welfare by sustaining long-term commitments, enabling banks to channel resources to the most constrained firms, thereby improving overall equilibrium allocation. These arrangements recover only a portion of the welfare losses since the estimated model is still characterized by limited commitment, as switching is costly but possible.

Finally, I cast the model in a New Keynesian setup to study the transmission of macroeconomic policy. Solving the response of the economy to an aggregate shock in a framework with heterogeneous agents and state-contingent contracts is challenging. Existing methods[5] mostly study the perfect foresight transition path in economies with incomplete markets (”MIT shocks”). I show that a simple and computationally tractable extension to the sequence space method, as implemented in Boppart, Krusell, and Mitman (2018), can accommodate state-contingent contracts. The method leverages the insight that aggregate shocks that occur with infinitesimal probability do not affect the steady state, but can still be contracted upon, meaning that we can solve exactly for the optimal contract, including lending rates and credit quantity, contingent on the realization of aggregate shocks. I use this solution method to study the impulse response to monetary policy—an innovation to the Taylor rule—and fiscal policy—a lump sum transfer to firms akin to the large corporate subsidy schemes implemented during Covid. I find that relationships dampen the transmission of both monetary and fiscal policy, as banks optimally absorb a portion of these shocks as part of the risk-sharing agreement. Quantitatively, the effect of relationships for monetary policy transmission is large for financial variables, more than halving the fraction of firms forced to raise costly equity after the shock, but smaller for real variables, with only a moderate dampening in the transmission of monetary policy to firm investment. Relationships strongly dampen the effects of fiscal policy, reducing its transmission to investment by nearly 20 percent.

Related Literature. My paper contributes to several strands of the literature. First of all, it contributes to, and builds upon, the literature that has documented the conse-

[4]Notice that in the model these firms are tempted to switch, meaning they are indifferent to staying or switching, but never actually prefer to switch, as the optimal contract is designed to prevent inefficient separations. Therefore, all switching in equilibrium occurs due to exogenous separation shocks.

[5]Such as Boppart, Krusell, and Mitman (2018) and Auclert, Bardóczy, Rognlie, and Straub (2021).

quences of relationships between banks and firms, from an individual firm perspective but also from the aggregate point of view. In particular, an empirical literature has documented that firms with close relationships, typically measured using the frequency of past interactions, receive more abundant credit and at better terms during recessionary periods (Bolton, Freixas, Gambacorta, and Mistrulli (2016), Beck, Degryse, De Haas, and Van Horen (2018), Karolyi (2018), Banerjee, Gambacorta, and Sette (2021)) or during episodes of firm distress (Giometti, 2022). Dougal, Engelberg, Parsons, and Van Wesep (2015) and Demiroglu, James, and Velioglu (2022) have found evidence of path dependence in loan spreads—the relevance of past spreads for current loan pricing. My main empirical contribution relative to this literature is to directly measure the pass-through of default risk to loan rates leveraging the risk assessments provided by banks, revealing a striking difference between bank stayers and switchers. My focus on bank switching is facilitated by the use of supervisory data covering many small and medium sized firms with single bank relationships, as opposed to the syndicated loans data often used in the literature, where a consortium of banks pool funds to provide large-scale financing. The strong evidence of history dependence that I document—the relevance of past default risk for pricing, which goes over and beyond that of past spreads—, is deeply related to the path dependence, but is typically considered more direct evidence of long-term contracts (Beaudry and DiNardo, 1991). While this literature is mostly empirical, I also build a structural model that rationalizes my findings through an insurance mechanism and delivers new predictions for loan pricing and bank switching that I validate in the data. In Section 6.5 I leverage my empirical and quantitative results to compare in detail the insurance mechanism with other theories that could plausibly explain some of my results or that have been discussed in the literature, such as behavioral anchoring and information frictions.

My empirical results also relate to a literature that has documented an imperfect pass-through of monetary policy to corporate loan rates (Berger and Udell (1992), Cao, Dubuis, and Liaudinskas (2023)), as well as mortgages (Scharfstein and Sunderam, 2016) and deposits (Drechsler, Savov, and Schnabl, 2017), typically attributed to bank market power. I complement this literature by focusing on the pass-through of default risk, and by providing new evidence consistent with an insurance mechanism, most notably, the stark difference between stayers and switchers, and the strong history-dependence. On the theory side, a small literature studies monetary policy with banking relationships, as in Hachem (2011) and Rocheteau, Wright, and Zhang (2018), which focus on information asymmetries and limited pass-through due to bank market power, and Bethune, Rocheteau, Wong, and Zhang (2022), that studies an environment with money where banking relationships reduce the need for precautionary cash holdings by firms. I contribute to this literature by bringing banking relationships to a standard New Keynesian setup and by studying how the insurance mechanism affects the pass-through of monetary and fiscal policy.

The risk-sharing mechanism in my model is related to an older strand of the banking literature, surveyed by Cetorelli et al. (2001) and Elyasiani and Goldberg (2004), which used informal discussions or stylized one-period models to explore how banks can provide financial flexibility to firms, as in Fried and Howitt (1980), Berger and Udell (1992), and Berlin and Mester (1999). I contribute to this literature by incorporating the insurance insight into a model with long-term relationships, firm dynamics, and default, making it possible to derive testable predictions for loan pricing throughout relationships and in response to shocks, and with general equilibrium, enabling the quantification of the allocation and welfare consequences of relational pricing.

My results on intertemporal risk-sharing, whereby banks deliver cheap loans early in the relationship and higher rates later, resonate with the hold-up mechanism in Petersen and Rajan (1995), which can be beneficial by expanding the set of firms banks are willing to lend to, and for which Schäfer (2019) has recently provided empirical support. Similarly, Ioannidou and Ongena (2010) document that firms receive loan rate discounts at the onset of a relationship and high rates later, which contrasts with the opposite result found in Berger and Udell (1995). I contribute to this literature in two ways. First, empirically, I find that the pattern of increasing rates along the tenure profile holds both unconditionally and conditionally on the firm's risk assessment, attenuating possible selection concerns in previous studies. Second, theoretically, I show that these patterns are also consistent with an optimal risk-sharing contract, as opposed to the hold-up mechanism discussed in these papers.

This paper is also related to the literature on evergreening and zombie lending that emerged from the seminal work of Hoshi (2006) and Caballero, Hoshi, and Kashyap (2008) in the Japanese setting, showing that banks often extend loans to insolvent firms at subsidized rates to prevent their bankruptcy (Artavanis, Lee, Panageas, and Tsoutsoura (2022), Acharya, Crosignani, Eisert, and Steffen (2022), Hamano, Schnattinger, Shintani, Uesugi, and Zanetti (2024)). I contribute to the empirical strand of this literature by documenting a widespread limited pass-through of firm default risk to lending rates within relationships, which is not confined to nearly-defaulted firms, and by uncovering strong evidence of history-dependence, consistent with an insurance view of relationships but not with loan evergreening. My model is related to Faria-e Castro, Paul, and Sánchez (2024), which develops a setup with bank lending and evergreening with heterogeneous firms à la Hopenhayn, and Martin, Mayordomo, and Vanasco (2023), which studies the interaction of evergreening and macroeconomic policy. These papers highlight a notion of evergreening as implicit restructuring, which I relate to in an extension of my model. I thus provide the first joint formalization of evergreening and insurance, contributing to recent efforts to clarify the definition of zombie lending in Álvarez, García-Posada, and Mayordomo (2023). Leveraging on this extension, I confirm that an evergreening mechanism delivers limited pass-through from default risk, but little history-dependence, suggesting it is insufficient to fully explain the data.

My model lies at the intersection of two literature strands studying firm dynamics, financial constraints, and their macroeconomic implications. I provide a schematic overview of these two strands, and the novelties of my model, in Appendix B.5, summarized in Table 37. The first strand encompasses a growing class of heterogeneous firm models with endogenous default, as in Cooley and Quadrini (2001), Khan, Senga, and Thomas (2014), Gomes, Jermann, and Schmid (2016), Arellano, Bai, and Kehoe (2019), Ottonello and Winberry (2020), and Guntin (2023). In this literature, markets are incomplete: firms can only borrow through non-contingent debt, typically with one-period maturity, and financial frictions arise due to various forms of bankruptcy costs. Debt is priced by competitive lenders offering zero-profit schedules, following a framework originally developed in the sovereign debt literature by Eaton and Gersovitz (1981). Perfect competition and firm mobility imply that lending rates move one-for-one with firm default risk and the risk-free rate. I nest this class of models as a special case when switching costs approach zero and extend them by studying the role of long-term relationships in alleviating financial frictions and determining loan pricing.

The second strand of literature on firm financing studies optimal lending contracts under limited commitment, as in Cooley, Marimon, and Quadrini (2004), Albuquerque and Hopenhayn (2004), Rampini and Viswanathan (2010), Kovrijnykh (2013), and Ai, Bhandari, Chen, and Ying (2019). Here, the contracting space is not exogenously restricted to non-contingent debt, but financial frictions arise because no financial contracts are enforceable, as the entrepreneur can renege on the outstanding debt, abscond with the cash flow or the capital, and possibly borrow again from a new bank. I extend these models by introducing enforceable debt, meaning that firms have to repay even if they switch to a new bank. This allows to have well-defined notions of lending rates and default, typically absent in contracting models. The limited commitment in my model pertains instead to the promises on future credit terms, which both banks and firms can renege upon by starting a new relationship. Incorporating an enforceable asset in state-contingent contracts has some parallels with recent work in Souchier (2024), which studies household savings and firm-worker contracts.

Layout. The paper starts by describing the credit registry data in Section 2. Section 3 presents the main empirical findings. Section 4 outlines the model. Section 5 analytically characterizes the optimal contract. Section 6 outlines the calibration strategy and validates the model predictions for loan pricing. Finally, Section 7 presents the macroeconomic implications of relationship lending.

2 Data and Institutional Background

To study the loan pricing decisions of lenders and their connection to firm fundamentals, I use administrative data from Mexico's Credit Registry, which covers the near universe of corporate loans originated by Mexican banks.

The registry was established in the credit reporting market reform of 1995 (Gil Hubert (2004)), and fully implemented in the early 2000s under the regulation of Banco de México and Comisión Nacional Bancaria y de Valores (CNBV), the Mexican bank supervisory authority. The data are reported at the monthly frequency, from December 2003 to December 2022, and are organized at the loan level. Each loan is linked to one firm and one bank[6].

In my sample, I include loans to non-financial corporations, including personal loans to entrepreneurs, and exclude those to public entities and financial firms. I also exclude loans to firms headquartered outside Mexico and those denominated in foreign currencies. These restrictions retain the majority of loans from the original data, resulting in a final sample of over 10 million loans by 145 lenders to 1.1 million firms.

The data contain rich information on each loan, such as loan rates, principal amount, collateral, maturity, and loan type (term loans or credit lines). Loans are observed both at origination and during subsequent periods, allowing for an evaluation of firm repayment performance. On the firm side, the credit registry contains information on firm sales and on the number of employees, self-reported by firms in the credit registry. I complement this data using Orbis Mexico, which provides a rich array of balance sheet data for a sample of large Mexican firms. Starting in 2016, banks must report to the regulator the default probability of each borrower, expressed in percentage points. These assessments, which I observe in the credit registry, are critical for my empirical analysis.

In Section 2.1, I provide the key summary statistics of borrowers and loans in my sample. Next, in Section 2.2, I describe in greater detail the institutional aspects of the risk assessment, and validate its ability to predict subsequent loan losses.

2.1 Summary Statistics

There is no single way to approach the data, instead, we can organize it in four different ways: at the bank-level, at the firm-level, at the loan-level, and finally at the bank-firm relationship level. I will now illustrate these approaches one by one.

Overview of banks. There are 145 lenders in the sample. The banking sector in Mexico is concentrated, with a few large banks dominating the market. Over 80 percent of banking assets are concentrated in the top seven banks, five of which are Mexican subsidiaries of American and European banks. The 145 lenders are divided into three groups: *Banco*

[6] A small fraction of loans, around 0.1% of the sample, are syndicated, with each loan linked to a single firm but multiple banks. I exclude this small fraction of loans from my analysis.

Multiple, privately-owned commercial banks that originate almost 90 percent of the loans in my sample; *Sociedad Financiera*, non-bank lenders typically specialized in niche lending areas, accounting for about 10 percent of the loan originations, and *Banco de Desarrollo*, government-owned banks that originate less than 1 percent of the loans in my sample. Following standard practice in the banking literature, if two banks merge, I assign them a unique identifier for the entire sample. Furthermore, I assign a unique identifier to all entities in the same banking group (there are several instances of *Sociedades Financieras* that are subsidiaries of *Banco Multiple*), which is useful in order to avoid misclassifying switching across lenders within the same group as starting a new relationship.

Overview of firms. There are 1.1 million firms that borrow at least once over the 2003-2022 period. In table 1, I show the descriptives of these firms, based on the last time they took out a new loan. The typical firm in the sample is small, with only 5 employees, but there is a large cross-sectional dispersion, with very large firms represented as well. The average default risk assessed by banks, a variable I describe in detail in the next Section, is 8 percent, with a fat right tail. Borrowing rates are high, reflective of both a high risk-free rate (the interbank rate averaged at over 6 percent during the 2004-2022 period), and the compensation for risk.

Variable	Mean	p10	p50	p90
Employees	23.88	1	5	30
Sales ($'000)	6,356.55	5.77	73.81	2,642.01
Firm Age (Years)	18.19	2	17	35
Default Probability (%)	8.89	0.68	2.16	16.40
Borrowing Rate (%)	17.46	10.18	15.13	26.50
Spread (%)	10.56	4.28	9.99	17.72

Table 1: Summary statistics of firm real variables in the credit registry. Statistics are computed at the firm-level with each firm considered at the time of its last borrowing. p10, p50 and p90 are the percentiles of the distribution of each variable.

Overview of loans. The typical loan size is small, with a median value of 10 thousand dollars and an average of 100 thousand dollars. The typical loan maturity is two years, and almost two-thirds of loans have no collateral. Slightly over half of the loans are term loans, and the rest are credit lines. Interestingly, most loans are classified as revocable, meaning that the bank can call back the loan at will—a fact that underscores the relational nature of credit in this environment.

Overview of bank-firm relationships. Long-term bank-firm relationships are pervasive. Each month, over 80 percent of firms borrow from a single bank, and throughout the entire observed period from 2003 to 2024, over 60 percent of firms borrow exclusively from one bank. When we exclude firms that have just entered the credit market, the average length of

Variable	Mean	p10	p50	p90
Loan size ($'000)	94.94	0.00	8.70	74.85
Loan Maturity (Years)	2.57	0.08	2.83	4.92
Collateral (%)	38.74	0	0	100
		Term Loan	Revocable	Fixed Rate
Fraction of Loans		0.48	0.91	0.37

Table 2: Summary statistics of loan characteristics. Statistics are computed at the firm-level with each firm considered at the time of its last borrowing. p10, p50 and p90 are the percentiles of the distribution of each variable.

a banking relationship is 4.5 years, a value which is likely underestimated since we only observe relationships starting in 2003 and ending in 2022. For context, this value is more than twice the average loan maturity. The switching rate is around 10 percent. More specifically, when a firm borrows at time t, and it borrows again at time $t > \tau$, 10 percent of the time this loan is originated by a bank with which the firm had no pre-existing relationship, 73 percent of the time by the same bank as the previous loan, and 17 percent of the time by a different bank from which it had borrowed in the past.

2.2 Risk Assessments

One of the unique advantages of data from the credit registry is that I observe the risk assessment that banks are required to report for each loan. I will refer to this variable interchangeably as risk assessments or probability of default (PD), but this measure is technically broad in scope and should be viewed as the probability that the loan becomes delinquent and is not repaid according to the initial terms, a scenario that is not necessarily associated with the legal bankruptcy and exit of the firm.

Banks calculate risk assessments based on a combination of hard (quantitative) and soft (qualitative) information, under strict regulatory guidelines. The hard information component combines balance-sheet data with past repayment history. The soft information component includes assessments of the quality of the firm's management and organizational structure, industry prospects, and the diversification of suppliers and customers. The headline risk assessment is then computed by combining the two components, with a weight on the soft information component which increases with firm size. For small firms, banks are instructed to only use hard information to perform the assessments. Instead, for medium and large firms the weight of soft information increases to 25 and 35 percent respectively. Interestingly, I can observe the soft and hard information assessments separately, making it possible to obtain some results separately for the two measures.

Because I will rely heavily on these risk assessments for my results, in Appendix A.1 I provide evidence that these assessments are predictive of subsequent loan losses. I measure loan losses for each loan as the fraction of the principal that is written-off or forgiven to the firm[7], and I show that a percentage point increase in the probability of default assessment leads to approximately a 0.4 percent increase in expected loan losses.

3 Empirics

I test the key prediction of competitive loan pricing, which posits that loan spreads should move one-for-one with firm default risk. I then explore how banking relationships shape this sensitivity. To measure default risk, I rely on the bank-reported risk assessments detailed in Section 2.2.

3.1 Relationships Determine Pass-Through

I show that whether a firm stays within a bank-firm relationship is a crucial determinant of the pass-through of default risk to lending rates. To establish this fact, I use pairs of consecutive loans obtained by the same firm in different months. I restrict my sample to uncollateralized loans, which represent over two thirds of the loans, and are a natural benchmark for my analysis, because a low pass-through is expected when a loan has good collateral, as banks have limited losses in case of default. Suppose that a firm f borrows at time τ and then again at time $t > \tau$. I denote the change in the loan spread between loans at time τ and t as $\Delta Spread_{ft}$, and the change in the risk assessment as ΔPD_{ft}. I then distinguish between firms that keep borrowing from the same bank, and firms that switch to a new lender with which they have no pre-existing relationship. I study the default pass-through as follows:

$$\Delta Spread_{ft} = \alpha + \beta \Delta PD_{ft} + \varepsilon_{ft} \tag{1}$$

Figure 1 shows that the pass-through β dramatically depends on lending relationships. When a firm keeps borrowing from the same bank, its borrowing rate is almost entirely disconnected from changes in its default risk. Instead, when a firm switches to a new bank, the response of spreads to firm risk is much larger and quantitatively close to one, as predicted by models of competitive loan pricing.

[7] Both charge-offs and forgiven loans represent a balance sheet loss for the bank, with the difference that with a charge-off the bank does not give up its right to collect the proceeds of the loan if the firm returns to solvency in the future. For the denominator, I use the maximum outstanding principal over the course of the loan.

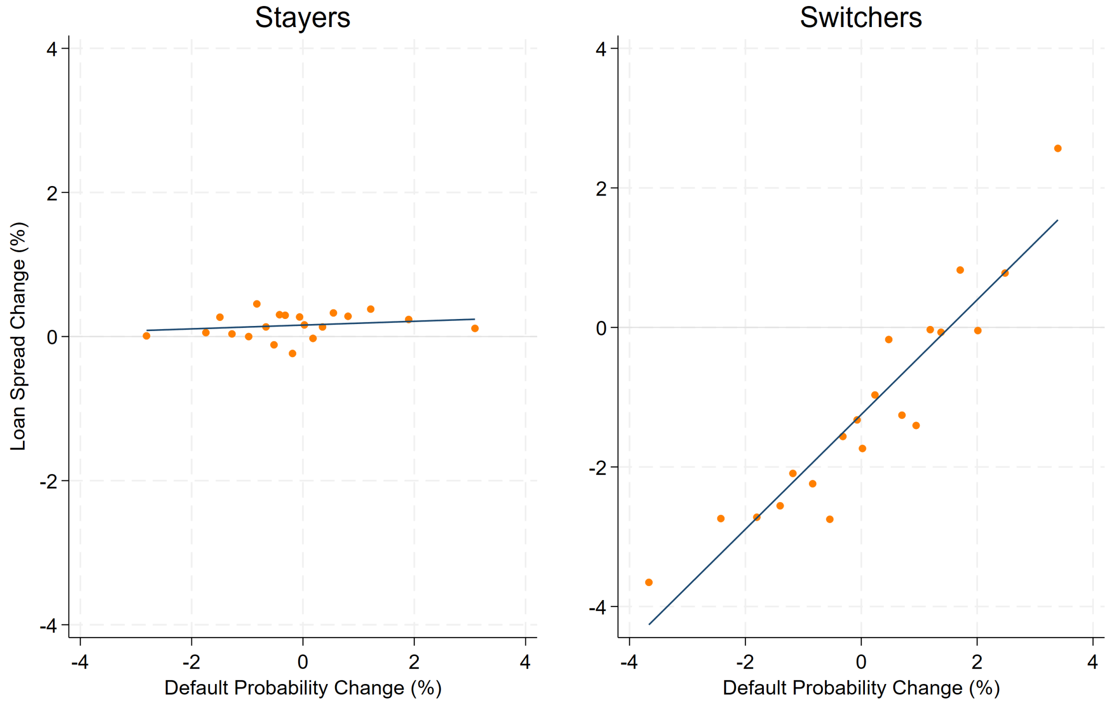


Figure 1: Scatterplot of changes in loan spread (y-axis) and changes in default probability (x-axis). Loan spread is the difference between loan rate and Mexico interbank rate TIIE 28. The default probability is the assessment provided by the lender. Changes are computed between two consecutive loans to the same firm originated either by the same bank (left plot) or by two different banks (right plot).

Notice that all loans in Figure 1 are newly originated, so there is no explicit commitment preventing the bank from adjusting rates to reflect the new information incorporated in the risk assessments.

To formally test the difference in pass-through between stayers and switchers, I regress changes in loan spreads on changes in default risk, and interact these with a dummy variable indicating whether firms stay with the same bank:

$$\Delta Spread_{ft} = \alpha + \beta_1 \Delta PD_{ft} + \beta_2 \Delta PD_{ft} \times \mathbb{1}_{ft}^{SameBank} + \beta_3 \mathbb{1}_{ft}^{SameBank} + \varepsilon_{ft} \tag{2}$$

The coefficient β_1 captures the pass-through for firms that switch to a new bank, while β_2 captures the dampening associated with staying with the same bank. The pass-through for stayers is thus $\beta_1 + \beta_2$. Results are shown in Table 3.

The main identification concern challenging these results is that switching to a new bank is not a random event. That is, firms are not randomly assigned to the left or the right side of Figure 1, and the dummy $\mathbb{1}_{ft}^{SameBank}$ may be related to unobserved drivers of $\Delta Spread_{ft}$ not captured by ΔPD_{ft}. I will discuss and address these identification concerns in the next subsection. But first, I outline a large battery of robustness tests which demonstrate how the

	(1) $\Delta Spread$	(2) $\Delta Spread$	(3) $\Delta Spread$
ΔPD	0.848***	0.609***	0.389***
	(0.025)	(0.026)	(0.041)
$\mathbb{1}_{SameBank} \times \Delta PD$	−0.787***	−0.560***	−0.359***
	(0.027)	(0.027)	(0.043)
Bank x Time FE		✓	✓
Firm FE			✓
R^2	0.0259	0.143	0.374
N	192,496	192,128	125,391

Table 3: This Table reports the OLS estimates of Equation (2). Pass-through of default risk assessments to loan spreads for switchers (top row) and dampening for stayers (bottom row). Robust standard errors in parenthesis.

presence of a loan pricing disconnect is broad-based across types of loans and firms, and not specific to a subset of the data.

Robustness tests. As documented in Section 2.1, loans are heterogeneous in their type, maturity and collateral. For my baseline results, I consider all loans, with the only exception of collateralized loans, which are excluded from the sample. In Appendix A.2, I show that, while the magnitudes change slightly across specifications, the dampening result holds when including collateralized loans or excluding credit lines, or considering only loans with short or long maturity. The result is also robust to excluding pairs of loans that have exactly the same loan spread or lending rate in both loans, which limits concerns related to renegotiation frictions, or to keeping only pairs of loans that are at least 12 months apart. Furthermore, it holds even when focusing exclusively on the component of risk assessments arising from hard or soft information.

3.2 Endogenous Switching

The results in the previous section uncover a stark role for relationships in determining default risk pass-through, but also raise questions about the role played by the switching decision. In this section, I first outline the key summary statistics for stayers and switchers, which shed light on the drivers of the switching decision and on its interaction with the insurance mechanism. Then, I show that the key role of relationships in determining default pass-through is robust to three approaches to deal with endogenous selection.

3.2.1 Who are the switchers?

To better understand the extent of selection into switching, I compare the key summary statistics for firms that stay with the same bank and those that switch, effectively those appearing on the left and the right panels of Figure 1.

Table 4 displays the key summary statistics for firms. It shows that stayers and switchers are comparable, which alleviates concerns that we may be confronting two different type of firms. While there are some differences driven by the very right tail, the center of the distribution (median) is almost identical among the two groups for all variables. Notice that the summary statistics differ from those reported in Table 1 since here they are computed at the firm-month level rather than at the firm-level, as switching is not a permanent characteristic of a firm, and this overweights firms that borrow more often, which tend to be larger and safer.

Variable	Group	Mean	p10	p50	p90
Employees	Stayers	93.62	1	5	69
	Switchers	41.33	1	5	50
Sales ($'000)	Stayers	32,143.38	8.28	268.37	10,195.01
	Switchers	10,822.02	5.59	274.69	4,731.45
Firm Age	Stayers	21.37	6	21	36
	Switchers	19.47	4	19	35
Spread	Stayers	9.84	3.35	9.00	18.71
	Switchers	9.01	4.01	8.56	15.00
PD	Stayers	6.53	1.07	2.09	8.15
	Switchers	3.97	0.53	2.35	4.45

Table 4: Summary statistics of firm characteristics. Statistics are computed at the firm-month level, with each firm considered in all months they borrow. Stayers are firms that borrow from the same bank of their previous loan. Switchers are firms that borrow from a new bank with which they have no preexisting relationship, excluding first-time borrowers. p10, p50 and p90 are the percentiles of the distribution of each variable.

Table 5 displays additional moments that are more informative about the source of the switching decision. First, in the top two rows, we notice that firms tend to receive a large rate reduction on average when they switch, which suggests that they are searching for better deals. Second, the bottom two rows highlight that switchers are firms that should indeed on paper benefit more from switching, as they tend to be charged high rates compared to their fundamental—a positive pricing wedge—in their prior relationship. This also ties back to the insurance mechanism, as switchers tend to be firms that have experienced a decline in their probability of default since the onset of their prior relationship ($PD_{t-1} < P_0$), and thus are likely to be disadvantaged by the insurance mechanism, which may leave them overcharged compared to their improved risk profile.

Variable	Group	Mean	p10	p50	p90
$\Delta Spread$	Stayers	0.03	-2.03	0	2.01
	Switchers	-0.90	-7.59	-0.67	5.33
$\Delta LoanRate$	Stayers	0.17	-2.21	0	2.75
	Switchers	-0.78	-8.86	-0.55	6.67
Pricing wedge[8]:	Stayers	0.19	-5.10	-0.78	8.09
$R^{wedge}_{ft} = (Spread_{t-1} - PD_{t-1})$	Switchers	0.48	-4.51	-0.25	7.48
Cum. ΔPD over relationship	Stayers	-0.25	-1.78	-0.44	1.66
$(PD_{t-1} - P_0)$	Switchers	-0.42	-1.91	-0.58	1.32

Table 5: Summary statistics of loan characteristics. Statistics are computed at the firm-month level, with each firm considered in all months they borrow. Stayers are firms that borrow from the same bank of their previous loan. Switchers are firms that borrow from a new bank with which they have no preexisting relationship, excluding first-time borrowers. p10, p50 and p90 are the percentiles of the distribution of each variable.

3.2.2 Identification

In this section, I first clarify the nature of the identification challenge arising from endogenous switching for correctly estimating the pass-through of default risk, and then I propose three approaches to address it.

Threats to Identification. The identification concern is that relationships may be irrelevant, but firms whose spread is more sensitive to risk assessments—perhaps due to the assessments being more informative—self-select as switchers. I interpret the exercise in Equation 2 first and foremost as a test of the theoretical prediction that $\Delta Spread = \Delta PD$. If this prediction held exactly in the data, and our measure of default risk was accurate, then we would always obtain $\beta_1 = 1$ and $\beta_2 = 0$, even if $\mathbb{1}_{SameBank}$ is correlated with ΔPD or other unobservables, which could occur if firms are more likely to switch when their probability of default is declining. If all banks priced loans competitively, the relationship $\Delta Spread = \Delta PD$ would still hold for all firms, and we would still estimate $\beta_1 = 1$ and $\beta_2 = 0$ regardless of selection issues.

The true identification concern, instead, emerges if endogenous switching is correlated with inaccuracies in risk assessments, which could arise because assessments are more informative for certain firms, for certain loan types or for certain periods, leading to a true competitive pricing pass-through that differs from one. In the following paragraphs, I present three approaches to address this form of endogeneity.

Pass-through before and after the switch. If relationships were irrelevant, and switchers were simply high pass-through firms, then we would expect a high pass-through also before and after the switch. Instead, Table 6 shows that high pass-through only occur when a firm is separating from its bank. Immediately before and after the separation, when the firm remains with the same bank, the pass-through is low.[9] Thus, we can rule out that switchers are inherently higher pass-through firms.

	(1)	(2)	(3)	(4)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
	stayers	before switch	at switch	after switch
ΔPD	0.001	-0.148	0.672***	0.091
	(0.008)	(0.102)	(0.057)	(0.051)
R^2	0.000124	0.00165	0.0267	0.000629
N	149,839	1,297	5,864	4,661

Table 6: Pass-through from changes in default risk to changes in loan spread. First column: all pairs of consecutive loans from the same bank. Second column: firms that are observed at least two periods before switching to a new bank. Third column: firms that switch to a new bank. Fourth column: firms that are observed at least two periods after switching to a new bank. Robust standard errors in parenthesis.

Pass-through for multi-bank firms. The exercise above shows that switchers are not firms with a permanently higher pass-through, but one might still worry that firms may self-select into switching during periods when their pass-through would have been high even within a relationship. The core issue, common to all selection problems, is that we only observe the outcome either when a firm switches banks or when it does not. For a subset of firms, however, we have the unique opportunity to observe outcomes in both cases, since some firms start borrowing from a new bank, but simultaneously take out a new loan from their old bank. Specifically, I restrict the sample to firms that initially borrow from only one bank (A), and then, in a subsequent month, from both their old bank (A) and a new bank (B). Table 7 reports the results, showing that yet again there is a low pass-through for the new loan obtained from the same bank, but a high pass-through for the one obtained from a different bank.

[9] The fact that the pass-through is low even for the first period immediately following separation is related to the finding in Appendix A.9 that being in a relationship is much more important in determining the pass-through than the length of the relationship itself.

	(1)	(2)
	$\Delta Spread$	$\Delta Spread$
ΔPD	0.926^{***}	0.970^{***}
	(0.043)	(0.056)
$\Delta PD \times \mathbb{1}_{SameBank}$	-0.903^{***}	-1.103^{***}
	(0.113)	(0.144)
Bank x Time FE		✓
R^2	0.0455	0.221
N	11,502	11,221

Table 7: Pass-through of default risk to loan spreads. The sample is restricted to firms that within the same month borrow from a new bank and take out a loan from the same bank of their last loan prior to the month. The dummy variable identifies the loan taken from the same bank. Robust standard errors in parenthesis.

In Appendix, I show that the results in Table 7 are robust when using changes in default risk driven only by the hard information component, and excluding the soft one, which make assessments identical across banks.

Instrumenting for separation. I propose two new instruments providing exogenous variation in the decision to switch banks. These are based on the insight that shocks to the old bank which are plausibly orthogonal to firm fundamentals may induce firms to switch. The first instrument uses variation in the availability of bank branches in the municipality where the firm is located. The second leverages bank credit-supply shocks estimated using the methodology from Amiti and Weinstein (2018). Note that in equation (2) the possibly endogenous variable $\mathbb{1}_{ft}^{SameBank}$ appears twice: once by itself and once interacted with the change in firm risk. Following Wooldridge (2010),[10] I instrument the interaction term using the product of the instrument and the change in firm risk, $(Z \times \Delta PD)$.

To construct the first instrument, I use branch-level data at the municipality level described in Appendix A.11, which I obtained from the *Inclusión Financiera* reports of CNBV, the banking supervisor. For each firm, I construct the change in the fraction of bank branches in its municipality which are operated by their old bank. Formally, for a firm f headquartered in municipality m, borrowing at time τ from bank b, and then borrowing again from any bank at time $t > \tau$, the instrument is defined as: $Z_{ft}^{branch} = (MS_{bmt} - MS_{bm\tau})$, where $MS_{bmt} = N_{bmt} / \sum_{b'} N_{b'mt}$ with N denoting the number of active branches. This instrument captures changes in the old bank's market share in the municipality, which reflects both the old bank's expansion or contraction and the entry or exit of other banks. In Appendix, I show that using only changes in bank branches of the firm's old bank, without incorporating information about its competitors, yields a weak first stage and an unintuitive sign, likely

[10] Chapter 6.2.1.

because variation in this instrument is concentrated in rural areas with few banks.

The second instrument is based on bank credit supply shocks, constructed using the methodology from Amiti and Weinstein (2018), and relies entirely on credit registry data. These shocks are identified by leveraging the presence of firms that borrow from multiple banks, along with firm and bank fixed effects, which make it possible to isolate variation in credit supply that is specific to each bank and orthogonal to firm demand. I estimate these credit supply shocks for each bank-month, denoted as $\tilde{\beta}_{bt}$, with the methodology of Amiti and Weinstein (2018), as summarized in Appendix A.12. Then, I construct the firm-specific switching instrument as the cumulative sum of shocks to the old bank (bank b) over the period between the firm's previous loan at time τ and its subsequent borrowing at time t: $Z_{ft}^{AW} = \sum_{j=\tau+1}^{t} \tilde{\beta}_{bj}$.

The results from the two instrumental regressions, alongside the OLS estimates, are reported in Table 8. Once again, we observe a large pass-through for switchers and a significant dampening for stayers, suggesting that, if anything, the OLS may understate the true effect of relationships.

	OLS	Branch IV	AW IV
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	0.848^{***}	2.651^{***}	3.133^{***}
	(0.025)	(0.937)	(1.055)
$\Delta PD \times \mathbb{1}_{SameBank}$	-0.787^{***}	-3.145^{***}	-3.820^{***}
	(0.027)	(1.214)	(1.257)
First stage F-stat		152.8	74.11

Table 8: Pass-through of default risk to loan spreads. First column: OLS. In the last two columns, the dummy for stayers is instrumented. The instrument is the branch IV in the second column and the credit supply shocks of Amiti and Weinstein (2018) in the third column. Robust standard errors in parenthesis.

The exclusion restriction assumes that changes in the instrument (market share or credit supply shocks) are not correlated with unobserved factors affecting loan spreads beyond what is predicted by risk assessments. I argue that it is plausible that these shocks are orthogonal to the firm. For example, the credit shocks are based on a structural decomposition between demand and supply. A violation of this would require that shocks to the bank also explain changes in firm risk beyond what is predicted by risk assessments. There is, however, a possibility that the instruments may violate the exclusion restriction if they change the bank loan pricing policies conditional on firm risk, for instance if banks that experience a contraction in branches or credit supply also raise their loan rate pass-through. To address this, I provide in Appendix A.4 an estimation of the pass-through using Heckman's correction to account for selection into switching, which relies on the weaker assumption that the instrument from the old bank (*e.g.* its credit supply shock) is unrelated to loan pricing at other banks.

In the next sections, I present evidence of two additional pricing patterns that complement the result on limited pass-through, and help differentiate between competing theories: history-dependence and pricing reconnect upon switching.

3.3 History Dependence

If loans were priced competitively, only the contemporaneous risk - the future default risk of the loan originated today - should determine the spread, and past information should be irrelevant for loan pricing. Instead, I find that spreads charged by banks depend not only on contemporaneous default risk, but also, and even more prominently, on the risk assessment made at the onset of the bank-firm relationship.[11]

To formally test for history-dependence, I regress the spread charged by banks in a given period, $Spread_{ft}$, on both the contemporaneous default risk, PD_{ft}, and the default risk at the start of the bank-firm relationship, denoted PD_{f0}:

$$Spread_{ft} = \alpha + \beta_1 PD_{ft} + \beta_2 PD_{f0} + \varepsilon_{ft} \tag{3}$$

	(1)	(2)	(3)
	$Spread_t$	$Spread_t$	$Spread_t$
PD_t	0.258^{***}	-0.001	0.004
	(0.068)	(0.002)	(0.017)
PD_0	0.750^{***}	0.312^{***}	0.208^{***}
	(0.057)	(0.046)	(0.048)
Cons	5.591^{***}	6.388^{***}	6.636^{***}
Firm FE			✓
Bank x Time FE		✓	✓
R^2	0.0494	0.415	0.856
N	146,988	146,677	103,643

Table 9: Test for history-dependence. PD_t is the contemporaneous default risk. PD_0 is firm default risk at onset of relationship between the firm and the bank originating the loan at time t. Robust standard errors in parenthesis.

[11] This result echoes Dougal et al. (2015), who test for anchoring in credit spreads using syndicated loan data, showing a role for past spreads in determining loan pricing, which is a form of path-dependence. Instead, I test for history-dependence, the relevant of past fundamental conditions, which is a typical test for long-term contracts (Beaudry and DiNardo (1991)). Section 6.5 compares the insurance mechanism with behavioral anchoring and other possible explanations.

Table 9 shows that loan spreads depend on past information, which should be irrelevant for pricing purposes in a competitive market. This past information is specifically the risk assessment at the onset of a relationship PD_0, even after controlling for the contemporaneous default risk PD_t. This result is important because the presence of history-dependence is a typical test for implicit long-term contracts performed in labor economics (Beaudry and DiNardo (1991)), as in an insurance arrangement banks and firms agree on the credit conditions throughout the relationship at the onset of the match. Instead, the relevance of past conditions is inconsistent not only with competitive pricing, but also with a large class of models with static bank market power or with loan evergreening to zombie firms. One possible concern with the interpretation of the history-dependence result as driven by relationship factors is that the updating rule may be imperfect. With proper Bayesian updating the most recent assessment should capture all relevant information, but it is an empirical question whether this happens in practice. In Appendix A.1, I show that, after accounting for contemporaneous default risk, past assessments have minimal predictive power for loan losses. In Appendix A.5, I also show that past default risk is relevant for loan pricing both conditionally and unconditionally on past loan spreads.

3.3.1 Competition

To further corroborate the role of limited mobility in affecting loan pricing, I study how my results of limited pass-through and history-dependence depend on the ability of a firm to switch bank. I measure this ability using branch data at the municipality-level, and for each firm I construct the market share of branches of their bank in the municipaility where they are headquartered. I find that when firms have limited outside banking options, then loan pricing exhibits weaker pass-through and stronger history-dependence. Results from this exercise are reported in Appendix A.3.

3.4 Switching and Pricing Reconnect

Due to limited pass-through and history-dependence, deviations of loan rates from their competitive pricing benchmark, which I refer to as pricing wedge $R^{wedge}_{ft} = Spread_{ft} - PD_{ft}$, can accumulate over the course of a relationship. I now show that a substantial fraction of such wedge is closed upon switching. For example, suppose that a firm is currently paying a spread above its default risk, perhaps because it has become significantly safer since the start of its relationship. Then, by switching to a new bank, such firm would benefit from a substantial rate reduction.

To establish this fact, I regress changes in loan spreads against the magnitude of the prior loan pricing wedge $R^{wedge}_{f\tau}$, where τ captures as usual the time at which the firm last borrowed:

$$\Delta Spread_{ft} = \alpha + \beta_1 R^{wedge}_{f\tau} + \varepsilon_{ft} \tag{4}$$

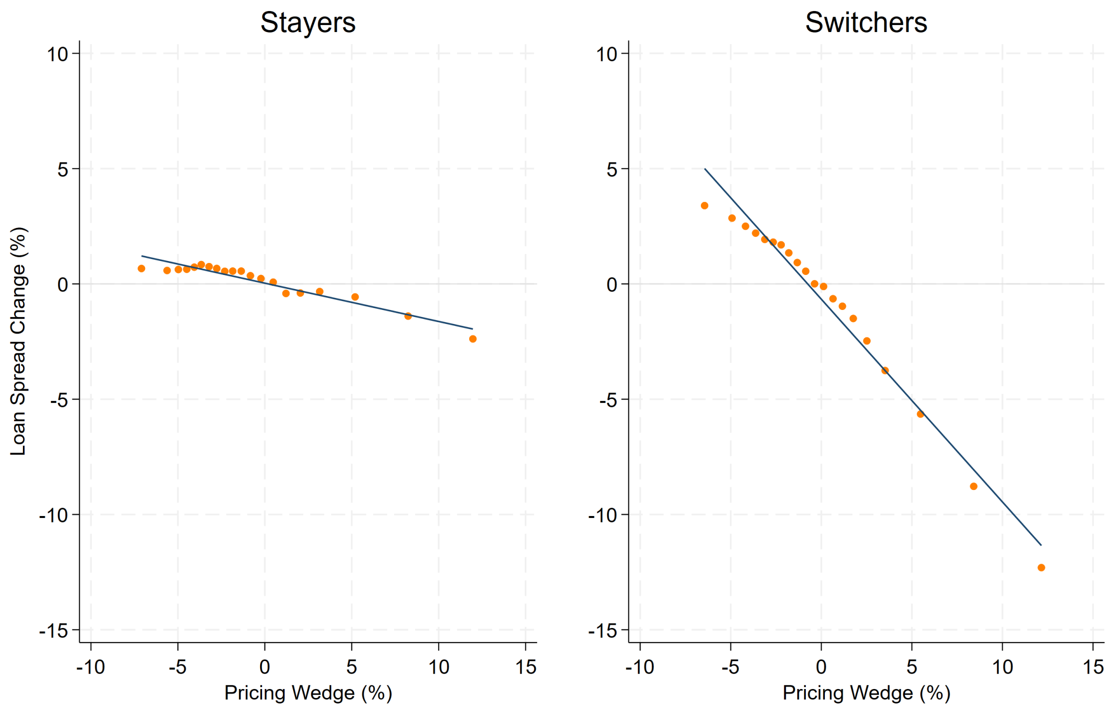


Figure 2: Scatterplot of changes in loan spread (y-axis) and loan pricing wedge in the prior loan (x-axis). The prior loan pricing wedge $R^{wedge}_{f\tau}$ is the difference between the loan spread and the default risk in the previous loan taken out by the firm.

Results in Figure 2 show that upon separation over 90 percent of the previous pricing wedge is corrected. Instead, when a firm borrows from the same bank, only about 10 to 20 percent of such wedge is closed, which suggests that wedges are persistent within relationships, consistently with the strong history dependence documented in section 3.3. A regression table for this plot is available in Appendix A.8.

3.5 Taking Stock of the Empirical Results

The case study outlined in Table 10 provides an effective summary of the loan pricing patterns documented in this section.

When the firm stays with the same bank, the pass-through from default risk to lending rates is small, and firm risk at the onset of the relationship with its bank is important for loan pricing also in subsequent loans. Instead, when it starts borrowing from a new bank, the new loan rate is close to the competitive benchmark. Therefore, changes in default risk upon separation are fully passed-through. On top of that, a large fraction of the pricing wedge

	t_0 Bank 1	t_1 Bank 1	t_2 Bank 1	t_3 Bank 2	t_4 Bank 2
PD	5%	4%	3%	2%	3%
Spread	5%	5%	4.5%	2%	2%
Pricing Wedge	0%	1%	1.5%	0%	1%
ΔPD	-	-1%	-1%	-1%	1%
$\Delta Spread$	-	0%	-0.5%	-2.5%	0%

Table 10: Case study of loan pricing. The top two rows show the probability of default assessed and the spread charged by the bank in any given period. The vertical line indicates a separation between the firm and the old bank.

in the previous relationship is corrected. Indeed, upon switching, the loan spread drops by 2.5%, which reflects both the 1% drop in risk and the correction of the 1.5% pricing wedge in t_2.

4 Model

In this section, I develop a general equilibrium model in which firms establish long-term relationships with banks. This is motivated by two objectives. First, to rationalize the empirical pricing patterns documented in the previous section, and to obtain new testable predictions consistent with the insurance mechanism. Second, to study and quantify the aggregate implications of relationships.

I formalize relationships as optimal long-term contracts, allowing banks and firms to negotiate future credit terms, potentially leading to deviations from the zero-profit pricing rule. The model introduces two key innovations compared to the existing literature on optimal contracts: first, it assumes that one-period debt is enforceable; second, it incorporates switching costs that firms face when starting new relationships.

Debt enforceability refers to the ability of a court to liquidate the firm if it fails to repay the bank. This means the bank can always force the firm to repay, even when it switches to a new bank, as long as the firm has sufficient financial capacity. This does not prevent default, which occurs when a firm is unable to repay, as banks cannot seize personal household assets. All other aspects of the contract, including promises on future credit terms, are subject to limited commitment: both parties can renege and establish a new relationship.[12] For instance, if a firm promises to pay a high lending rate when a particular state occurs, it may still seek lower rates from other banks when that state materializes. However, it remains obligated to extinguish its debt with the old bank.

[12] I view these promises as informal and not legally enforceable— for instance, because courts may be unable to verify the state that materializes, or more simply because it is very costly to write legal contracts for each contingency.

The assumption of enforceable one-period debt is not only realistic in many settings, but also crucial from a technical standpoint. In standard bank-firm contracting models,[13] the concepts of lending rates and default are not explicitly defined; rather, only a sequence of state-contingent transfers is determined.[14] Incorporating enforceable debt provides a clear rationale for setting debt within relationships: debt plays a direct role when the bank and firm separate, both *on-path* and *off-path*. By providing an explicit role for debt, lending rates and default, I can naturally map my framework to the data and study how lending rates are determined within banking relationships.

Switching costs reflect frictions that limit mobility in credit markets, such as red tape, the sunk screening costs sustained by the bank, or the cost of traveling to a more distant bank branch. In the model, switching costs play a critical role in determining the structure of financial markets. I will show that as switching costs approach zero, the model converges into a competitive credit market, where shocks are fully passed through to loan rates. Conversely, with positive switching costs, deviations from complete pass-through can be sustained, which enables optimal risk-sharing between banks and firms.

In Appendix B.5, I provide a detailed examination of each assumption's assumption role and how different assumptions map into different models in the literature. This overview is succinctly summarized in Table 37 of Appendix B.5.

In the following sections, I formally describe the model setup and characterize the optimal contract.

4.1 Setup

The economy features a continuum of heterogeneous firms and identical banks. Time is discrete. The production side of the economy is standard: firms produce using a neoclassical production technology $y(z,k,l) = zk^{\alpha}l^{\nu}$, where z is the firm's idiosyncratic productivity, which follows an AR(1) process. They start their life with zero capital and debt, but can accumulate both over time. Capital depreciates at rate δ, and investment is subject to convex capital adjustment costs.

The financial side of the economy is more involved. When a firm enters the economy, banks compete to start a new relationship by offering long-term contracts to firms, denoted

[13]For instance, Cooley, Marimon, and Quadrini (2004), Albuquerque and Hopenhayn (2004), and Kovrijnykh (2013).

[14]Optimal contracts typically pin down the sequence of state-contingent transfers $T(h^t) = b(h^t) - Q(h^t)b'(h^t)$. Infinitely many combinations of $\{Q(h^t), b(h^t)\}_t$ could theoretically achieve such transfers. One approach, in the tradition of DeMarzo and Sannikov (2006), is to identify ex-post which securities can implement these transfers. Alternatively, physical capital is sometimes used to represent debt, but this approach is inherently limited to stylized environments where (i) capital is non-durable, meaning the firm begins each period with no capital, and (ii) the firm cannot retain earnings, so it starts each period without savings.

$\mathscr{C}$. These contracts specify not only the current period's credit terms, but also informal agreements on future liquidation and credit decisions. As discussed in the previous section, the key novelty lies in the commitment structure: contracts $\mathscr{C}$ are subject to limited commitment, as both parties can terminate their current relationship, but firms remain liable to repay their outstanding debt b. Contracts are complex objects, detailing the firm's credit and investment policies for every history realization: $\mathscr{C} = \{Q_t(h^t), b_{t+1}(h^t), k_{t+1}(h^t), l_t(h^t), d_t(h^t), Exit_t(h^t)\}_{h^t, t>t_0}$. Here, Q is the lending rate, b is the borrowed amount, k is physical capital, l is labor, d represents dividends, and $Exit$ indicates a recommendation to liquidate the firm. Debt and capital have a $(t+1)$ subscript because they are predetermined upon meeting with the bank. Contracts are summarized by the promised value v to the firm. New firms select the contract with the highest promised value from those offered by banks.

Incumbent firms are matched with a bank. At the beginning of each period, a firm's state is captured by (z, b, k, v), where z denotes productivity, b is debt, k is capital and v is the promised value summarizing the contract $\mathscr{C}$ with the bank. For notational convenience, I denote by $s = (z, b, k)$ the triplet of productivity, debt and capital of the firm. This (s, v) notation is useful because the firm's outside option will solely depend on s, while v represents the implicit promises exchanged with the current bank, which become irrelevant in the event of a separation. Each period, a fraction θ of randomly selected bank-firm matches is terminated exogenously, leading the firm to search for a new lender in the banking market. When separation occurs in state s, either endogenously or exogenously, the firm receives value $v^m(s)$ while the bank obtains value $b^m(s)$. These values are endogenously determined by the free-entry condition for new banks to start relationships, but for now it is easier to consider them as given functions.

Switching decision. At the beginning of each period, when a match is not exogenously terminated, both parties can decide to either honor their implicit agreement and continue the relationship or form a new one. If a firm switches, it must repay its debt b to the old bank, and simultaneously start a new relationship. Doing so, the firm obtains value $v^m(s)$ minus a switching cost ψ^f. Similarly, the bank can choose to renege on its past promises and enforce repayment of debt b. Deviating banks incur a penalty ψ^b, which can be interpreted either as a reputation cost, a sunk screening cost for starting a new relationship, or as the loss of profits from cross-selling multiple products to the firm, such as deposits. I will later discuss in greater detail the situation in which the firm is insolvent, meaning that it is unable to repay b without the support of its current bank. Firms can also exit costlessly and obtain a value $\underline{v}$, in which case they are not liable for their legacy debt b. Setting $\underline{v} \geq 0$ ensures that the firm's equity stays non-negative. If neither party opts to deviate, they adhere to the terms of their contract, which specifies a liquidation choice and credit policy in the event of continuation.

The timeline of the model within each period is summarized in Figure 3.

Recursive Formulation. I solve the problem using a recursive formulation, following

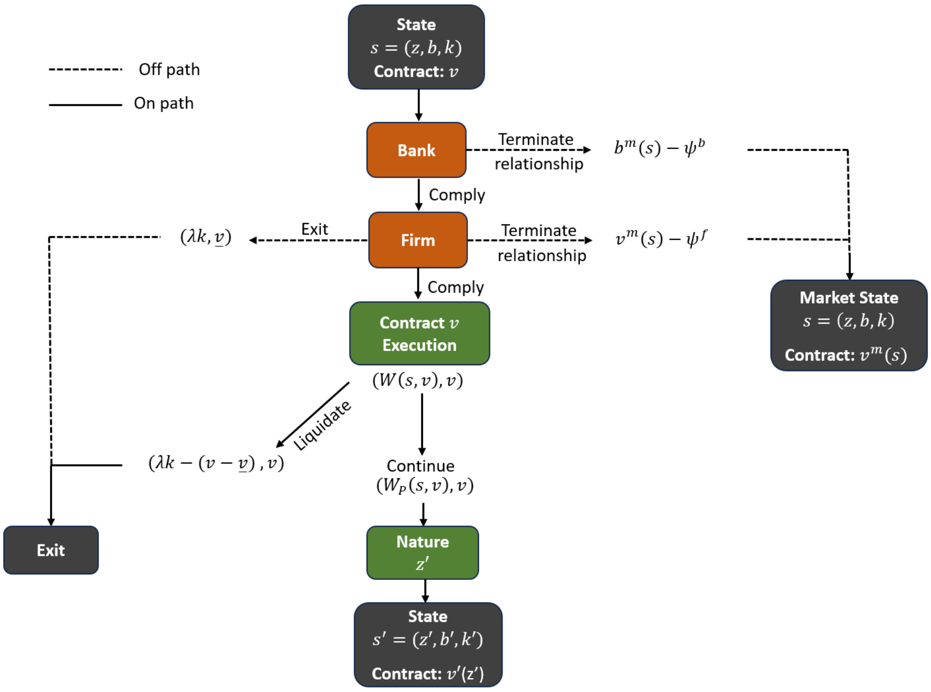


Figure 3: Timeline of each period. A bank-firm match starts with state $s = (z, b, k)$ and with a promised value v, which summarizes the liquidation and credit policy today and in all future histories. The bank can always end the relationship, but incurs a cost ψ^b. Separating firms can either start a new relationship, incurring a cost ψ^f, or exit costlessly.

Spear and Srivastava (1987): solving for the optimal long-term contract can be reduced to a sequence of static problems. Starting from state $s = (z, b, k)$ and promised value v, the bank chooses how to optimally deliver value v to the firm. I impose participation constraints (PCs) when choosing tomorrow's values $v'(z')$. Therefore, at this stage we can focus on finding how the bank can best deliver value v to the firm, without worrying whether v is incentive-compatible.

Liquidation decision. The first decision of the period is whether to shut down the firm. If the firm is liquidated, the bank recovers a fraction λ of its capital, and a cash transfer is used to settle the remaining promised value $(v - \underline{v})$, if positive. The bank's value in liquidation is thus: $\lambda k - (v - \underline{v})$. Alternatively, the bank can choose to continue to the production phase and obtain a value $W_P(s, v)$, defined below. The beginning-of-period value $W(s, v)$ is thus defined as follows:

$$W(s, v) = \max\{W_P(s, v), \lambda k - (v - \underline{v})\}$$

Production Phase. If a bank-firm match reaches the production phase with state $s =$

(z,b,k) and promised value v, the bank faces the following problem:

$$\underbrace{W_P(s,v)}_{\text{Bank value}} = \max_{Q,b',k',\{v(z')\}} \underbrace{b-Qb'}_{\text{Bank flow}} + \underbrace{\beta(1-\theta)E_{z'|z}[W(s,v'(z'))] + \beta\theta E_{z'|z}[b^m(s')]}_{\text{Bank cont. value}}$$

Subject to:

$$\underbrace{d}_{\text{Dividend}} + \underbrace{\beta(1-\theta)E_{z'|z}[v(z')] + \beta\theta E_{z'|z}[v^m(s')]}_{\text{Firm cont. value}} \geq v \qquad (\mu\text{: PK})$$

$$\underbrace{v(z')}_{\text{Promised value}} \geq \underbrace{\underline{v}}_{\text{Exit value}} \qquad (g(z')\text{: PC-Exit Firm})$$

$$\underbrace{v'(z')}_{\text{Promised value}} \geq \underbrace{v^m(s')}_{\text{Firm outside value}} - \underbrace{\psi^f}_{\text{Switching cost}} \qquad (\eta(z')\text{: PC-Mkt Firm})$$

$$\underbrace{W(s',v')}_{\text{Implied bank value}} \geq \underbrace{b^m(s')}_{\text{Firm outside value}} - \underbrace{\psi^b}_{\text{Switching cost}} \qquad (q(z')\text{: PC-Mkt Bank})$$

Where the budget,[15] is:

$$\underbrace{py^*(z,k,w)-f}_{Output} = \underbrace{(k'-(1-\delta)k+\Phi(k,k'))}_{Capex} + \underbrace{(b-Qb')}_{Net\ Repayment} + \underbrace{d(1+\tau\mathbb{1}_{d<0})^{-1}}_{Dividends} \qquad (5)$$

The bank's objective is to maximize the sum of its cash-flow today $(b-Qb')$ and its continuation value. This is achieved by choosing the lending rate Q, the debt level b', the installed capital k', and future value promises to the firm $\{v'(z')\}$, contingent on each shock realization z'. Several constraints define this problem. First, the promise-keeping constraint (PK): the bank must deliver the promised value v to the firm, through either dividends d or future promised value. Additionally, the participation constraint (PC) requires that all promises $\{v'(z')\}$ must be incentive-compatible. The firm must prefer to comply with the contract rather than exit the market or start a new relationship. The bank must prefer to comply rather than terminate the relationship and enforce the repayment of the debt. I next describe in greater detail the outside options of banks and firms.

Outside Options and Firm Solvency. In equilibrium, $v^m(s)$ and $b^m(s)$ are determined endogenously in the market for new relationships. At the beginning of this section, I briefly discussed the case of an entrant firm, that by definition has zero legacy debt ($b=0$) and no old bank. This will no longer be the case when an incumbent firm enters the banking market, making the problem more complex. I now formally outline the functioning of the banking

[15] In Section 7.2 I will cast the model in a New Keynesian framework to study monetary and fiscal policy. There, the price at which firms sell their output, p, will play a key role. For now, it suffices to know that it is a constant equal to $p=\frac{\gamma_{nk}-1}{\gamma_{nk}}$, where γ_{nk} is a markup term of the retailers that purchase the output of the firms. For theoretical purposes, when studying the steady state we can simply consider $\gamma_{nk}\to\infty$, which implies $p=1$, and is equivalent to assuming that firms directly produce the final good. The markup γ_{nk} is the only element of the New Keynesian block that is relevant for the steady state and the contracting problem.

market for a generic firm, which will nest the case of entrant firms with $b = 0$ as a special case.

Ultimately, the solution for $b^m(s)$ will be straightforward. For a solvent firm, $b^m(s) = b$, as the bank recovers its outstanding loan. Instead, if the firm is insolvent, the bank cannot recover the full amount, leading to $b^m(s) < b$. I will now outline the contracting protocol that results in this outcome.

If the bank and the firm separate, they face competitive banks offering contracts to attract the firm. In standard models, contracts are typically summarized by a promised value (e.g., V), but in the banking context of my model the situation is more complex because there are three parties involved: the new bank which is posting the contract, the firm, and the old bank. As a result, contracts posted by a new bank are summarized by a pair $\{B, V\}$, where B represents the monetary transfer to the old bank used to settle the legacy debt b, and V denotes the promised value to the firm. Because of perfect competition to start new relationships, the new bank must break even: its value from starting the new relationship must equal the monetary transfer to the old bank, $W(s,V) - B = 0$. However, for each state s many combinations of $\{B, V\}$ may satisfy this zero-profit condition, and could thus be offered by new lenders.

Which party, the firm or the old bank, chooses which contract to pick among those posted by new banks? I make an assumption that is arguably quite natural and ensures that the model converges into the competitive market case as switching costs approach zero. Specifically, if there are contracts that offer $B \geq b$, the firm is considered *solvent* and is free to select the contract with the highest V among those that provide $B \geq b$ to the old bank. Conversely, if no contract offers $B \geq b$, the firm is deemed *insolvent* and it is captive of the old bank. In this case, the old bank chooses the contract with the highest B.

Let $\overline{b}(s)$ represent the maximum amount that new banks would be willing to pay the old bank, by making the firm indifferent to exit or continuation:

$$\overline{b}(s) = W(s, \underline{v})$$

Solvent firms have $b \leq \overline{b}(s)$, as the amount they can raise from a new bank is sufficient to repay their legacy debt. The value these firms can achieve by starting a new relationship solves:

$$W(s, \overline{v}(s)) = b$$

Firms with $b > \overline{b}(s)$ are *insolvent*. In this case, the old bank can pick the contract with the highest B, which equals $\overline{b}(s)$, and delivers value $\underline{v}$ to the firm. I make the additional assumption that in the event of insolvency, the old bank only recovers a fraction γ^{sep} of this amount, resulting in $\overline{b}^m(s) = \gamma^{sep}\overline{b}(s)$, with the remaining fraction representing a deadweight loss. This *bankruptcy cost* introduces an external finance premium and is a crucial financial constraint that prevents firms from reaching the first-best level of capital.

Imposing $\gamma^{sep} < 1$ in models with debt is standard to prevent costless renegotiation.[16] I provide additional details on this assumption in Appendix B.5, and I formalize how to add such *bankruptcy costs* also within a match in Appendix B.3.1, which is important for both quantitative and theoretical reasons.[17]

The equilibrium outside values are summarized as follows:

$$v^m(s) = \begin{cases} \bar{v}(s) & \text{if } \bar{b}(s) \geq b \\ \underline{v} & \text{if } \bar{b}(s) < b \end{cases} \tag{6}$$

$$b^m(s) = \begin{cases} b & \text{if } \bar{b}(s) \geq b \\ \gamma^{sep}\bar{b}(s) & \text{if } \bar{b}(s) < b \end{cases} \tag{7}$$

where $\bar{b}(s) = W(s, \underline{v})$ and $W(s, \bar{v}(s)) = b$.

4.2 Household and Equilibrium

The representative household owns all the firms and the banks in the economy through a mutual fund. The household's preferences over consumption and labor are described by the following utility function:

$$\mathscr{U} = \sum_t \beta^t E_t \left[\frac{C^{1-\sigma}}{1-\sigma} - \psi_L L_t\right] \tag{8}$$

The consumption-saving problem determines the stochastic discount factor $\Lambda_{t+1} = \beta \frac{C_{t+1}^{-\sigma}}{C_t^{-\sigma}}$, which in steady state simplifies to $\Lambda = \beta$.

Denoting as usual $s = (z, b, k)$, an equilibrium consists of value functions $W(s, v)$ and $W_P(s, v)$; decision rules for capital $k'(s, v)$, debt $b'(s, v)$, dividends $d(s, v)$, labor $l(s, v)$, and contractual promises contingent on the idiosyncratic state $v'(s, v; z')$; a measure of bank-firm matches $\mu(s, v)$; and prices w, p, such that (i) all contracts are written optimally, (ii) the household optimizes, (iii) the steady state distribution is consistent with decision rules, and (iv) the labor, asset and good markets clear.

For most of the paper, I will focus on the steady state of the economy described above. Then, in Section 7.2, I will define the equilibrium with aggregate shocks and nominal rigidi-

[16]For instance, setting $\gamma^{sep} = 1$ would lead to the unrealistic scenario in which firms choose excessively high debt levels b', only to renegotiate them costlessly in most states, effectively making debt a contingent instrument.

[17]For quantitative reasons, having *bankruptcy costs* only in the event of separation imply that such financial friction only bites with probability θ, the exogenous separation rate. When θ is small, the model converges into the first best case. From a theoretical viewpoint, the proof the model converges into a competitive market as switching costs approach zero only holds if either $\gamma^{sep} = 1$ or $\gamma^{sep} = \gamma^{cont}$, as both conditions ensure that there is no intrinsic difference between the old and the new bank.

ties to study how relationships affect the economy's response to monetary and fiscal policy. The steady state equilibrium is the special case of this more general economy when there are no aggregate shocks.

5 Equilibrium Characterization

In this section, I outline the key analytical results characterizing loan pricing within relationships.

5.1 Optimality Conditions

To appreciate the optimality conditions, it is useful to recall the notion of the marginal value of a dollar for the firm. This is also referred to as the shadow value of equity and corresponds to Tobin's marginal q when there are no adjustment costs of capital. Because of financial frictions (equity injection costs and bankruptcy costs), firms may value one additional dollar of funding differently: constrained firms value dollars more. The shadow value of equity, denoted by ξ, ranges from 0 to the cost of equity injections, τ. When the firm is unconstrained and distributes dividends, $\xi = 0$. If $\xi > 0$, the firm is constrained and does not distribute dividends. The value of ξ is bound above by τ, since a larger value would lead firms to immediately inject equity until sufficient funds are raised to restore $\xi = \tau$. If τ approaches infinity, raising equity becomes prohibitively costly, effectively imposing a constraint $d \geq 0$, with ξ acting as the Lagrange multiplier for this constraint.

The optimality condition for Q links the shadow value of equity ξ to another notion of financial constraint: the Lagrange multiplier on the Promise-Keeping (PK) constraint μ. These two are related as follows:

$$\mu = 1 - \xi \tag{9}$$

The above equation, though tautological when understood, connects theories of firm dynamics with financial frictions—such as those of Khan, Senga, and Thomas (2014) and Ottonello and Winberry (2020)—with bank-firm contracting models, as in Albuquerque and Hopenhayn (2004) and Kovrijnykh (2013). The Lagrange multiplier μ on the promise-keeping constraint is the slope of the Pareto frontier ($\mu = -\frac{\partial W(s,v)}{\partial v}$), and indicates the cost for the bank of providing one additional dollar of value to the firm. Constrained firms with $\xi > 0$ lie on the non-diagonal part of the Pareto frontier, meaning that it only takes $\mu < 1$ dollars from the bank to generate one dollar of value for the firm.

Choice of state-contingent promises $v'(z')$. The choice of state-contingent promises $v'(z')$ is central to the risk-sharing problem, as it summarizes the promises for future lending rates and credit quantities across all histories following the realization of state z'. How are these choices optimally made within a relationship? The optimal contract aims to keep the firm's shadow value of equity constant across time and states. Formally, the choice for $v'(z')$ solves:[18]

$$\xi'(z') = \xi - \underbrace{\eta(z')}_{\text{Firm-Mkt PC}} + \underbrace{q(z')(1-\xi'(z'))}_{\text{Bank PC}} \tag{10}$$

When neither Participation Constraint is binding, it is optimal to set $v'(z')$ such that $\xi'(z') = \xi$. If $\xi'(z') > \xi$, then the bank could improve its profits by delivering more value to the firm (e.g., through lower rates) in state z' tomorrow, when the firm's valuation of a dollar is higher, in exchange for higher rates today. However, $\xi'(z') > \xi$ could be optimal if the bank's PC binds in state z', as the bank may be unable to credibly offer better terms in state z'. Conversely, $\xi'(z') < \xi$ is optimal only if the firm's PC binds in state z', meaning that harsher conditions in that state would lead the firm to separate.

Figure 4 illustrates the choice of $v'(z')$. The left panel shows an S-shape typical of an insurance setting: firms receive a high value $v'(z')$, compared to their outside option $v^m(s')$, when productivity z' is low, and a low value when productivity is high. The magnitude of such insurance is bound above by the bank's PC, and below by the firm's PC. In the left panel, the upper PC is not flat because it is expressed in terms of bank value rather than firm value, and the "exchange rate" between bank and firm value units, μ, varies across states—which is what makes insurance valuable in the first place.

On the right panel, the same policies are expressed in terms of bank value rather than firm value, which causes the firm's PC to appear curved.

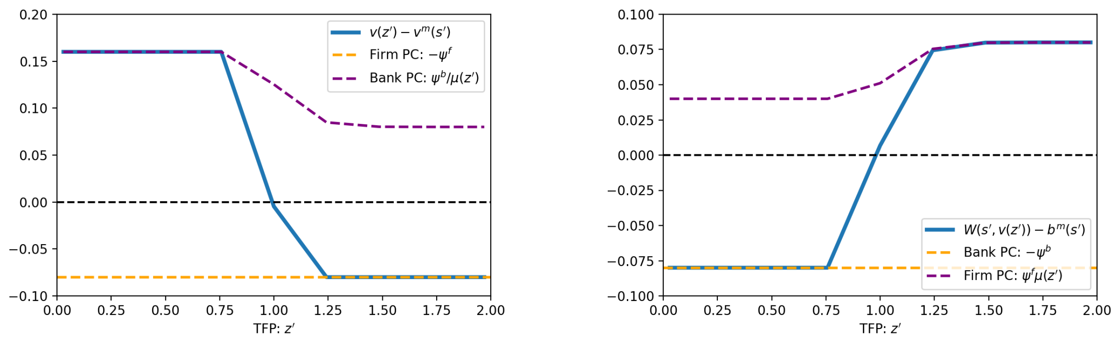


Figure 4: Visualization of the policy function for state-contingent promises, expressed in terms of firm value in the left panel and bank value in the right panel.

Choice of debt b' and capital k'. Debt b' is crucial, as it influences the outside options $v^m(s)$ and $b^m(s)$. First, in the event of exogenous separation, b' dictates how value is split between the old bank and the firm in the banking market. Second, by altering the outside

[18] Note that I am normalizing the true Lagrange multipliers to have cleaner optimality conditions. The Lagrange multiplier on the Bank PC for state z' is $\beta(1-\theta)\tau(z'|z)q(z')$. Similarly for $g(z')$ and $\eta^{sep}(z')$.

values, b' affects the set of enforceable promises $v'(z')$. High debt b' reduces $v^m(s)$, making it easier to retain the firm even for low values of $v'(z')$. Conversely, it raises $b^m(s)$, which might tempt the bank to terminate the relationship unless $W(s, v'(z'))$ is large enough. Thus, the model uncovers a new role for bank debt as a retention device.

The choice of capital k' integrates standard elements, the trade-off between investment costs and increased production, with the retention role of debt described earlier, which emerges because also capital affects the outside options. The optimality conditions for both variables are reported in Appendix B.6.

5.2 Relationships and Loan Pricing Disconnect

I now outline the core results on loan pricing within relationships. Throughout this section, I focus on the case of solvent firms,[19] and defer the treatment of insolvent firms to Section 5.4.

To understand loan pricing, it is useful to first illustrate two benchmarks. The first, $Q^{zero,w}$, is the zero-profit rate that makes the bank break-even by engaging in relationship from the current period onward:

$$Q^{zero,w} = \beta \frac{E_{z'|z}[w'(z')]}{b'}$$

where $w'(z') = W(s', v'(z'))$ for brevity. This rate fully accounts for default risk and endogenous recovery rates, and also incorporates any profits or losses the bank expects to make in subsequent interactions with the firm according to their contractual promises.

The second benchmark, $Q^{zero,b}$, is the zero-profit rate that would make a bank break-even by engaging in the relationship for the current period only, ignoring future interactions with the firm:

$$\begin{aligned} Q^{zero,b} &= \beta \frac{E_{z'|z}[b^m(s')]}{b'} \\ &= \beta \left[(1 - PD) + PD \times E_{z'|z}[b^m(s')/b' | Def] \right] \end{aligned}$$

where $PD = P(\bar{b}(s') < b')$. If the endogenous recovery rate is always zero, then $Q^{zero,b}$ simplifies to $\beta(1 - PD)$.

We can interpret $Q^{zero,b}$ as the lending rate acceptable to a secondary-market investor purchasing a single loan from a bank (e.g. a CLO investor) who will not benefit from future interactions with the firm.

The two benchmarks are inherently connected: in Appendix B.2, I show that $Q^{zero,w}$ converges into $Q^{zero,b}$ as switching costs approach zero.

[19] Firms with $b \leq \bar{b}(s)$, and thus $b^m(s) = b$.

Next, I leverage these benchmarks to characterize loan pricing within a relationship. The following propositions assume that the exogenous separation rate θ is zero. This assumption is not essential, but simplifies the results. I provide the full expression for the general case in Appendix.

Proposition 1. *Consider a firm in state $s = (z, b, k)$ that is solvent ($\bar{b}(s) \geq b$). The equilibrium loan rate is given by:*

$$Q = \underbrace{\beta \frac{E_{z'|z}[w'(z')]}{b'}}_{Q^{zero,w}} + \underbrace{\frac{b^m(s) - w}{b'}}_{\chi:\ subsidy} \tag{11}$$

where $w'(z') = W(s', v'(z'))$ and $w = W(s, v)$ for brevity.

Proof: *See Appendix B.1.*

The pricing equation in Proposition 1 summarizes loan pricing within a relationship. The rate $Q^{zero,w}$ would make the bank break-even, accounting for future interactions with the firm. Deviations from this benchmark can occur due to past promises. For instance, if the bank committed to favorable terms for the firm, accepting a value w lower than what its outside option $b^m(s)$,[20] then we would obtain $Q > Q^{zero,w}$.

Equation (27) offers a concise representation of the loan pricing disconnect. However, when considering the intertemporal dimension of risk-sharing along the relationship tenure, it is useful to further decompose Q into $Q^{zero,b}$ and $Q^{zero,w}$, as shown in the following Corollary.

Corollary 1: *The equilibrium loan rate can also be expressed as:*

$$Q = \underbrace{\beta \frac{E_{z'|z}[b^m(s')]}{b'}}_{Q^{zero,b}} + \underbrace{\beta \frac{E_{z'|z}[w'(z') - b^m(s')]}{b'}}_{Q^{zero,w} - Q^{zero,b}:\ intertemporal\ subsidy} + \underbrace{\frac{b^m(s) - w}{b'}}_{\chi:\ subsidy} \tag{12}$$

The *intertemporal subsidy* in Equation (12) captures deviations from the zero-profit condition on today's loan, in either direction, that the bank expects to make up for in subsequent interactions with the firm. For instance, in the quantitative results, I will show that banks provide cheap loans early in the relationship, when the firm is more constrained, in exchange for higher rates later. When focusing on Equation (27) alone, we would miss this result because, due to perfect competition to start new relationships, the first loan in a relationship is always given at $Q = Q^{zero,w}$. However, when examining individual loans through Equation 12, we can see that even in the first loan it is possible that $Q \neq Q^{zero,b}$, as the bank has to break-even over the course of the relationship, not just on the first loan.

[20] Here, since we are focusing on solvent firms, we have $b^m(s) = b$: the outside option for the bank is simply to enforce the repayment of b.

5.3 Competitive Lenders Limit

In this section, I analyze the model in the extreme case of zero switching cost ($\psi^{f,b} \to 0$). Proposition 2 shows that, in this scenario, the optimal contract converges into the simpler case of competitive credit markets, and shocks are fully passed-through to loan rates. Intuitively, any promise on future credit conditions other than what would prevail in a competitive market becomes unenforceable. The proof concept relies on showing that zero switching costs imply $v(z') = v^m(s')$. As a result, all the flexibility coming from state-contingent promises summarized by $v(z')$ is lost, and the values in each state are determined directly from the choice of the non-contingent debt b', together with z' and k'.

Proposition 2. *Suppose that switching costs approach zero ($\psi^{f,b} \to 0$). Consider a firm in state $s = (z,b,k)$ that is solvent ($\overline{b}(s) \geq b$). Then, the optimal contract prescribes that loans are priced competitively:*

$$Q = \underbrace{\beta \frac{E_{z'|z}[b^m(s')]}{b'}}_{Q^{zero,b}}$$

That is,

$$Q = \beta \left[(1-PD) + PD \times E_{z'|z}[\overline{b}(s')/b|Def] \right] \tag{13}$$

where $PD = P(\overline{b}(s') < b')$.
Furthermore, the allocations are identical to those in an economy where firms face lenders offering competitive pricing schedules each period, solving the following problem:

$$V_P(s) = \max_{b',k',d} \quad d + \beta E_{z'|z}[V(s')]$$

Subject to:

$$Qb' = \beta E_{z'|z}[b^m(s')] \qquad (\mu^{-1}\text{: PK-Bank})$$

and the usual budget constraint.

Proof: *See Appendix B.2.*

Proposition 2 shows that as switching costs approach zero, the contracting framework described in Section 4 becomes a redundant description of a competitive credit market, like the one in Cooley and Quadrini (2001). This convergence applies both to loan pricing and equilibrium allocations. Furthermore, Proposition 3 highlights that the zero-profit loan schedule in that class of models can be viewed as the limiting case of a promise-keeping constraint in a long-term contract with outside options.[21]

[21] To see this more clearly, in Proposition 3 I have outlined the dual formulation of the problem, where the firm maximizes its value subject to a Bank PK constraint, rather than the opposite.

So far, I have focused on solvent firms. The results in Propositions 1 and 2 extend to insolvent firms when insolvency is resolved through restructuring of the legacy debt. In the next subsection, I will discuss an alternative hypothesis in which insolvency is resolved by the bank through the issuance of a new subsidized loan. This will clarify the distinction between *insurance* and *evergreening*.

5.4 Insolvent Firms and Evergreening

In this section, I examine the case of insolvent firms, where $b < \bar{b}(s)$. When a firm is insolvent, exit is not always optimal, as the lender may prefer to keep the firm operational in expectation of future repayments. In such cases, we must specify how insolvency is resolved. For my main quantitative results, I assume that insolvency is addressed by reducing debt from b to $b^m(s)$, as in Cooley and Quadrini (2001). Under this assumption, the theoretical results from Sections 5.2 and 5.3 extend to insolvent firms. Alternatively, the bank could demand the full repayment of b, but simultaneously offer a new subsidized loan, thus engaging in *evergreening*. In my framework, the choice between restructuring and *evergreening* is merely an accounting one, and does not affect allocations.[22] The notion of *evergreening* as implicit debt restructuring is recently explored in Martin, Mayordomo, and Vanasco (2023) and Faria-e Castro, Paul, and Sánchez (2024).

If we assume that insolvency is resolved through loan evergreening, a disconnect from competitive loan pricing can emerge even in the absence of switching costs, but it would be confined to insolvent firms. This result is detailed in the next proposition.

Proposition 3. *Suppose that switching costs are zero ($\psi^{f,b} \to 0$). Consider a firm in state $s = (z,b,k)$ that is insolvent ($\bar{b}(s) < b$). Then, under loan evergreening, the optimal contract prescribes that the firm receives a subsidized loan:*

$$Q = \underbrace{\beta \frac{E_{z'|z}[b^m(s')]}{b'}}_{Q^{zero,b}} + \underbrace{\frac{b - b^m(s)}{b'}}_{Evergreening} \tag{14}$$

Furthermore, the loan subsidy equals the capital loss the bank would incur if selling the loan on the secondary market:

$$\underbrace{Qb' - \beta E_{z'|z}[b^m(s')]}_{Loan\ Subsidy} = \underbrace{(b - b^m(s))}_{Capital\ Loss} \tag{15}$$

[22] This is also the case in Martin, Mayordomo, and Vanasco (2023), as the firm's size is fixed. Instead, in Faria-e Castro, Paul, and Sánchez (2024), *evergreening* can cause distortions because the firm can borrow any amount at the posted Q. This does not happen in my model, as I allow banks and firm to contract on both rates and quantities.

Proof: *See Appendix B.3.*

Proposition 3 clarifies the conceptual difference between *insurance* and *evergreening*. In Proposition 1, a disconnect from competitive pricing arises as an *insurance* mechanism when $w \neq b^m(s)$, where $w = W(s,v)$ is the value the bank obtains by honoring past promises, and $b^m(s)$ is its outside option. This difference benefits one party, either the bank or the firm, but harms the other. Thus, it can only be sustained with positive switching costs, which enforce the commitment to the original terms. In contrast, Proposition 3 shows that *evergreening* occurs when $b \neq b^m(s)$, and specifically when $b > b^m(s)$, indicating insolvency. Evergreening does not require imperfect mobility, as providing a subsidized loan is in the bank's interest to avoid the firm's exit.

In Appendix B.8, I show that when insolvencies are resolved through loan evergreening, the model displays limited pass-through from default risk to loan rates but minimal history-dependence, which contradicts empirical evidence. Positive switching costs, combined with the *insurance* mechanism, are needed to deliver large history-dependence. I examine the quantitative performance of the model in the next section.

6 Quantitative Results

In this section, I illustrate the model's calibration and evaluate its ability to explain the empirical loan pricing patterns documented in Section 3. I then derive and empirically validate new predictions about when firms should receive cheap loans.

6.1 Calibration

I divide the parameters into two groups: fixed, which I either calibrate from the literature (household parameters) or externally estimate using Mexican Orbis data (firm parameters), and internally calibrated, which I estimate to match key moments from the data through the lens of the model.

6.1.1 Fixed Parameters

The fixed parameters are reported in Table 11. I calibrate the economy to the annual frequency, aligning with the firm-level balance sheet from both Orbis and the credit registry. This also approximates well the average loan maturity in the data. I set the discount factor to $\beta = 0.96$, corresponding to a real rate of 4 percent in the steady state.

For the firm-specific parameters, I set $\alpha = 0.25$ and $\nu = 0.5$, which yield a labor share of two-thirds and total returns to scale of 0.75, consistent with Ottonello and Winberry (2020), Jeenas (2018), and Cooper and Ejarque (2001). I set the depreciation rate to $\delta = 0.17$, matching the average investment rate of firms in Mexico, and the separation rate at $\theta = 0.1$, matching its empirical counterpart.[23] I estimate the TFP process externally, using Orbis data from Mexico, as detailed in Appendix B.9.2.

Parameter	Description	Value	Target
Household			
β	Discount factor	0.96	Real rate 4%
σ	EIS inverse	2	Macro literature
ϕ_L	Labor Disutility	1.57	Employment rate 0.6
Firm			
α	Capital coefficient	0.25	Capital share one-third
ν	Labor coefficient	0.5	Returns to scale 0.75
δ	Depreciation	0.17	Orbis mean investment Rate
θ	Exogenous separation	0.1	Separation rate in R04
TFP Process			
ρ_z	Persistence	0.881	Estimated in Orbis
σ_z	Volatility	0.152	Estimated in Orbis

Table 11: Fixed and externally estimated parameters.

6.1.2 Internally Estimated Parameters

I estimate the remaining parameters to match several moments from Mexican micro-data, which are reported in Table 12. The key moment is the pass-through of default risk to loan rates within relationships. I choose to target a value of 0.35 for this pass-through, at the conservative end of the estimates from Section 3. Before returning to the pass-through, I briefly review the other targeted moments.

Targeted moments. Most of the moments I estimate using Mexican microdata are standard in the firm dynamics literature. One exception is the elasticity of default risk to sales. This elasticity is crucial as it captures when firms are constrained: a positive value would suggest that firms are more constrained when their sales increase, as they need to borrow to expand. Instead, the negative empirical estimate suggests they are more constrained as sales decline, as their financial resources are depleted. Introducing capital adjustment costs is pivotal to match this moment, as they make it costly for firms to downsize when their sales decline, preventing them from deleveraging, and suboptimal to expand too quickly as sales grow. Matching this moment is important for loan pricing in the model, because firms that are more constrained tend to receive cheap loans from the bank.[24]

[23] Notice that, as shown in Appendix B.2, this parameter is irrelevant for loan pricing in the limit when switching costs are zero. The key parameters affecting loan pricing are the switching costs $\psi^{f,b}$.

[24] Failing to match this moment would imply that firms are more constrained when their sales increase, and thus banks would try to help these firms. It would also imply that banks would potentially raise the

Average default risk is high in the data, and it is well known that AR(1) TFP shocks alone often fail to match this moment. However, incorporating fixed operating costs and capital adjustment costs significantly raises average default risk without needing to add new shocks.[25] The TFP disaster causing firm productivity to drop to zero permanently is akin to an exogenous exit shock, but is represented more compactly as an endogenous choice.[26] It serves a dual role. First, it ensures the stationarity of the model. Second, it also increases default risk, even though it does not always lead to default, as firms with little debt are able to satisfy their creditor by liquidating capital.

Moment	Description	Data	Model
Loan Pricing			
$PT(\Delta R, \Delta PD)$	Risk pass-through stayers	0.35	0.39
Firm Constraints			
$E_i[PD]$	Mean default probability	0.08	0.09
$E_i[D/\pi]$	Mean payout ratio	0.26	0.37
$E_i[b/k]$	Median leverage	0.48	0.31
Sales			
$\varepsilon(\Delta PD, g_s)$	PD-sales elasticity	-0.07	-0.06
$E_i[v/(k-b)]$	Median market-to-book Ratio	1.51	2.31
Exit rate	Firm exit rate	0.085	0.087
Investment			
$corr(I/K, I/K)$	Investment autocorrelation	0.11	0.17
$Frac(I>0)$	Fraction of positive investments	0.75	0.61

Table 12: Matched moments. The parameters in Table 13 are estimated to match the empirical moments in the Data column. $PT(\Delta R, \Delta PD)$ is the pass-through from default risk to loan spreads.

Switching costs and pass-through. The key parameters I aim to estimate are the switching costs, which I assume to be equal in my estimation ($\psi^f = \psi^b$). These are identified by the pass-through of default risk to loan pricing. The identification strategy is illustrated in Figure 5, which shows the pass-through of default risk for stayers and switchers in the model. In the figure, I hold all parameters at their estimated values while varying the switching costs $\psi^{f,b}$. On the left side of the plot, as switching costs approach zero, full pass-through emerges and there is no difference between stayers and switchers, representing a competitive spot market for loans. In this scenario, the pass-through is actually slightly

pass-through of monetary policy shocks above one, since after a policy hike firms could simply downsize, deleverage and become less constrained.

[25] For instance, Ottonello and Winberry (2020) add capital quality shocks. Adding new shocks is computationally much more expensive than in typical heterogeneous firms models: since agents write contracts contingent on the realization of each shock, new shocks raise the number of choice variables exponentially. Therefore, generating default risk through other mechanisms is preferable.

[26] While exit in the model is always endogenous, all firms that receive the TFP disaster shock choose to exit in my calibration.

below 1, as it is proportional to β, which is 0.96 in my calibration.

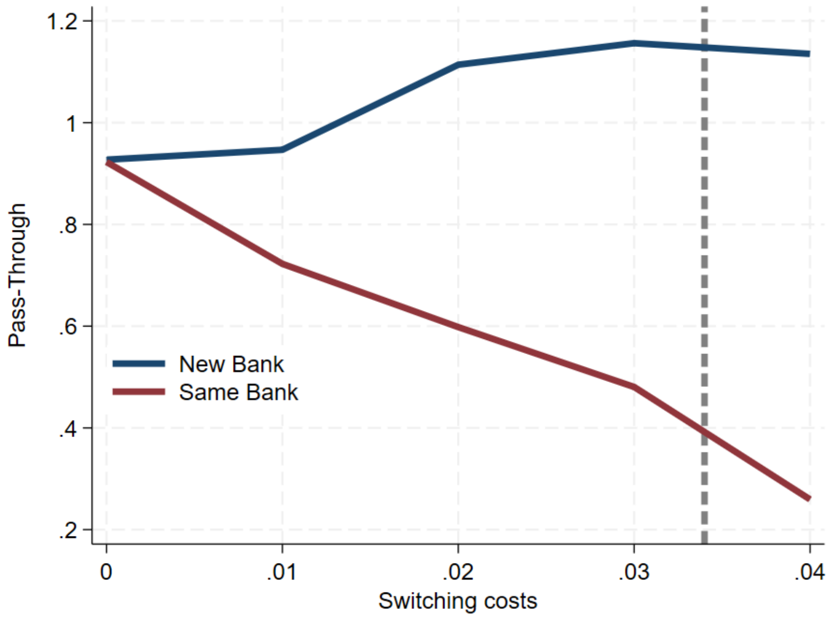


Figure 5: This figure shows how switching costs affect the pass-through of default risk to loan spreads, differentially for stayers and switchers. The pass-through is computed on simulated data as the regression coefficient of $\Delta Spread_{ft} = \alpha + \beta \Delta PD^w + \varepsilon_{ft}$, where default risk is defined as $PD^w = \beta - Q^{zero,w}$.

As switching costs increase, the *loan pricing disconnect* emerges, and loan pass-through declines sharply for firms that stay with the same bank. To formalize the patterns in Figure 5, we can rearrange Equation (27) as follows:

$$\Delta Spread = \Delta PD^w - \Delta \chi \tag{16}$$

where χ is the subsidy term defined in Proposition 1, and $PD^w = (\beta - Q^{zero,w})$ is the zero-profit rate accounting for default risk. The limited pass-through can be understood as an omitted variable problem: by regressing $\Delta Spread$ on ΔPD^w, we are ignoring the subsidy term χ, which might covary with the other terms in equilibrium.

In the next proposition, I leverage this insight to derive an analytical expression for the pass-through.

Proposition 4: *Suppose that $\beta \to 1$. The pass-through of default risk to loan rates for stayers, obtained from regressing $\Delta Spread_{ft}$ on ΔPD^w_{ft}, is given by:*

$$PT^{stayers} = 1 - \frac{Cov(\Delta PD^w_{ft}, \chi_{ft})}{Var(\Delta \chi_{ft})} + \frac{Cov(\Delta PD^w_{ft}, \chi_{f,t-1})}{Var(\Delta \chi_{ft})} \tag{17}$$

where $\chi = \frac{b^m(s) - w}{b'}$ is the subsidy defined in Proposition 1, $PD^w = (\beta - Q^{zero,w})$ is the zero-profit loan spread accounting for default risk, and $Spread = (\beta - Q)$ is the loan spread

actually charged by the bank.
The pass-through of default risk to loan rates for switchers is:

$$PT^{switchers} = 1 + \frac{Cov(\Delta PD^{w}_{ft}, \chi_{f,t-1})}{Var(\chi_{f,t-1})} \tag{18}$$

Banks allocate subsidies to more constrained firms that have a high marginal valuation of a dollar. Since firms with increased default risk are typically more constrained, the covariance $Cov(\Delta PD^{w}_{ft}, \chi_{ft})$ in Equation (17) is positive, which reduces the pass-through of default risk to loan spreads. When switching, new loan is priced competitively ($Q = Q^{zero,w}$) but the previous loan was not. Therefore, pass-through for switchers can differ from one if the old subsidy term $\chi_{f,t-1}$ covaries with ΔPD^{w}_{ft}. Quantitatively, this term tends to be small and positive, pushing the pass-through for switchers slightly above one.

Parameter	Description	Value
Switching Costs		
$\psi^{f,b}$	Switching Costs	0.034
Production		
h_C	Convex Adjustment Cost	0.85
f	Fixed Production Cost	0.037
π_0	TFP disaster	0.08
Financial Frictions		
τ	Equity Injection Cost	0.5
$1-\gamma$	Bankruptcy Cost	0.125

Table 13: Estimated Firm parameters. The parameters in this Table are chosen to match the moments reported in Table 12.

6.2 History-Dependence and Pricing Reconnect

I now evaluate the model's performance in generating the two additional pricing patterns documented in Section 3: history-dependence and the role of switching bank in closing the prior loan pricing wedge.

History-Dependence. In the data, past risk assessments significantly predict loan pricing, even after accounting for current default risk. I replicate this analysis in the model by regressing the spread charged by the bank on both the contemporaneous and the lagged default risk four periods in the past, matching the median relationship tenure. While the model does not display as much history-dependence as observed in the data, it succeeds in generating a significant role for past risk. This success is even starker when comparing the model's performance with perfect mobility,[27] displayed in the last column of Table 14.

[27]The coefficient on PD_{ft} in the model with zero switching costs is slightly below one, reflecting the level

	Data	Model	Model
$\psi^{f,b}$		0.034	0.00
	$Spread_{ft}$	$Spread_{ft}$	$Spread_{ft}$
PD_{ft}	0.258^{***}	0.571^{***}	0.883^{***}
	(0.068)	(0.009)	(0.001)
PD_{f0}	0.750^{***}	0.327^{***}	0.000
	(0.057)	(0.008)	(0.000)

Table 14: History-dependence in the data (first column) and in the model (last two columns). The second column contains the results from the model with positive switching costs, as estimated in Section 6.1.2. The last column presents results for the model calibration with zero switching costs. Robust standard errors in parenthesis.

Correcting Past Mispricings. In the data, switching to a new bank is associated with a significant reversal of past mispricing, defined as $Mispr_{ft} = PD_{ft} - Spread_{ft}$. Table 15 shows that the model successfully replicates this pattern, along with a more muted correction within relationships. Similar to the history-dependence results, the model falls short of capturing as much persistence as in the data within relationships, but it does explain a substantial portion.

	Data	Model
	$Spread_{ft}$	$Spread_{ft}$
$R^{wedge}_{f,t-1}$	-0.936^{***}	-1.039^{***}
	(0.009)	(0.006)
$R^{wedge}_{f,t-1} \times \mathbb{1}_{SameBank}$	0.766^{***}	0.252^{***}
	(0.010)	(0.006)

Table 15: The pricing wedge is defined as $R^{wedge}_{ft} = Spread_{ft} - PD_{ft}$, which captures the difference between the actual loan rate and the one implied by competitive pricing.

6.3 Who Receives Cheap Loans?

In the model, banks offer cheap loans to firms when their marginal value of a dollar is high, aiming to reduce inefficient fluctuations in the value of funding. Since the marginal value of a dollar is not directly observable, testing the model's mechanism is challenging. However, we can assess whether the model captures the covariance between loan subsidies and other observable firm characteristics.

of $\beta < 1$ and possible numerical imprecisions that would bias the estimate downward.

Cheap loans and firm shocks. Dividend distributions are the most direct empirical proxy for the marginal value of a dollar, since in the model constrained firms with $\xi > 0$ do not distribute dividends, which is consistent with evidence in Kaplan and Zingales (1997). Additionally, I analyze how loan subsidies relate to sales growth and investment rates. To this end, I compute the pricing residual as $Res_{ft} = \Delta Spread_{it} - \Delta PD_{ft}$, which corresponds to the term $-\Delta\chi_{ft}$ in Equation 16, and evaluate how this term covaries with firm observables.

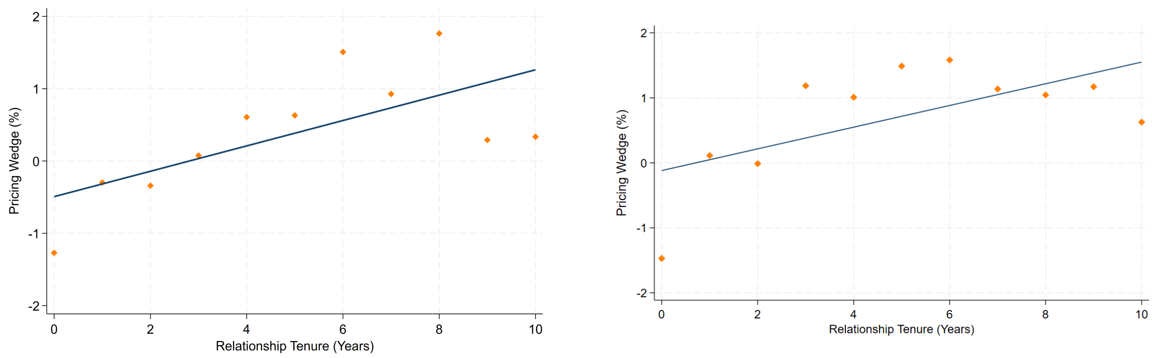
	Data	Model
Payout ratio	2.844**	1.317***
	(2.67)	(76.52)
Sales growth	3.993***	1.694***
	(12.12)	(191.00)

Table 16: Comovement of the residual pricing term R_{ft}^{wedge} with some key firm observables. The dividend payout ratio is the ratio of dividends to earnings, sales growth is the log growth in sales compared to the previous year.

As predicted by the model, firms that distribute dividends face higher lending rates, as they are unconstrained, making it optimal to charge them more. Similarly, positive sales shocks are associated with tougher conditions. Matching the negative elasticity of default risk to sales, as documented in Section 6.1.2, is crucial for the success of the model to capture this pattern: firms with declining sales become more risky and constrained, which makes them more likely to receive loan subsidies. In the data, I also observe a weak negative relation between investment and loan rates. The model replicates this pattern and reveals a U-shaped relationship due to capital adjustment costs: firms investing heavily receive subsidies, as do those significantly reducing their capital, that are likely undergoing costly downsizing.

Cheap loans by match tenure. I now turn to the model predictions for loan pricing along the tenure profile. Figure 6 shows that in the model banks deliver cheap loans early on in their relationship with the firm, anticipating the ability to charge higher markups later. This form of *intertemporal subsidy* is optimal because firms are typically more constrained when they are young, while financial frictions lessen as firms accumulate retained earnings.

The result in Figure 6 aligns with the model's mechanism but contrasts with the literature on optimal contracts, which typically emphasises *backloading*: payments to the firm are deferred, while payments to the bank are frontloaded. In my model, there are two notions of backloading. The first, common in the literature, pertains to cash flows. In this sense, the model aligns with the literature, as dividends are *backloaded* and cash flows initially accrue to the bank. The second notion, unique to my model, emerges from comparing loan rates to break-even rates, as in Figure 6. According to this notion, loan rates are *frontloaded*:



Data

Model

Figure 6: Pricing wedge $R^{wedge}_{ft} = (Spread_{ft} - PD_{ft})$ by match tenure (years since the beginning of the banking relationship).

the firm benefits early in the relationship, while high loan rate payments to the bank are deferred.

6.4 When Are Firms Tempted to Switch?

I now study when firms are tempted to switch to a new bank in the model. Notice that the optimal contract is designed to prevent switching,[28] but studying when firms are tempted to switch—that is, when the firm participation constraint is binding—is still informative to understand the switching decision.

Figure 7 shows that in the model firms are tempted to switch when their default risk has declined since the onset of their banking relationship. Intuitively, the insurance mechanism is penalizing these firms, which may find it desirable to switch and see their improved risk profile better reflected in their borrowing conditions. The data confirm this model prediction.

[28]One possibility to achieve endogenous separation *on-path* would be to introduce some trembling, which would require the firm switching cost ψ^f to be both random and privately observed, which makes it non-contractible, an approach used in Hopenhayn and Werning (2008) and Müller, Storesletten, and Zilibotti (2019) to obtain default on-path.

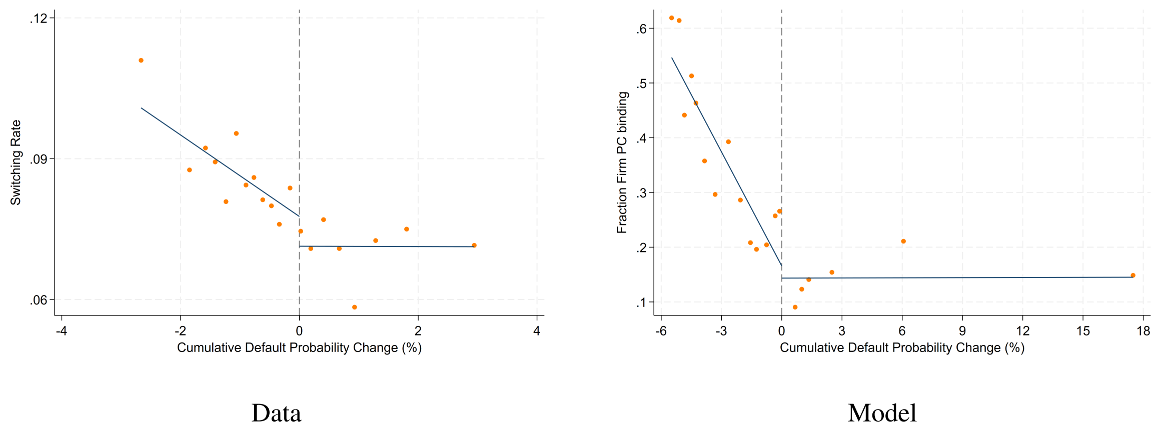


Figure 7: Left plot: empirical switching rate by cumulative default probability change since onset of the relationship ($PD_{t-1} - PD_0$). The switching rate is residualized against firm-level characteristics (outstanding credit, firm age). Right plot: fraction of firms tempted to switch (binding firm PC) by cumulative default probability change since onset of the relationship.

6.5 Discussion of Alternative Theories

The model, which formalizes relationships as optimal long-term contracts with risk-sharing, rationalizes the main empirical result of Section 3: limited pass-through for stayers, history-dependence, and pricing reconnect upon switching switching, and correctly predicts when firms should receive cheap loans and when they should be tempted to switch.

I now present and assess potential alternative explanations for the empirical findings. While I do not aim to dismiss the existence of other mechanisms, I highlight that a risk-sharing explanation is more likely to fully account for the empirical pricing patterns I have documented.

Evergreening. A prominent theory which has received increasing attention after the financial crisis is that banks may evergreen loans, providing cheap credit to ”zombie firms” (Caballero, Hoshi, and Kashyap (2008)). My model provides, to my knowledge, the first conceptual distinction of *insurance* and *evergreening*, which is formalized in Section 5.4. *Insurance* emerges within relationships because banks provide cheap credit to constrained firms to improve allocative efficiency, while *evergreening* arises when firms are insolvent and banks continue lending at subsidized rates to prevent firms’ bankruptcy.

Leveraging this formalization of *insurance* and *evergreening*, I can assess their relative performance in explaining the data. In the model, *evergreening* delivers limited pass-through but shows minimal history-dependence, which contradicts the data. Indeed, papers that have studied evergreening have looked at loan pricing in the cross-section (Artavanis et al. (2022), Faria-e Castro, Paul, and Sánchez (2024)), documenting that firms with low profits or high risk often obtain cheap credit. Instead, I look at loan pricing over time for the same firm within relationships, and find evidence not only of limited pass-through, but also of strong history dependence. Furthermore, in the data, I find a muted pass-through of default risk even when firms do not have loans maturing with their bank in the current or the three subsequent months, which are cases in which evergreening concerns would be milder, as reported in Appendix B.8. Finally, it is to be noticed that around 90% of firms keep borrowing from their old bank: attributing the extent of limited pass-through exclusively to zombie firms may overstate their importance.

Aside from evergreening, there are three other theories whose relation to my empirical results is worth discussing: market power, information asymmetries, and anchoring.

Market Power is often associated in economics with limited pass-through,[29] but traditional market power frameworks do not deliver history-dependence. Implicit long-term contracts, on the other hand, are a specific way to model imperfect competition, common in the labor literature, which can deliver both limited pass-through and history-dependence.

Information asymmetries are potentially important in banking. Because banks themselves report the risk assessments, clearly they are aware of the change in firm risk, so bank ignorance cannot explain limited pass-through. However, the concern could be that limited pass-through emerges because the assessment is noisy and banks rely on superior soft information, a point supported in Demiroglu, James, and Velioglu (2022). In my data, I highlight five empirical observations which are at odds with the ability of this mechanism to explain my data. First, under this theory, we would expect spreads to have a superior explanatory power in predicting loan losses. Instead, in Appendix A.1, I show that risk assessments seem to be more important for forecasting defaults. Second, an information-based explanation would be incompatible with history-dependence, unless past assessments have predictive power for defaults beyond current assessments, which I rule out in Appendix A.1. Third, in Appendix A.2.2, I show that lending rates are also unresponsive to changes in risk assessments due exclusively to the bank's soft information, as reported in the credit registry. Fourth, when firms switch to a new bank, a large fraction of the previous mispricing is closed, as shown in Section 3.4, which directly invalidates the soft information story if we believe these assessments to be accurate. If the pricing disconnect reflected the superior private information of the old bank, we would not expect new banks to correct it. Indeed, new banks can observe firms' borrowing history through the credit registry, including their borrowing rates, and could use this information for their loan pricing. Fifth, full pass-through emerges only as firms are switching, and immediately after the new bank starts pricing with relational patterns. This could only be compatible with information asymmetries if the new bank learns about the firm very quickly.

Anchoring. Finally, an alternative explanation is that banks may simply stick to previous loan terms to avoid renegotiating. This notion of *anchoring* has been introduced in Dougal

[29]Loan market institutions, however, are very different from canonical monopolistic competition frameworks, for instance because quantities and prices are typically contracted together, and because the source of the shock is specific to the firm, not to the bank.

et al. (2015), and is compelling because it is the only alternative explanation compatible with both limited pass-through and history dependence. Some behavioral anchoring is probably present in the data, since around four percent of consecutive loans pairs have exactly the same spread and almost twenty percent have the same loan rate. However, loan spreads are ultimately far from anchored in my data, as the volatility of loan spreads is almost twice that of default risk: they are simply disconnected from firm risk. This is compatible with my firm dynamics setup, where *insurance* does not necessarily require flat lending rates, but rather allocating markups optimally across states, in a way which can be empirically validated. This ensures firms receive discounts from their bank when the value of having additional dollars to invest in the firm is highest, which does not necessarily translates into perfectly anchored loan rates.

7 Macroeconomic Implications

In this section, I study the aggregate implications of relationship lending. The model provides a clear framework for addressing this issue, since changing a single parameter—the switching cost—shifts the economy from a competitive to a relational equilibrium.

In Section 7.1, I quantify the steady state welfare gains from relationship lending. Next, in Section 7.2, I cast the model in a New Keynesian framework to study how relationships affect the transmission of monetary and fiscal policy.

7.1 Steady state

To evaluate the role of banking relationships, I solve for the steady state under three calibrations, as shown in Table 17. I fix all parameters—except the switching costs—at their calibrated values from Section 6, varying only the switching costs. First, I consider an economy with competitive spot credit markets ($\psi^{f,b} = 0$). Second, the calibrated model with relationship lending and positive switching costs ($\psi^{f,b} = 0.034$). Finally, a model with perfectly complete capital markets, which serves as a benchmark to quantify the potential gains from removing financial frictions. Perfect capital markets can be achieved by either setting switching costs to infinity—which is equivalent to imposing full commitment on the set of Arrow-Debreu securities—or by eliminating one of the two financial frictions ($\tau = 0$ to allow firms to raise equity, or $\gamma = 1$ to eliminate bankruptcy costs in debt markets).

I find that relationships reduce misallocation in the steady state relative to competitive markets. To see this, consider the shadow value of equity ξ, which measures the marginal value of one dollar invested in a firm. As the economy transitions to a relational equilibrium, the average firm becomes less constrained. Furthermore, the dispersion across firms in the value of a dollar declines sharply. The variance of ξ captures the potential gains in allocation that a planner could achieve by redistributing resources across firms. Banks in the model actively engage in this redistribution, realizing part of the potential allocative gains available in competitive markets.

This improvement in allocation leads to increased consumption, capital, and output. In the relational market, hours worked rise, whereas in the complete market economy they fall sharply due to higher equilibrium wages. These results are summarized by comparing the welfare of the representative household in equilibrium. In the economy with complete markets, consumption-equivalent welfare is 2.51% higher than in the spot market economy, which quantifies the welfare losses due to financial frictions. The calibrated model with relationships accounts for over 10 percent of these gains, as welfare increases by 0.29% compared to the model with spot markets. Therefore, higher switching costs raise equilibrium welfare in my economy.

	Competitive	**Relationships**	**Complete Markets**
Commitment	No	Partial	Full
$\psi^{f,b}$	0.0	0.034	∞
Allocation			
ξ	0.088	0.079	0.0
		(-0.9 ppt)	
$Var(\xi)$	0.026	0.025	0.0
		(-4.0%)	
Aggregates			
Capital	0.461	0.469	0.472
		(+1.78%)	(+2.40%)
Consumption	0.570	0.578	0.583
		(+1.40%)	(+2.28%)
Labor	0.589	0.600	0.584
		(+1.86%)	(-0.85%)
Welfare (CE)	0.00	+0.29%	+2.51%

Table 17: This table shows how key macroeconomic aggregates respond to changes in the structure of financial markets. ξ is the shadow value of equity. $Var(\xi)$ is the cross-sectional dispersion in the shadow value of equity.

7.2 Macroeconomic Policy

In this section, I embed the model in a New Keynesian framework to study how banking relationships affect the economy's response to monetary and fiscal policy. This requires some extensions to the setup described in Section 4, which I outline below.

7.2.1 Model extension with aggregate shocks

The extension of the optimal contract problem with aggregate shocks and inflation is described in Appendix B.4. The main difference lies in the fact that contracts must now condition also on aggregate shocks.[30]

New Keynesian Block. The New Keynesian block of the economy is standard and designed following Ottonello and Winberry (2020) to interfere minimally with the rest of the economy in steady state.

The heterogeneous firms outlined in Section 4 combine labor and the final good to produce an intermediate good, which they sell competitively at price p_t to a fringe of monopolistically competitive producers. These transform the intermediate good with a linear production technology $\tilde{y}_{jt} = y_{jt}$ and sell their output to a final good producer, which uses the production function $Y_t = \left(\int \tilde{y}_{jt}^{1-\gamma}\right)^{\frac{1}{1-\gamma}}$. The monopolistically competitive producers set prices for their output $\tilde{p}_{jt}$, but incur quadratic adjustment costs when changing prices: $\frac{\phi_{NK}}{2}\left(\frac{\tilde{p}_{jt}}{\tilde{p}_{j,t-1}} - 1\right)^2$.

The advantage of this formulation is that it does not directly interferes with the already complex heterogeneous firm block, but it generates a New-Keynesian Phillips curve:

$$\log \Pi_t = \frac{\gamma - 1}{\phi_{NK}} log \frac{p_t}{p^*} + \beta E_t log \Pi_{t+1} \tag{19}$$

where $p^* = \frac{\gamma-1}{\gamma}$ is the steady state level of p_t, the price of the intermediate input produced by the heterogeneous firms.

The central bank sets nominal interest rates according to the following Taylor rule: $\log R_t^{nom} = \log(\frac{1}{\beta}) + \phi_\pi \log \Pi_t + \varepsilon_t^m$, where $\varepsilon_t^m \sim N(0, \sigma_m^2)$ is the monetary policy shock.

Equilibrium with aggregate shocks. An equilibrium consists of a set of value functions $W_t(s,v)$ and $W_t^P(s,v)$; decision rules for capital $k_t'(s,v)$, debt $b_t'(s,v)$, dividends $d_t(s,v)$, labor $l_t(s,v)$, and contractual promises contingent on the idiosyncratic and aggregate state $v_t'(s,v;z',\varepsilon')$); a measure of bank-firm matches $\mu_t(s,v)$; and prices w_t, p_t, Π_t, Λ_{t+1} such that (i) all contracts are written optimally, (ii) the household optimizes, (iii) the evolution of the distribution is consistent with decision rules, and (iv) the labor, asset, and good markets clear.

In the absence of aggregate shocks, the problem reduces to the steady state defined in Section 4.2, and the price sequence simplifies to $p_t = \frac{\gamma-1}{\gamma}$, $\Pi_t = 1$, $\Lambda_{t+1} = \beta$. Consequently, the New-Keynesian structure becomes irrelevant, and the time dependence in all value and policy functions can be dropped.

[30] Notice that aggregate shocks change outside options, but idiosyncratic shocks do as well since firms bring their productivity to a new bank when they switch. This is different from recent models of relationships in the labor market (Souchier (2022)), where only aggregate shocks, and not firm shocks, affect workers outside options. This would be the case in my setting if there was a bank-specific shock.

Solution method. I solve the problem assuming that aggregate shocks occur with a near-zero probability, which allows agents to write contracts contingent on the shock's arrival, but preserves computational tractability. I study the transition of the economy back to steady state after receiving an aggregate shock. The solution method, which I describe in detail in Appendix B.6.3, extends the sequence space approach in Boppart, Krusell, and Mitman (2018), and in Auclert et al. (2021), to incorporate state-contingent contracts.

Calibration of the New Keynesian Block. I calibrate the New Keynesian parameters using standard values from Kaplan, Moll, and Violante (2018) and Ottonello and Winberry (2020). I set $\gamma = 10$, corresponding to an 11 percent markup in the steady state, the price rigidity parameter to $\psi = 100$, implying a slope of the Phillips Curve of 0.1, and the Taylor rule coefficient to $\phi_\pi = 1.25$.

7.2.2 Response to a monetary policy shock

I study the response to a 25 basis points monetary policy hike. Because long-term contracts are contingent on the monetary policy shock, banks and firms negotiate credit conditions following the rate hike in advance, potentially insulating firms from the shock. As for the steady-state, I present the results for three economies: the competitive spot market with zero switching costs, the relational economy with estimated switching costs, and the economy with complete markets.

Figure 8 shows the response of two key financial variables: equity injections and the fraction of firms in default. In the case of complete markets, both remain at zero. In the other two economies, they differ sharply. As the rate hike makes firms more constrained, in the competitive economy they rush to raise costly equity; in contrast, relationship banks provide cheap funding after the shock, reducing the need for costly equity injections and limiting the surge in defaults. Notably, the model exhibits a hump-shaped response of defaults. This indicates that rate hikes, rather than immediately pushing firms into default, increase their exposure to subsequent idiosyncratic shocks.

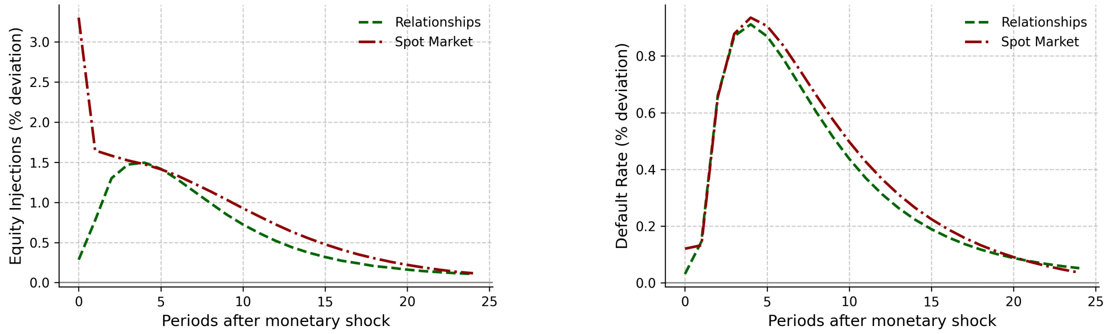


Figure 8: Impulse Response of Equity Injections (left panel) and fraction of firms in default (right panel) to a 25 basis point interest rate hike in three benchmark economies: complete markets (blue), relationships (orange), and competitive market (green). Values are expressed as percentage deviations from steady state.

Figure 9 shows that relationships also dampen the transmission of monetary policy to macro aggregates, with the relational market response lying between those of the competitive and complete markets. Even in the complete markets economy there is a large response to monetary policy, because firms respond to changes in the discount rate also when they are unconstrained.

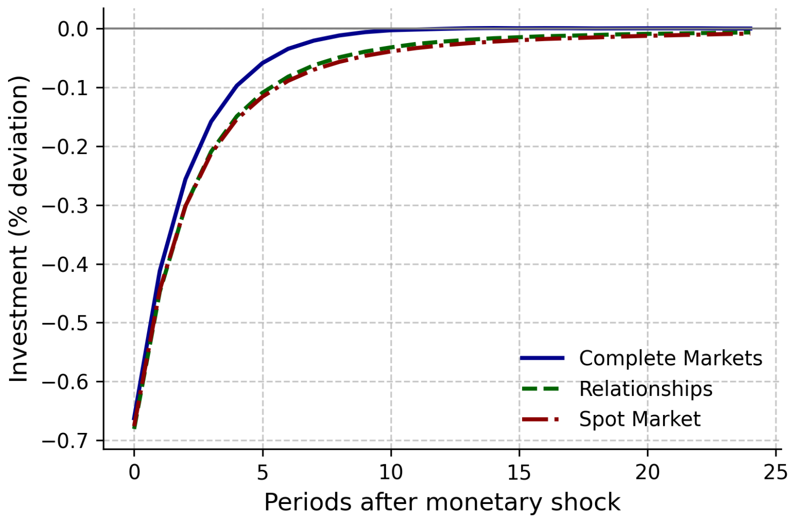


Figure 9: Impulse response of aggregate investment to a 25 basis points interest rate hike with a persistence of 0.6 in three benchmark economies: complete markets (full commitment), relationships (partial commitment), and competitive spot markets (lack of commitment).

7.2.3 Response to a fiscal policy shock

I now examine the case of fiscal policy, designed as a lump sum cash transfer from the representative household to the corporate sector. The transfer, totaling 10 percent of steady-state output, is distributed uniformly across firms. After the transfer, firms become less constrained as their net worth increases. In a relational market, banks optimally charge high spreads after the shock. Therefore, a fraction of the transfer effectively ends up as bank profits. In the aggregate, the insurance mechanism leads to a dampening in the peak response fiscal policy of almost 20 percent.

Compared to the case of monetary policy, the role of relationships is starker for fiscal policy. Fiscal transfers make firms less constrained without affecting their investment opportunities, and are thus a natural candidate for the insurance mechanism. Instead, the case of monetary policy is more ambiguous. Higher rates deplete firms net worth, which makes them more constrained, but also lower their investment demand, which makes them less constrained. Indeed, the response to monetary policy in the cases of competitive spot market and complete markets is quantitatively similar.

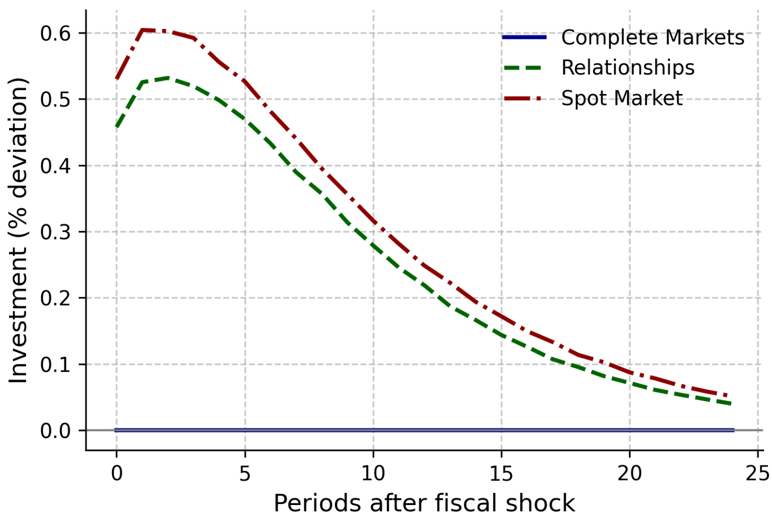


Figure 10: Impulse response of aggregate investment to corporate fiscal transfer policy uniformly distributing 10 percent of output to firms. Three benchmark economies are displayed: complete markets (full commitment), relationships (partial commitment), and competitive spot markets (lack of commitment).

8 Conclusion

I document several novel patterns of loan pricing consistent with an insurance view of banking relationships. When a firm borrows again from the same bank, the pass-through from its default risk to its borrowing rate is close to zero, and lending rates exhibit strong history-dependence: past default risk is more important than contemporaneous risk for loan pricing. Instead, when a firm switches to a new bank, the pass-though is close to one, as predicted by competitive pricing. Furthermore, a large fraction of the prior wedge with competitive pricing is eliminated upon switching.

I rationalize these findings through a new model in which firms establish long-term relationships with banks. Relationships act as an implicit contract over future credit conditions. Starting a new relationship is always possible but it is costly. These switching costs create commitment to the relationship, enabling risk-sharing: firms pay higher rates in good times, in exchange for cheap credit in bad times, when they are more constrained. As switching costs approach zero, the model converges into a competitive credit market, with shocks fully passed through to the lending rate.

The estimated model matches the empirical patterns of limited pass-through and history dependence. It also generates new predictions on when firms receive cheap loans and when instead they are tempted to switch to a new bank, for which I find strong empirical support. Firms receive cheap credit, compared to their default risk, when their sales decline are when they are not distributing dividends, which is associated with tighter financial constraints. Throughout a relationship, firms receive low rates early and pay high rates later in the tenure, as they become less constrained by accumulating earnings. Furthermore, in both the data and the model, firms are tempted to switch to a new bank when their default risk has declined throughout the relationship and thus find themselves disadvantaged by the insurance agreement.

Switching costs, by enabling risk-sharing, improve the allocation of capital in the economy: relationships restore over 10 percent of welfare losses due to financial frictions. Finally, relationships dampen the pass-through of monetary and fiscal policy, since a portion of these policy shocks is absorbed by banks as part of the optimal risk-sharing agreements with firms, which weakens the transmission of macroeconomic policy to investment.

References

Acharya, Viral V, Matteo Crosignani, Tim Eisert, and Sascha Steffen. 2022. "Zombie lending: Theoretical, international, and historical perspectives." *Annual Review of Financial Economics* 14 (1):21–38.

Ai, Hengjie, Anmol Bhandari, Yuchen Chen, and Chao Ying. 2019. "Capital misallocation and risk sharing." *Available at SSRN 3521553* .

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

Álvarez, Laura, Miguel García-Posada, and Sergio Mayordomo. 2023. "Distressed firms, zombie firms and zombie lending: a taxonomy." *Journal of Banking & Finance* 149:106762.

Amiti, Mary and David E Weinstein. 2018. "How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data." *Journal of Political Economy* 126 (2):525–587.

Arellano, Cristina, Yan Bai, and Patrick J Kehoe. 2019. "Financial frictions and fluctuations in volatility." *Journal of Political Economy* 127 (5):2049–2103.

Artavanis, Nikolaos, Brian Jonghwan Lee, Stavros Panageas, and Margarita Tsoutsoura. 2022. "Cross-subsidization of bad credit in a lending crisis." Tech. rep., National Bureau of Economic Research.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. 2021. "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models." *Econometrica* 89 (5):2375–2408.

Auclert, Adrien, Matthew Rognlie, Ludwig Straub, and Tomáš Tapák. 2024. "When do Endogenous Portfolios Matter for HANK?" *Unpublished Manuscript* .

Banerjee, Ryan, Leonardo Gambacorta, and Enrico Sette. 2021. "The real effects of relationship lending." *Journal of Financial Intermediation* 48:100923.

Beaudry, Paul and John DiNardo. 1991. "The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data." *Journal of political Economy* 99 (4):665–688.

Beck, Thorsten, Hans Degryse, Ralph De Haas, and Neeltje Van Horen. 2018. "When arm's length is too far: Relationship banking over the credit cycle." *Journal of Financial Economics* 127 (1):174–196.

Berger, Allen N and Gregory F Udell. 1992. "Some evidence on the empirical significance of credit rationing." *Journal of Political Economy* 100 (5):1047–1077.

———. 1995. "Relationship lending and lines of credit in small firm finance." *Journal of business* :351–381.

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

Bethune, Zachary, Guillaume Rocheteau, Tsz-Nga Wong, and Cathy Zhang. 2022. "Lending relationships and optimal monetary policy." *The Review of Economic Studies* 89 (4):1833–1872.

Bolton, Patrick, Xavier Freixas, Leonardo Gambacorta, and Paolo Emilio Mistrulli. 2016. "Relationship and transaction lending in a crisis." *The Review of Financial Studies* 29 (10):2643–2676.

Boppart, Timo, Per Krusell, and Kurt Mitman. 2018. "Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative." *Journal of Economic Dynamics and Control* 89:68–92.

Caballero, Ricardo J, Takeo Hoshi, and Anil K Kashyap. 2008. "Zombie lending and depressed restructuring in Japan." *American economic review* 98 (5):1943–1977.

Cao, Jin, Pierre Dubuis, and Karolis Liaudinskas. 2023. "Relationship Lending and Monetary Policy Pass-Through." .

Cetorelli, Nicola et al. 2001. "Competition among banks: Good or bad?" *Economic Perspectives-Federal Reserve Bank of Chicago* 25 (2):38–48.

Cooley, Thomas, Ramon Marimon, and Vincenzo Quadrini. 2004. "Aggregate consequences of limited contract enforceability." *Journal of political Economy* 112 (4):817–847.

Cooley, Thomas F and Vincenzo Quadrini. 2001. "Financial markets and firm dynamics." *American economic review* 91 (5):1286–1310.

Cooper, Russell and Joao Ejarque. 2001. "Exhuming Q: Market power vs. capital market imperfections."

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

Demiroglu, Cem, Christopher James, and Guner Velioglu. 2022. "Why are commercial loan rates so sticky? The effect of private information on loan spreads." *Journal of Financial Economics* 143 (2):959–972.

Diamond, Douglas W. 1984. "Financial intermediation and delegated monitoring." *The review of economic studies* 51 (3):393–414.

Dougal, Casey, Joseph Engelberg, Christopher A Parsons, and Edward D Van Wesep. 2015. "Anchoring on credit spreads." *The Journal of Finance* 70 (3):1039–1080.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2017. "The deposits channel of monetary policy." *The Quarterly Journal of Economics* 132 (4):1819–1876.

Eaton, Jonathan and Mark Gersovitz. 1981. "Debt with potential repudiation: Theoretical and empirical analysis." *The Review of Economic Studies* 48 (2):289–309.

Elyasiani, Elyas and Lawrence G Goldberg. 2004. "Relationship lending: a survey of the literature." *Journal of Economics and Business* 56 (4):315–330.

Faria-e Castro, Miguel, Pascal Paul, and Juan M Sánchez. 2024. "Evergreening." *Journal of Financial Economics* 153:103778.

Fried, Joel and Peter Howitt. 1980. "Credit rationing and implicit contract theory." *Journal of Money, Credit and Banking* 12 (3):471–487.

Gil Hubert, Johanna. 2004. "The Mexican Credit Reporting Industry Reform: A Case Study." .

Giometti, Marco. 2022. "Relationship Lending when Borrowers are in Distress." *University of Pennsylvania The Wharton School, ed* .

Gomes, Joao, Urban Jermann, and Lukas Schmid. 2016. "Sticky leverage." *American Economic Review* 106 (12):3800–3828.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi. 2005. "Insurance within the firm." *Journal of Political Economy* 113 (5):1054–1087.

Guntin, Rafael. 2023. "Firms' Rollover Risk and Macroeconomics Dynamics." *Unpublished manuscript* .

Hachem, Kinda. 2011. "Relationship lending and the transmission of monetary policy." *Journal of Monetary Economics* 58 (6-8):590–600.

Hamano, Masashige, Philip Schnattinger, Mototsugu Shintani, Iichiro Uesugi, and Francesco Zanetti. 2024. "Credit Market Tightness and the Share of Zombie Firms: Theory and Evidence from Japan." .

Heckman, J. 1979. "Sample selection bias as a specification error." *Econometrica* .

Heckman, James J and Edward J Vytlacil. 2007. "Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments." *Handbook of econometrics* 6:4875–5143.

Hopenhayn, Hugo and Ivan Werning. 2008. "Equilibrium default." *Manuscript, MIT* 4 (21):6.

Hoshi, Takeo. 2006. "Economics of the living dead." *The Japanese Economic Review* 57 (1):30–49.

Ioannidou, Vasso and Steven Ongena. 2010. ""Time for a change": loan conditions and bank behavior when firms switch banks." *The Journal of Finance* 65 (5):1847–1877.

Jeenas, Priit. 2018. "Monetary policy shocks, financial structure, and firm activity: A panel approach." *Financial Structure, and Firm Activity: A Panel Approach (January 5, 2018)* .

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. "Monetary policy according to HANK." *American Economic Review* 108 (3):697–743.

Kaplan, Steven N and Luigi Zingales. 1997. "Do investment-cash flow sensitivities provide useful measures of financing constraints?" *The quarterly journal of economics* 112 (1):169–215.

Karolyi, Stephen Adam. 2018. "Personal lending relationships." *The Journal of Finance* 73 (1):5–49.

Khan, Aubhik, Tatsuro Senga, and Julia K Thomas. 2014. "Default risk and aggregate fluctuations in an economy with production heterogeneity." *Unpublished manuscript* 5.

Kovrijnykh, Natalia. 2013. "Debt contracts with partial commitment." *American Economic Review* 103 (7):2848–2874.

Krusell, Per and Anthony A Smith, Jr. 1998. "Income and wealth heterogeneity in the macroeconomy." *Journal of political Economy* 106 (5):867–896.

Martin, Alberto, Sergio Mayordomo, and Victoria Vanasco. 2023. "Banks vs. firms: who benefits from credit guarantees?" .

Moll, Benjamin. 2014. "Productivity losses from financial frictions: Can self-financing undo capital misallocation?" *American Economic Review* 104 (10):3186–3221.

Mukoyama, Toshihiko. 2021. "MIT shocks imply market incompleteness." *Economics Letters* 198:109666.

Müller, Andreas, Kjetil Storesletten, and Fabrizio Zilibotti. 2019. "Sovereign debt and structural reforms." *American Economic Review* 109 (12):4220–4259.

Ottonello, Pablo and Thomas Winberry. 2020. "Financial heterogeneity and the investment channel of monetary policy." *Econometrica* 88 (6):2473–2502.

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

Rampini, Adriano A and S Viswanathan. 2010. "Collateral, risk management, and the distribution of debt capacity." *The Journal of Finance* 65 (6):2293–2322.

Rocheteau, Guillaume, Randall Wright, and Cathy Zhang. 2018. "Corporate finance and monetary policy." *American Economic Review* 108 (4-5):1147–1186.

Schäfer, Larissa. 2019. ""Forgive but not forget": The behavior of relationship banks when firms are in distress." *Review of Finance* 23 (6):1079–1114.

Scharfstein, David and Adi Sunderam. 2016. "Market power in mortgage lending and the transmission of monetary policy." *Unpublished working paper. Harvard University* 2.

Souchier, Martin. 2022. "The pass-through of productivity shocks to wages and the cyclical competition for workers." *Unpublished Manuscript* .

———. 2024. "Insurance Inside and Outside the Firm." *Unpublished Manuscript* .

Spear, Stephen E and Sanjay Srivastava. 1987. "On repeated moral hazard with discounting." *The Review of Economic Studies* 54 (4):599–617.

Stiglitz, Joseph E and Andrew Weiss. 1981. "Credit rationing in markets with imperfect information." *The American economic review* 71 (3):393–410.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. MIT press.

A Empirics

A.1 Risk Assessments validation

In Table 18, I validate the risk assessments by assessing their ability to predict future loan losses. Each additional percentage probability of default is associated with approximately 0.4 percent loan losses, defined as the sum of forgiven loan and write-offs. Loan spread does not seem to be more weakly associated to loan losses.

	Stayers	Switchers	Stayers	Switchers
	Loan Losses	Loan Losses	Loan Losses	Loan Losses
Default Risk	0.364^{***}	0.386^{***}	0.323^{***}	0.366^{***}
	(0.010)	(0.025)	(0.010)	(0.029)
Spread			0.074^{***}	0.012^{***}
			(0.003)	(0.006)
Cons	-0.002^{***}	-0.001^{***}	-0.006^{***}	-0.002^{***}
R^2	0.000395	0.00670	0.000541	0.00749
N	259,130	26,432	249,151	24,077

Table 18: Loan Losses are the sum of charge-offs and forgiven loans. Default Risk is the assessment by lenders in R04.

In Table 19, I evaluate the role of past risk assessments in predicting loan losses. The Table shows that loan losses are far more sensitive to contemporaneous default risk than to the default risk assessed at the onset of a relationship. This reinforces the interpretation of the result on history-dependence in Section 3 as being driven by relationship considerations, as opposed to the effective relevance of past information.

	(1)	(2)
	Loan Losses	Loan Losses
PD_t	0.435***	0.419***
	(0.026)	(0.028)
PD_0		0.0542*
		(0.025)
Constant	−0.000234	−0.00131
R^2	0.00144	0.00147
N	187,400	187,400

Table 19: Test for the role of past information in predicting loan losses.Loan Losses are the sum of charge-offs and forgived loans. Default Risk is the assessment by lenders in R04. Sample in the first column is restricted to firms that have PD_0.

A.2 Default Risk Pass-Through

A.2.1 Controls

In Table 20, I replicate the pass-through analysis controlling for loan- and firm- specific characteristics. Since the regression specification is in changes, there is not an obvious way to control for these characteristics. I follow Guiso, Pistaferri, and Schivardi (2005) and residualize both spreads and default risk on the set of controls before running the pass-through regressions.

	(1)	(3)
	$\Delta Spread$	$\Delta Spread$
$\mathbb{1}_{SameBank}$	1.407^{***}	0.559^{***}
	(0.042)	(0.050)
ΔPD	0.822^{***}	0.492^{***}
	(0.025)	(0.034)
$\mathbb{1}_{SameBank} \times \Delta PD$	-0.796^{***}	-0.380^{***}
	(0.026)	(0.034)
Constant	-1.248^{***}	-0.344^{***}
	(0.042)	(0.050)
Controls	No	Yes
R^2	0.0243	0.00572
N	233815	211493

Table 20: Pass-through of firm default risk to loan spread. In the second column, both spreads and default risk are residualized against a set of controls, as in Guiso, Pistaferri, and Schivardi (2005). The controls include maturity and loan type (credit line or term loan).

	(1)	(2)	(3)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
$\mathbb{1}_{SameBank}$	0.620^{***}	0.774^{***}	1.058^{***}
	(0.0553)	(0.0605)	(0.1037)
ΔPD	0.626^{***}	0.576^{***}	0.378^{***}
	(0.0309)	(0.0314)	(0.0563)
$\mathbb{1}_{SameBank} \times \Delta PD$	-0.736^{***}	-0.692^{***}	-0.493^{***}
	(0.0318)	(0.0326)	(0.0584)
Constant	-0.052	-0.121^{*}	-0.252^{*}
	(0.0561)	(0.0614)	(0.1086)
Firm FE			✓
Bank x Time FE		✓	✓
R^2	0.0187	0.189	0.477
N	110211	109860	62277

Table 21: Pass-through of default risk to loan spreads, excluding credit lines.

	(1)	(2)	(3)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
$\mathbb{1}_{SameBank}$	1.985^{***}	0.999^{***}	0.664^{***}
	(0.0273)	(0.0326)	(0.0423)
ΔPD	0.518^{***}	0.242^{***}	0.206^{***}
	(0.0168)	(0.0170)	(0.0213)
$\mathbb{1}_{SameBank} \times \Delta PD$	-0.459^{***}	-0.176^{***}	-0.140^{***}
	(0.0176)	(0.0179)	(0.0228)
Constant	-1.665^{***}	-0.826^{***}	-0.554^{***}
	(0.0266)	(0.0299)	(0.0373)
R^2	0.0263	0.145	0.348
N	476112	475779	350764

Table 22: Pass-through of default risk to loan spreads, including also collateralized loans.

A.2.2 Soft and hard assessments separately

In this section, I provide the pass-through and the pricing reconnect results distinguishing between the headline risk assessment (23), and keeping in isolation the hard and the soft information components, respectively in Table 24 and 25.

I also report separately the results for bank stayers and switchers. This also allows to examine the R^2 of these regressions more closely. These are quite low when focusing on bank stayers, consistent with the loan pricing disconnect narrative. Instead, they are much higher for bank switchers. Specifically, the correction of the prior pricing wedge (R_{t-1}^{wedge}) alone accounts for more than a third of the change in lending rate upon switching. Together, ΔPD and R_{t-1}^{wedge} account for over 40 percent of such variation.

	Stayers			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	0.026***		0.116***	0.822***		0.734***
	(0.0061)		(0.0059)	(0.0245)		(0.0176)
R^{wedge}_{t-1}		0.166***	0.176***		0.879***	0.896***
		(0.0012)	(0.0016)		(0.0066)	(0.0066)
Constant	0.159***	0.033***	0.196***	−1.248***	−0.662***	−0.721***
	(0.0080)	(0.0055)	(0.0075)	(0.0415)	(0.0295)	(0.0298)
R^2	0.0000909	0.0582	0.0592	0.0352	0.368	0.404
N	202894	332880	192446	30921	30780	28751

Table 23: Pass-through of default risk and persistence of the pricing wedge R^{wedge}_{t-1} (loan pricing reconnect). The first three columns are bank stayers. The last three columns are bank switchers. This table uses the headline risk assessments reported by banks, which combines hard and soft information.

	Stayers			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	0.045***		0.109***	0.847***		0.802***
	(0.0070)		(0.0067)	(0.0284)		(0.0204)
R^{wedge}_{t-1}		0.172***	0.183***		0.887***	0.899***
		(0.0012)	(0.0017)		(0.0066)	(0.0069)
Constant	0.169***	1.434***	1.714***	−1.351***	6.562***	6.556***
	(0.0085)	(0.0105)	(0.0159)	(0.0437)	(0.0624)	(0.0660)
R^2	0.000211	0.0608	0.0617	0.0308	0.375	0.407
N	194064	330862	183778	27987	30398	26093

Table 24: Pass-through of default risk and persistence of the pricing wedge R^{wedge}_{t-1} (loan pricing reconnect). The first three columns are bank stayers. The last three columns are bank switchers. This table uses the hard information component of the risk assessments only.

	Stayers			Switchers		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	0.015		0.105^{***}	-0.020		0.293^{**}
	(0.0159)		(0.0149)	(0.1057)		(0.1035)
R^{wedge}_{t-1}		0.115^{***}	0.160^{***}		0.432^{***}	0.561^{***}
		(0.0025)	(0.0081)		(0.0441)	(0.0695)
Constant	-0.023	0.201^{***}	0.286^{***}	0.252	1.675^{***}	1.383^{***}
	(0.0220)	(0.0077)	(0.0265)	(0.2423)	(0.1658)	(0.2679)
R^2	0.000208	0.0491	0.0871	0.000132	0.0880	0.191
N	4188	42170	4114	288	998	279

Table 25: Pass-through of default risk and persistence of the pricing wedge R^{wedge}_{t-1} (loan pricing reconnect). The first three columns are bank stayers. The last three columns are bank switchers. This table uses the soft information component of the risk assessments only.

A.3 Competition at the Municipality-level

Does the loan pricing behavior of banks change depending on the availability of outside options for the firm? In this section, I test this hypothesis using geographical variation in the availability of outside banking options for firms. To do so, I use branch data at the municipality level to measure the market share of each lender in each of Mexico's 2500 municipalities. I then proceed to test whether the loan pricing disconnect documented in previous sections is stronger when the bank has a large market share. I find that when a bank has a large market share, it exhibits both lower pass-through and stronger history-dependence, consistent with the idea that when firms have plenty of outside options, loan pricing patterns are closer to that of a competitive market. Instead, when it is hard to a firm to switch, the relational pricing patterns emerge more strongly.

I compute the market share of bank b in municipality m at time τ, the time of the previous loan taken by the firm, as the share of the total branches in that municipality that belong to that bank: $MS_{bm\tau} = N_{bm\tau}/\sum_b N_{bm\tau}$. Then, I restrict the sample to firms that remain with the same bank, the stayers, and run the following regression:

$$\Delta Spread_{ft} = \alpha + \beta_1 \Delta PD_{ft} + \beta_2 \Delta PD_{ft} \times MS_{bm\tau} + \beta_3 MS_{bm\tau} + \varepsilon_{ft} \tag{20}$$

I find that β_2 is negative and significant, suggesting that when a bank has a large market share in a municipality, its pricing decisions are more disconnected from firm default risk.

Importantly, I also show that the lack of outside banking options plays a role in the history-dependence patterns. As discussed in more detail in Section 6.5, a simple market power model could deliver limited pass-through which is stronger in geographies with less competition, but would not be consistent with history dependence, nor with history dependence being stronger when firms have worse outside options.

To test the role of competition in determining the extent of history-dependece documented in Section 3.3, I run the following regression:

$$Spread_{ft} = \alpha + \beta_1 PD_{ft} + \beta_2 \Delta PD_{f0} + \beta_3 \Delta PD_{f0} \times MS_{bm\tau} + \beta_4 MS_{bm\tau} + \varepsilon_{ft} \tag{21}$$

First of all, I confirm the result of Section X that β_2 is positive, suggesting that past information matters for loan pricing even after controlling for the contemporaneous default risk. More importantly, I find that $\beta_3 > 0$, which implies that history-dependent pricing is more pronounced when the bank has a large market share in the municipality.

	(1)	(2)	(3)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	0.221***	0.147***	0.122***
	(0.015)	(0.015)	(0.021)
$MS_{b,t-1}$	−1.393***	−0.582**	−1.351
	(0.153)	(0.185)	(0.938)
$\Delta PD \times MS_{b,t-1}$	−1.187***	−0.493***	−0.709***
	(0.145)	(0.141)	(0.206)
Cons	0.252***	0.176***	0.278***
	(0.017)	(0.019)	(0.081)
Firm FE			✓
Bank x Time FE		✓	✓
R^2	0.00242	0.109	0.324
N	131,227	130,967	84,398

Table 26: Default risk pass-through by branch market share of the bank in the municipality where the firm is headquartered, $MS_{b,t-1}$. Robust standard errors in parentheses.

	(1)	(2)	(3)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	0.221^{***}	0.146^{***}	0.287^{***}
	(0.0145)	(0.0140)	(0.0216)
$\Delta PD \times MS_{b,t-1}$	-1.187^{***}		-1.767^{***}
	(0.145)		(0.188)
$\Delta PD \times HFI_{t-1}$		-0.182^{*}	-0.299^{**}
		(0.0890)	(0.0984)
R^2	0.00242	0.00147	0.00266
N	131227	131231	131227

Table 27: History-dependence by branch market share of the bank in the municipality where the firm is headquartered, $MS_{b,t-1}$. Robust standard errors in parentheses.

A.4 Heckman Correction

In this section, I provide an alternative approach to use the two instruments for bank switching described in Section 3.2.2. There, I used such instruments as simple IVs. However, when dealing with a selection problem, two approaches are possible: IV and selection models, such as the Heckman correction (Heckman, 1979).

In my setting, using IV is possible because I observe the outcome both if the firm is selected into switching or not. The most natural applications of the selection corrections are instead those cases where the outcome is observed only conditional on selection, such as with labor income and selection into the workforce. Still, the Heckman correction has some advantages even in my setup. Most notably, because the pass-through equations are estimated separately for stayers and switchers, the exogeneity requirement applies one equation at a time. This is helpful as a robustness against the possibility that the shocks to the old bank which I use as an instrument may be related to the loan pricing policy of that bank.

As discussed in **?** and in Heckman and Vytlacil (2007), this procedure relies on much weaker identification requirements to identify slope coefficients, rather than the intercept. Since I am interested in the pass-through, the coefficient on ΔPD, the slope coefficient is exactly the object of interest in my case. I present the results in Table 28. Again, while the magnitude change compared to the OLS and the IV, we get the familiar result of low pass-through for stayers and high pass-through for switchers.

	Branch IV		AW IV	
	Stayers	Switchers	Stayers	Switchers
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
Main Equation				
ΔPD	0.0517***	0.407***	0.0155*	0.404***
	(0.0086)	(0.0350)	(0.0078)	(0.0302)
Constant	1.032***	−14.83***	1.056***	19.24***
	(0.0451)	(1.2775)	(0.0441)	(0.5344)
Selection Equation				
Z	1.812***	−0.814***	0.452***	−0.250***
	(0.1597)	(0.2103)	(0.0161)	(0.0191)
Constant	0.423***	−1.645***	0.413***	−1.627***
	(0.0031)	(0.0038)	(0.0028)	(0.0035)
ρ	−0.416***	0.861***	−0.399***	−1.285***
	(0.0223)	(0.0757)	(0.0227)	(0.0298)
σ	1.358***	2.222***	1.322***	2.442***
	(0.0048)	(0.0408)	(0.0046)	(0.0176)

Table 28: Pass-through of default risk to loan spreads. Heckman selection correction.

A.5 History-Dependence

Table 29 shows that loan spreads depend on both past risk assessments and past loan spreads. Past loan spreads capture explain a large fraction of current loan spreads. Current and past default risk are similarly important for loan pricing, both conditionally and unconditionally on past spreads.

	(1)	(2)
	$Spread_t$	$Spread_t$
PD_t	0.619***	0.129***
	(0.0133)	(0.0080)
PD_{t-1}	0.620***	0.092***
	(0.0109)	(0.0062)
$Spread_{t-1}$		0.810***
		(0.0019)
Constant	6.090***	1.308***
	(0.0311)	(0.0218)
R^2	0.0455	0.644
N	210450	205963

Table 29: Dependence of loan spreads on past default risk and past loan spreads.

A.6 State-Dependence

In this section, I show an additional feature of the data: the pass-through of default risk to lending rates depend on the state of the bank-firm relationship. The intuition is as follows. When the pricing wedge $R^{wedge}_{ft} = (Spread_{ft} - PD_{ft})$ is negative ($R^{wedge}_{ft} < 0$), the bank is already charging to the firm a rate which is lower than what is justified by its fundamental default risk. Therefore, if the conditions of the firm deteriorate further ($\Delta PD_{ft} > 0$), the bank might be inclined to pass-through a large portion of this shock. Conversely, when the pricing wedge is positive ($R^{wedge}_{ft} > 0$), the firm is already being charged a high rate, and banks could thus more easily neglect an increase in firm risk. The opposite pattern would occur when the firm improves ($\Delta PD_{ft} < 0$). In this case, we expect the bank to lowers rates and pass-through the shock more strongly when the firm is tempted to leave if the bank does not improve their terms, that is, if $R^{wedge}_{ft} > 0$.

To assess such state-dependent pass-through, I first define a dummy $\mathbb{1}^{Load}_{ft}$ capturing whether a shock is loading in the same direction of the previous mispricing. Specifically, I define:

$$\mathbb{1}^{Load}_{ft} = \begin{cases} 0 & \text{if} \quad \Delta PD_{ft} \times R^{wedge}_{ft} > 0 \\ 1 & \text{if} \quad \Delta PD_{ft} \times R^{wedge}_{ft} < 0 \end{cases} \tag{22}$$

Formally, the empirical question is whether the pass-through from ΔPD_{ft} to $\Delta Spread_{ft}$ depends on the prior pricing wedge R^{wedge}_{ft}. To test this hypothesis, I run the following regression:

$$\Delta Spread_{ft} = \alpha + \beta_1 \Delta PD_{ft} + \beta_2 \Delta PD_{ft} \times \mathbb{1}^{Load}_{ft} + \beta_3 \mathbb{1}^{Load}_{ft} + \varepsilon_{ft} \tag{23}$$

A positive sign for β_2 suggests the presence of state-dependent pass-through in the expected direction. When a shock loads in the same direction of a previous pricing wedge, which would make the wedge larger if not passed-through, the pass-through is larger.

	(1)	(2)	(3)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	-0.349^{***}	-0.403^{***}	-0.576^{***}
	(0.0109)	(0.0116)	(0.0198)
$\mathbb{1}^{Load}$	-0.319^{***}	-0.358^{***}	-0.430^{***}
	(0.0199)	(0.0204)	(0.0305)
$\mathbb{1}^{Load} \times \Delta PD$	0.961^{***}	1.201^{***}	1.653^{***}
	(0.0178)	(0.0201)	(0.0401)
Constant	0.447^{***}	0.473^{***}	0.354^{***}
	(0.0139)	(0.0141)	(0.0192)
R^2	0.0286	0.158	0.376
N	144,029	143,665	91,500

Table 30: State-dependence in the default risk pass-through, depending on whether the shock worsen or reduce the prior pricing wedge, which is captured by $\mathbb{1}^{Load}$ defined in Equation (22).

Notice that, because of the patterns highlighted in Section 3.4, simply using $\Delta Spread_{ft}$ may lead to mechanical results. Indeed, when $R^{wedge}_{ft} < 0$, the firm is borrowing at a subsidized rate, and there is a tendency for its borrowing rate to increase in subsequent loans, especially when it switches to a new bank. Therefore, we may for example erroneously estimate a large pass-through in cases when $\Delta PD_{ft} > 0$ and $R^{wedge}_{ft} < 0$, simply because $\Delta Spread_{ft}$ is positive and large to correct a past mispricing. To address such concern, I estimate Equation (23) using on the left-hand side the residuals of Equation (4), which captures the surprise component in the change in spread, after controlling for past information contained in the prior pricing wedge. Results are reported in Table 31.

	(1)	(2)	(3)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
ΔPD	0.106^{***}	-0.029^{**}	-0.338^{***}
	(0.0103)	(0.0108)	(0.0178)
$\mathbb{1}^{Load}$	-0.171^{***}	-0.218^{***}	-0.359^{***}
	(0.0191)	(0.0191)	(0.0276)
$\mathbb{1}^{Load} \times \Delta PD$	-0.007	0.446^{***}	1.272^{***}
	(0.0169)	(0.0187)	(0.0362)
Constant	0.305^{***}	0.332^{***}	0.103^{***}
	(0.0133)	(0.0132)	(0.0174)
R^2	0.00170	0.178	0.421
N	144,029	143,665	91,500

Table 31: State-dependence in the default risk pass-through, depending on whether the shock worsen or reduce the prior pricing wedge, which is captured by $\mathbb{1}^{Load}$ defined in Equation (22).

A.7 Dynamic Pass-Through

How persistent are the pass-through over time? I have shown that changes in loan spreads are not immediately passed-through to loan rates within a relationship, but does this pass-through becomes larger overtime? To answer this question, I adapt the analysis in (2) but rather than focusing on the contemporaneous pass-through, I measure the effects over the subsequent five loans that a firm obtains. The notion of switching in this case is slightly different. Before, I was focusing on pairs of consecutive loans. To study the pass-through after s loans, I use the specification in Equation (24), thus regressing the cumulative spread over the next h loans after time t on the change in probability of default that the firm experienced right before time t. I separate the sample in firms that stay with the same bank throughout the s loans, and firms that switch to a new bank immediately at time t, and then keep borrowing from that bank throughout the s loans.[31] Results are reported in Figure 11.

$$\Delta Spread_{f,t+h} = \alpha + \beta_1 \Delta PD_{ft} + \varepsilon_{ft} \tag{24}$$

[31] For this exercise, I only keep firms that borrow from one bank at a time, and compute the average loan rate whenever the firm has multiple loans in the same month.

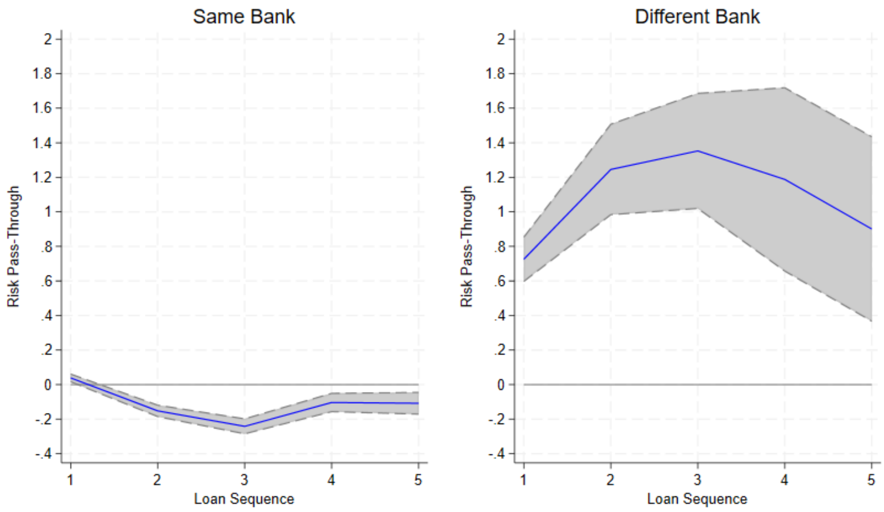


Figure 11: Computes the dynamic response of spreads to a change in firm default risk over the subsequent five loans, separately for stayers and switchers.

A.8 Pricing Reconnect

Table 32 reports the regression results for Figure 2, displayed in Section 3.4. The results show a stark loan pricing *reconnect* upon switching. The previous gap between how much a firm gets charged and its default risk, $R^{wedge}_{f,t-1}$, dissipates upon switching.

	(1)	(2)	(3)
	$\Delta Spread_{ft}$	$\Delta Spread_{ft}$	$\Delta Spread_{ft}$
$R^{wedge}_{f,t-1}$	-0.936^{***}	-0.943^{***}	-1.279^{***}
	(0.010)	(0.009)	(0.014)
$R^{wedge}_{f,t-1} \times \mathbb{1}_{SameBank}$	0.766^{***}	0.613^{***}	0.230^{***}
	(0.010)	(0.009)	(0.014)
Firm FE			✓
Bank x Time FE		✓	✓
R^2	0.142	0.342	0.694
Obs	165,914	165,533	119,943

Table 32: Regression of changes in loan spreads on the previous loan pricing wedge $R^{wedge}_{f,t-1}$. The wedge is measured as the difference between the loan spread and the assessed probability of default. Robust standard errors in parenthesis.

	(1)	(2)	(3)
	$\Delta Spread$	$\Delta Spread$	$\Delta Spread$
$\mathbb{1}_{SameBank}$	1.893***	1.014***	0.649***
	(0.0286)	(0.0336)	(0.0438)
ΔPD	0.518***	0.242***	0.205***
	(0.0168)	(0.0170)	(0.0213)
$\mathbb{1}_{SameBank} \times \Delta PD$	−0.425***	−0.140***	−0.119***
	(0.0187)	(0.0191)	(0.0250)
$\mathbb{1}_{LongTerm}$	0.153***	−0.022***	0.079***
	(0.0136)	(0.0150)	(0.0340)
$\mathbb{1}_{LongTerm} \times \Delta PD$	−0.064***	−0.058***	−0.030***
	(0.0108)	(0.0108)	(0.0156)
Constant	−1.665***	−0.830***	−0.586***
	(0.0266)	(0.0300)	(0.0384)
Bank x Time FE		✓	✓
Firm FE			✓
R^2	0.0267	0.145	0.350
N	269,906	269,573	144,338

Table 33: Pass-through from changes in default risk to changes in loan spreads. $\mathbb{1}_{SameBank}$ is a dummy capturing whether the firm is staying with the same bank or switching. $\mathbb{1}_{LongTerm}$ is a dummy capturing whether the firm has stayed with its current bank for more than two years. Robust standard errors in parenthesis.

A.9 Duration of the relationship

In this section, I show that the extensive margin of relationships is far more important than the intensive margin in determining risk pass-through. Specifically, I divide my sample not only in stayers and switchers, as in Section 3, but I also subdivide stayers depending on the length of their relationship with their bank. I define a dummy for long-term relationship, $\mathbb{1}_{ft}^{LongTerm}$, which is equal to zero if the length of the relationship is below 24 months (which includes all the switchers), and one when the length is above 24 months.

$$\begin{aligned}\Delta Spread_{ft} = \alpha + \beta_1 \Delta PD_{ft} + \beta_2 \Delta PD_{ft} \times \mathbb{1}_{ft}^{SameBank} + \beta_3 \mathbb{1}_{ft}^{SameBank} \\ + \beta_4 \Delta PD_{ft} \times \mathbb{1}_{ft}^{LongTerm} + \beta_5 \mathbb{1}_{ft}^{LongTerm} + \varepsilon_{ft}\end{aligned} \tag{25}$$

Results shows that the results for stayers and switchers remain broadly unchanged. Instead, being in a long-term relationship has a quantitatively small effect on the pass-through, and a sign which is the opposite of what one would expect: pass-through are slightly higher for *stayers* in relationships with a long tenure than with *stayers* in a relationship with a short tenure.

On the other hand, firms that have been in a relationship for longer tend to obtain lower rates on average, even though this pattern is entirely explained by firm fixed effects.

A.10 Aggregate Risk Pass-Through

	Idiosyncratic $\Delta Spread$	Aggregate $\Delta Spread$
$\mathbb{1}_{SameBank}$	1.407^{***}	1.324^{***}
	(0.042)	(0.048)
ΔPD	0.822^{***}	
	(0.025)	
$\mathbb{1}_{SameBank} \times \Delta PD$	-0.796^{***}	
	(0.026)	
$\Delta \overline{PD}$		1.877^{***}
		(0.175)
$\mathbb{1}_{SameBank} \times \Delta \overline{PD}$		-1.053^{***}
		(0.181)
Constant	-1.248^{***}	-0.955^{***}
	(0.042)	(0.048)
R^2	0.0243	0.0147
N	233,815	219,905

Table 34: Pass-through of idiosyncratic firm default risk *PD* (left column) and aggregate default risk $\overline{PD}$ (right column). The aggregate default risk is computed as the average default risk across all firms, excluding those in the top fifth percentile.

A.11 Branch Data

In this section, I describe Since the core of my empirical analysis is to study the pricing patterns for firms that stay with the same bank (*stayers*) and contrast it with that of firms that switch to a new bank (*switchers*), I collect bank branch data at the municipality level to construct a separation instrument.

The data is available online, and provided by the *Equipo de Inclusión Financiera* at CNBV, the banking supervisor. The data is provided subdivided by month, and I scrape the whole data and merge together. Once the data is merged, the geographical dimension is organized using INEGI codes (postal codes), which divide Mexico into 32 states, 2,560 municipalities and approximately 250,000 localities. Municipalities are comparable in size to US counties, while localities are approximately the size of a neighbourhood. Because the R04 Credit Registry data reports the geography of firms at the municipality-level, I choose to use this level of aggregation when studying bank branch data. This is the most natural choice for two reasons. First, the vast majority of localities do not have bank branches. Second, the municipality-level provides an ideal level of aggregation to study banking relationship, while localities may be too fine to matter.

I observe branch data for 145 individual banks, or 107 banking groups once I apply the filter described in Section 2. Therefore, in total, I observe branch data for 108 months (January 2013 to December 2022), across 2,432 municipalities, and 107 banks, so that I have a total of 3,957,118 bank-month-municipality observations. Only 1,005 municipalities actually have had a bank branch at some point during the 2013 to 2022 period. The median municipality had branches from 3 banks, and on average each municipality had 4.6 banks with active branches. Firms in the Credit registry are domiciled across 1,900 municipalities.

A.12 Amiti and Weinstein (2018) credit shocks construction

In this section, I briefly illustrate the identification setup borrowed from Amiti and Weinstein (2018) to estimate bank-specific credit supply shocks. For additional details, I defer to the original paper.

The first step is to aggregate data at the bank-firm relationship level, rather than at the loan-level. I compute L_{fbt} as the total outstanding credit from bank b to firm f at month t, which includes both current period originations and previous outstanding loans which has not matured yet. Then, I compute $D(L_{fbt}/L_{fb,t-1})$ as the percentage growth in L_{fbt} within the relationship.

The class of models studied in Amiti and Weinstein (2018) can be summarized as follows:

$$D(L_{fbt}/L_{fb,t-1}) = \alpha_{ft} + \beta_{bt} + \varepsilon_{fbt} \tag{26}$$

the credit growth within a relationship is the sum of a firm-specific demand component and a bank-specific supply component. The approach in Amiti and Weinstein (2018) allows to overcome two challenges with estimating Equation (26). The first is identification, which is crucial for the application in this paper and I will cover in detail. The second is aggregation: by using the correct regression weights and treatment the formation of new relationship appropriately, bank shocks and firm shocks add up naturally to account exactly for aggregate credit growth. This aspect is not particularly relevant for my application, and I defer its treatment to the original paper.

Identification in Equation (26), which is estimated using a fixed-effect WLS regression, is obtained from the presence of multi-bank firms. Intuitively, if a firm has an established relationship with two banks, the relative credit growth in the two relationships will be informative of the credit supply at each bank. The authors also show that omitting the interaction terms Z_{fbt} on the right-side of Equation (26) is not problematic, as long as the bank and firm shocks are interpreted as also encompassing a component of the interaction term.

The idiosyncratic component of the bank shock which I use to construct my instrument for separation in Section 3.2, which is denoted as $\tilde{\beta}_{bt}$ in Amiti and Weinstein (2018), is obtained by subtracting the median bank shock from the bank's fixed effect estimate β_{bt}.

A.13 First-Stage IV

Tables 35 and 36 report the first-stage of the IV reported in Section 3, respectively the branch IV and the credit shock IV constructed as in Amiti and Weinstein (2018). Notice that, following Wooldridge (2010), I instrument for all terms where the potentially endogenous dummy $\mathbb{1}_{SameBank}$ enters: the dummy and its interaction with ΔPD. Columns 2 and 3 are effectively the first-stages used in the regression. I also report the first-column to assess the significance of the instrument in predicting the staying vs switching decision before adding the other controls.

	(1)	(2)	(3)
	$\mathbb{1}_{SameBank}$	$\mathbb{1}_{SameBank}$	$\mathbb{1}_{SameBank} \times \Delta PD$
Z^{branch}	1.614***	0.645***	0.187
	(0.0826)	(0.0511)	(0.130)
$Z^{branch} \times \Delta PD$		0.195**	2.670***
		(0.0633)	(0.0995)
ΔPD		0.00610***	0.760***
		(0.00066)	(0.00104)
Constant	0.860***	0.911***	−0.0101***
	(0.00049)	(0.00049)	(0.00134)
F	152.8	159.0	179,230.7
R^2	0.00268	0.000472	0.759
Obs	170,610	336,901	170,610

Table 35: First-state of the branch market share IV.

	(1)	(3)	(2)
	$\mathbb{1}_{SameBank}$	$\mathbb{1}_{SameBank}$	$\mathbb{1}_{SameBank} \times \Delta PD$
Z^{AW}	0.0410***	0.0681***	0.166***
	(0.0048)	(0.0049)	(0.0080)
$Z^{AW} \times \Delta PD$		0.0467***	0.0950***
		(0.0036)	(0.0060)
ΔPD		−0.0189***	0.773***
		(0.0006)	(0.0010)
Constant	0.883***	0.877***	−0.0545***
	(0.0008)	(0.0008)	(0.0013)
F	74.11	353.2	195,579.6
R^2	0.000443	0.00630	0.778
Obs	167,218	167,218	167,218

Table 36: First-state of the Amiti and Weinstein (2018) IV.

A.14 Banks Mergers and Exit activity

The credit registry reports 146 banks with different lenders keys. One approach could be that of relying on the classification of the credit registry to isolate different banks. However, two caveats apply. First, Mexico has experimented some episodes of mergers and acquisition among banks (e.g. Banco Serfin was acquired by Banco Santander in 2000, and the two banks merged in 2004.). I obtained a list of bank mergers from Banco de Mexico. I follow standard procedure in the literature and attribute a unique identifier to all the banks that ended up forming the ultimate banking group. Second, some small lenders were reincorporated with a new name and new identifier (e.g. Agrofinanzas S.A. de C.V., a sociedad Financiera de Objecto Multiplo, reincorporated as Bankool, an Instituciones de Banca Multiple). I impute the same new identifier to the old entity as well. Finally, several banking groups are organized with a main bank (Banca Multiple) and some separate entities (Sociedades Financieras de Objecto Multiple). I attribute to all entities the identifiers of the Banca Multiple. At the end of this operation, I am left with 107 entities.

B Theory

B.1 Pricing Equation

Recall Proposition 1 from Section 5.3.

Proposition 1: *Consider a firm in state $s = (z, b, k)$ that is solvent ($\overline{b}(s) \geq b$). The equilibrium loan rate is given by:*

$$Q = \underbrace{\beta \frac{E_{z'|z}[w'(z')]}{b'}}_{Q^{zero,w}} + \underbrace{\frac{b^m(s) - w}{b'}}_{\chi:\ subsidy} \tag{27}$$

where $w'(z') = W(s', v'(z'))$ and $w = W(s, v)$ for brevity.

Proof:

The proof follows almost immediately from rearranging the value function of the bank. For a firm that receives a loan, and thus reaches the production phase, $W(s, v) = W_P(s, v)$. Therefore, the bank value reads:

$$W(s,v) = b - Qb' + \beta(1-\theta)E_{z'|z}[W(s, v'(z'))] + \beta\theta E_{z'|z}[b^m(s')]$$

For solvent firms, $b = b^m(s)$. Therefore, we can rearrange that equation to read:

$$Qb' = \big(b^m(s) - W(s,v)\big) + \beta(1-\theta)E_{z'|z}[W(s, v'(z'))] + \beta\theta E_{z'|z}[b^m(s')]$$

Now, denote $w'(z') = W(s', v'(z'))$ and $w = W(s, v)$ for brevity.

We thus obtain:

$$Q = \frac{b^m(s) - w}{b'} + \beta \frac{(1-\theta)E_{z'|z}[w'(z')] + \theta E_{z'|z}[b^m(s')]}{b'}$$

In the special case without exogenous separations ($\theta \to 0$) this yields:

$$Q = \underbrace{\beta \frac{E_{z'|z}[w'(z')]}{b'}}_{Q^{zero,w}} + \underbrace{\frac{b^m(s) - w}{b'}}_{\chi:\ \text{subsidy}}$$

B.2 Competitive Lenders

Recall Proposition 2 from Section 5.3.

Proposition 2: *Suppose that switching costs are zero ($\psi^{f,b} \to 0$). Suppose that a firm in state $s = (z, b, k)$ is solvent, that is, $\overline{b}(s) \geq b$. Then, the optimal contract prescribes that* *loans are priced competitively:*

$$Q = \beta \underbrace{\frac{E_{z'|z}[b^m(s')]}{b'}}_{Q^{zero,b}}$$

That is,

$$Q = \beta \Big[(1 - PD) + PD \times E_{z'|z}[\bar{b}(s')/b | Def] \Big] \tag{28}$$

where $PD = P(\bar{b}(s') < b')$.

Furthermore, the allocations are identical to the ones that would prevail in an economy where firms face lenders that offer competitive pricing schedules in each period, and solve the following problem:

$$V_P(s) = \max_{b',k'} \quad D + \beta E_{z'|z}[V(s')]$$

Subject to:

$$Qb' = \beta E_{z'|z}[b^m(s')] \qquad (\mu^{-1}\text{: PK-Bank})$$

and the usual budget constraint.

Proof:

The proof is divided into two parts. In the first part, I show the loan pricing result and the key auxiliary result that $v'(z') = v^m(s')$. In the second part - which is in turn divided into three steps - I show how that at the limit we can rearrange the optimal contract problem described in Section 4 to look like the firm problem in the literature of heterogeneous firms with endogenous default.

Step 1: Show that $Q = Q^{\mathbf{zero,b}} = Q^{\mathbf{zero,w}}$. Consider a bank-firm match in state (s, v) that has reached the production phase, and thus have $W(s, v) = W_P(s, v)$. Along the equilibrium path, we can write $W(s, v)$ as:

$$W(s,v) = b - Qb' + \beta(1-\theta)E_{z'|z}[W(s', v'(z'))] + \beta\theta E_{z'|z}[b^m(s')]$$

which we can then rearrange as:

$$Qb' = \Big(b - W(s,v)\Big) + \beta(1-\theta)E_{z'|z}[W(s', v'(z'))] + \beta\theta E_{z'|z}[b^m(s')] \tag{29}$$

To show that competitive pricing holds it is sufficient to show that, on-path, $b = W(s, v)$ and that $W(s', v'(z')) = b^m(s')$. To do so, we first need to show that the promised values in the relationship match the outside value for the firm: $v(z') = v^m(s')$.

One direction is trivial: $v(s') < v^m(s')$ would violate firm PC.

The other direction requires using the Bank PC. Suppose by contradiction that $v(s') > v^m(s')$. By definition of $v^m(s')$ and $b^m(s')$, we obtain that $W(s', v^m(s')) = b^m(s')$, which is true irrespectively of whether the firm is solvent or insolvent in state s'.

Then $v(z') > v^m(s')$ would imply: $W(s', v(z')) < W(s', v^m(s')) = b^m(s')$, which implies a violation of Bank PC.

This confirms our claim that $v(z') = v^m(s')$. That is, when switching costs go to zero, the only free margins are effectively b' and k', while $v'(z')$ is instead uniquely pin down by those two choices. Furthermore, $W(s', v^m(s')) = b^m(s')$ by definition of $v^m(s')$ a $b^m(s')$.

We can now use these results into Equation (29). Notice that on-path it must be that $v = v^m(s)$, and therefore $W(s, v^m(s)) = b^m(s)$. Furthermore, because we are focusing on the case of solvent firms, we get that $W(s, v^m(s)) = b^m(s) = b$. I will discuss the case of insolvent firms in Section B.3.

Therefore, Equation 29 becomes:

$$
\begin{aligned}
Qb' &= 0 + \beta(1-\theta)E_{z'|z}[b^m(s')] + \beta\theta E_{z'|z}[b^m(s')] \\
&= \beta E_{z'|z}[b^m(s')] \\
&= Q^{zero,b} b'
\end{aligned}
$$

That is: $Q = Q^{zero,b}$.

Finally, we showed that on-path we have that $W(s', v^m(s)) = b^m(s')$, or in the compact notation of Proposition 1, $w'(z') = b^m(s')$. Therefore, $E_{z'|z}[b^m(s')] = E_{z'|z}[w'(z')]$.

This concludes the proof that, when $\psi^{f,b} \to 0$, we obtain $Q = Q^{zero,b} = Q^{zero,w}$.

B.2.1 Second half of the Proof: Competitive Market Formulation

In the previous subsection, we have shown that when switching costs go to zero, $v'(z') = v^m(s')$, and loans are priced competitively.

We now show that in such limiting case, we can rewrite the optimal contract problem as a sequence of short-term optimization problems with a competitive pricing rule, as it is canonical in the literature of heterogenous firms with endogenous default risk (Cooley and Quadrini (2001), Khan, Senga, and Thomas (2014), Ottonello and Winberry (2020)).

Step 2a: Impose $v'(z') = v^m(s')$. By using the result that $v'(z') = v^m(s')$ and that $W(s', v^m(s')) = b^m(s')$, we can drop the two market participation constraints. Because $v^m(s') \geq \underline{v}$ by definition, we can also drop the Exit-PC. We can also drop the distinction in the continuation values depending on whether the separation shock hits or not, as the terms multiplying θ and $(1-\theta)$ converge.

We can thus rewrite the problem in the production phase as:

$$\underbrace{W_P(s,v)}_{\text{Bank value}} = \max_{Q,b',k'} \quad \underbrace{b - Qb'}_{\text{Bank flow}} + \underbrace{\beta E_{z'|z}[b^m(s')]}_{\text{Bank cont. value}}$$

Subject to:

$$\underbrace{D}_{\text{Dividend}} + \underbrace{\beta E_{z'|z}[v^m(s')]}_{\text{Firm cont. value}} \geq v \qquad (\mu\text{: PK})$$

Where the budget constraint is unchanged.

Step 2b: Consider the dual problem. Next, it is useful to consider the *dual* of the problem above, in which the optimal contract is solved from the perspective of the firm.

$$\underbrace{V_P(s,w)}_{\text{Bank value}} = \max_{Q,b',k'} \quad \underbrace{D}_{\text{Dividend}} + \underbrace{\beta E_{z'|z}[v^m(s')]}_{\text{Firm cont. value}}$$

Subject to:

$$\underbrace{b - Qb'}_{\text{Bank flow}} + \underbrace{\beta E_{z'|z}[b^m(s')]}_{\text{Bank cont. value}} \geq w \qquad (\mu^{-1}\text{: PK-Bank})$$

and where $V(s,w) = \max\{V_P(s,w), \underline{v} + \lambda k - w\}$.

Step 2c: Impose $v = v^m(s)$. Finally, we can use the fact that, on path, $v = v^m(s)$ and therefore $W_P(s,v) = b^m(s)$. In the *dual* formulation, this implies $w = b^m(s)$. Furthermore, because we are considering the case of solvent firms, we get $w = b^m(s) = b$.

Therefore, on path, the problem of the firm will always be of this form:

$$\underbrace{V_P(s)}_{\text{Bank value}} = \max_{b',k'} \quad \underbrace{D}_{\text{Dividend}} + \underbrace{\beta E_{z'|z}[v^m(s')]}_{\text{Firm cont. value}}$$

Subject to:

$$Qb' = \underbrace{\beta E_{z'|z}[b^m(s')]}_{\text{Bank cont. value}} \qquad (\mu^{-1}\text{: PK-Bank})$$

where to improve the clarity of the result I also used the fact that, on path, the PK constraint always binds, and dropped the optimal choice of Q, which is implied by the zero-profit loan pricing schedule in PK-Bank.

And where $V(s) = \max\{V_P(s), \max\{\underline{v} + \lambda k - b, \underline{v}\}\}$, where the second maximum term reflects the fact that when a firm exits it will use the proceeds from the capital sale to repay the debt, but cannot make liable to repay more than that. This is implied by the Firm-Exit PC in the main problem.

The formulation above is then virtually equivalent to the models in the endogenous default literature cited above, where the firm chooses debt and capital only, with the objective of maximizing its current dividend and future value, under the constraint that creditors will offer credit conditions Q to break-even (Q is typically called an *endogenous pricing schedule*).

The only difference with the models in this literature would be the appearance of the evergreening term

Capital Loss term appearing. This term disappears when the firm is solvent. When the firm is insolvent, it captures the fact that ongoing lenders could evergreen the past loan.

B.3 Insolvent Firms and Evergreening

Recall now Proposition 3 from Section 5.4:

Proposition 3: *Suppose that switching costs are zero ($\psi^{f,b} \to 0$). Consider a firm in state $s = (z,b,k)$ that is insolvent ($\overline{b}(s) < b$). Then, under loan evergreening, the optimal contract prescribes that the firm receives a subsidized loan:*

$$Q = \underbrace{\beta \frac{E_{z'|z}[b^m(s')]}{b'}}_{Q^{zero,b}} + \underbrace{\frac{b - b^m(s)}{b'}}_{Evergreening} \tag{30}$$

Furthermore, the loan subsidy equals the capital loss the bank would incur if selling the loan on the secondary market:

$$\underbrace{Qb' - \beta E_{z'|z}[b^m(s')]}_{Loan\ Subsidy} = \underbrace{(b - b^m(s))}_{Capital\ Loss} \tag{31}$$

When a firm is insolvent ($b > \overline{b}(s)$), then it must be that on-path $w = \overline{b}(s) < b$. Reconsider now the promise keeping constraint of the dual problem analyzed in the proof of Proposition 1:

$$b - Qb' + \beta E_{z'|z}[b^m(s')] \geq w$$

Since $w = \overline{b}(s)$, the b terms do not simplify in the promise-keeping constraint, and an Evergreening loan pricing schedule appears:

$$Qb' = \beta E_{z'|z}[b^m(s')] + (b - \overline{b}(s))$$

B.3.1 Bankruptcy costs within the relationship

In this subsection, I formalize a version of the model where bankruptcy costs occur not only when a firm starts new relationships, but also when it continues borrowing from the same bank, but is insolvent. I labeled as γ^{sep} the exogenous fraction of the value $b^m(s)$ which is actually recovered by the old bank when an insolvent firm separates. I will now label as γ^{cont} the equivalent fraction in continuation.

As briefly discussed in Section 4, this extension is important both for quantitative and theoretical reasons. For quantitative reasons, having *bankruptcy costs* only in case of separation imply that such financial friction only bite with probability θ, the exogenous separation rate. When θ is small, the model converges to the first best case. From a theoretical viewpoint, the proof that when switching costs go to zero the model converges to a competitive market only holds if either $\gamma^{sep} = 1$ or $\gamma^{sep} = \gamma^{cont}$, as both conditions ensure that there is no intrinsic difference between the old and the new bank, implying that $W(s, v^m(s)) = b^m(s)$.

I include such bankruptcy costs by assuming that when a firm is insolvent (in the sense that $b > \overline{b}(s)$), then even if the match continues the bank has to endure restructuring/bankruptcy costs proportional to the ones it would undergo if there was a separation. For generality, I will allow the two costs to be different (γ^{sep} and γ^{cont}), but in practice I calibrate the two to be equivalent, which is needed for my key convergence theorems.

Formally, I first define $W_P(s,v)$ exactly as in the main text in Section 4.

Then, I define

$$W_{ex.\gamma}(s,v) = \max\{W_P(s,v), \lambda k - (v - \underline{v})\}$$

this was called $W(s,v)$ in the main text, and "$excl - \gamma$" denotes the fact that this is the value of the bank gross of the bankruptcy costs.

When a firm starts with a new bank, the new bank value function is $W_{ex.\gamma}$, since the bankruptcy costs are weighting on the old bank. Therefore, we can use $W_{ex.\gamma}$ to compute the maximum amount that the firm can raise from a new bank $\overline{b}(s)$:

$$\overline{b}(s) = W_{ex.\gamma}(s, \underline{v})$$

Therefore we obtain, as before:

$$b^m(s) = \begin{cases} b & \text{if } \overline{b}(s) \geq b \\ \gamma^{sep}\overline{b}(s) & \text{if } \overline{b}(s) < b \end{cases} \tag{32}$$

Notice that this means that the value of the bank in case of separation, when the firm is insolvent, is:

$$b^m(s) = \gamma^{sep}\overline{b}(s) = \gamma^{sep} W_{ex.\gamma}(s, \underline{v}) = W_{ex.\gamma}(s, \underline{v}) - (1 - \gamma^{sep})\overline{b}(s)$$

Now, we assume that if the match is not separated, but $b > \overline{b}(s)$, then the bank must pay a cost $(1 - \gamma^{cont})\overline{b}(s)$.

Therefore, the true value to the bank is:

$$W(s,v) = \begin{cases} W_{ex.\gamma}(s,v) & \text{if } \overline{b}(s) \geq b \\ W_{ex.\gamma}(s,v) - (1-\gamma^{cont})\overline{b}(s) & \text{if } \overline{b}(s) < b \end{cases} \tag{33}$$

We can rationalize this as a penalty which is proportional to the market value of the outstanding debt to the firm (z,b,k).

It is now obvious to see that if $\gamma^{cont} = \gamma^{sep}$, we obtain that $W(s,v^m(s)) = b^m(s)$ also in states in which the firm is insolvent, a fact that is used in the Proof of Proposition 2. When the firm is solvent, this is true irrespectively of the choice of $\{\gamma^{cont},\gamma^{sep}\}$. When the firm in insolvent, then we have

$$W(s,v) = W_{ex.\gamma}(s,v) - (1-\gamma^{cont})\overline{b}(s)$$

and therefore, if $v = v^m(s) = \underline{v}$, we get:

$$\begin{aligned} W(s,\underline{v}) &= W_{ex.\gamma}(s,\underline{v}) - (1-\gamma^{cont})\overline{b}(s) \\ &= \overline{b}(s) - (1-\gamma^{cont})\overline{b}(s) \\ &= \gamma^{cont}\overline{b}(s) \end{aligned}$$

So that $W(s,\underline{v}) = b^m(s) \iff \gamma^{cont} = \gamma^{sep}$.

Furthermore, we can now obtain more transparently the result that relationships are no panacea for firm default. Indeed, if a firm is insolvent in state s, then it must be that *on path* the bank gets no more than $b^m(s)$:

$$b > \overline{b}(s) \implies W(s,v) \leq b^m(s)$$

This follows because *on path* we have $v \geq \underline{v}$, and therefore $W(s,v) \leq W(s,\underline{v}) = b^m(s)$.

B.4 Optimal Contract with Aggregate Shocks

When we introduce aggregate shocks, the following elements of the problem described in Section 4 have to be changed. (i) The value functions W and W_P are now dependent on time, and so are the outside value functions b^m and v^m that they imply; (ii) all state-contingent promises v' have to be chosen not only for each realization of the idiosyncratic shock z', but also for each realization of the aggregate shock, here generically identified by ε', in addition, there is now a PC for each combination of idiosyncratic and aggregate shock; (iii) discount rates are no longer equal to β, but the stochastic discount factor $\Lambda_{t+1}(\varepsilon')$ appears, which depend on the realization of the aggregate shock; (iv) all values are defined in units of current-period final consumption good, therefore an inflation term appears when evaluating the legacy nominal debt b, which was chosen in real terms in the previous period; (v) also the remaining prices p_t and w_t are no longer in steady-state.

$$W^t(s,v) = \max\{W_P^t(s,v), \lambda k - (v - \underline{v})\}$$

Production Phase. If a bank-firm match reaches the production phase with state $s = (z,b,k)$ and promised value v, then the problem of the bank is the following:

$$\underbrace{W_P^t(s,v)}_{\text{Bank value}} = \max_{Q,b',k',\{v(z',\varepsilon')\}} \underbrace{\frac{1}{\Pi_t}b - Qb'}_{\text{Bank flow}}$$
$$+ \underbrace{(1-\theta)E_t[\Lambda_{t+1}(\varepsilon')W_P^{t+1}(s,v'(z',\varepsilon'))] + \theta E_t[\Lambda_{t+1}(\varepsilon')b_{t+1}^m(s')]}_{\text{Bank cont. value}}$$

Subject to:

$$\underbrace{D}_{\text{Dividend}} + \underbrace{(1-\theta)E_t[\Lambda_{t+1}(\varepsilon')v(z',\varepsilon')] + \theta E_{z'|z}[\Lambda_{t+1}(\varepsilon')v_{t+1}^m(s')]}_{\text{Firm cont. value}} \geq v \quad (\mu\text{: PK})$$

$$\underbrace{v(z',\varepsilon')}_{\text{Promised value}} \geq \underbrace{\underline{v}}_{\text{Exit value}} \quad (g(z',\varepsilon')\text{: PC-Exit Firm})$$

$$\underbrace{v'(z',\varepsilon')}_{\text{Promised value}} \geq \underbrace{v_{t+1}^m(s')}_{\text{Firm outside value}} - \underbrace{\psi^f}_{\text{Switching cost}} \quad (\eta(z',\varepsilon')\text{: PC-Mkt Firm})$$

$$\underbrace{W^{t+1}(s',v'(z',\varepsilon'))}_{\text{Implied bank value}} \geq \underbrace{b_{t+1}^m(s')}_{\text{Firm outside value}} - \underbrace{\psi^b}_{\text{Switching cost}} \quad (q(z',\varepsilon')\text{: PC-Mkt Bank})$$

Where the budget is:

$$\underbrace{p_t y^*(z,k,w_t) - f}_{Output} = \underbrace{(k' - (1-\delta)k + \Phi(k,k'))}_{Capex} + \underbrace{(\frac{1}{\Pi_t}b - Qb')}_{Net\ Repayment} + \underbrace{D(1+\tau\mathbb{1}_{D<0})^{-1}}_{Dividends} \tag{34}$$

B.5 Discussion of Model Ingredients

In this section, I lay out and discuss three key assumptions in the model, which clarifies the role played by each of them. This exercise is also useful to illustrate how a different combination of these three assumptions would map into different models in the literature. This discussion is summarized in Table 37.

- *Assumption 1: debt enforceability.* Non-contingent debt b is enforceable in court, and firms with sufficient financial capacity must repay it even they switch to a new bank.
- *Assumption 2: bankruptcy costs.* When a firm is insolvent, there is a deadweight loss γ^{sep}.
- *Assumption 3: switching costs.* Contingent promises $\mathscr{C} = \{Q_t(h^t), b_{t+1}(h^t), k_{t+1}(h^t), l_t(h^t), d_t(h^t), Exit_t(h^t)\}_{h^t, t>t_0}$ are not enforceable in court, but starting a new relationship entails switching costs $\psi^{f,b}$.

I will now discuss these assumptions one-by-one. On a high-level, the combination of Assumption 1 and 2 (market incompleteness and costly bankruptcy), with zero switching costs lead to a canonical firm dynamics model with competitive lenders offering zero-profit loan pricing schedules in each period, as in Cooley and Quadrini (2001), Khan, Senga, and Thomas (2014), Ottonello and Winberry (2020). Introducing assumption 3 generalizes these models. When switching costs are zero, we retrive exactly the setup with firm dynamics and competitive lenders. Instead, when switching costs are positive, lending relationships emerge, meaning that long-term contracts can be sustained and used as a tool to alleviate financial frictions and misallocation. In summary, this is not a model that aims to explain financial frictions. On the contrary, it takes them as given in one of their most popular format, and studies how agents can use the presence of another market imperfections, the switching costs, to alleviate them.

Assumption 1. Without Assumption 1, the entrepreneur can always renege on the outstanding debt and start a new relationship. In such a case, debt is not sustainable in equilibrium, unless there is some cost in reneging (which we can think of as a bankruptcy cost or as a switching cost, the interpretation I use in Table 37). Models such as Albuquerque and Hopenhayn (2004), Cooley, Marimon, and Quadrini (2004), Rampini and Viswanathan (2010), Kovrijnykh (2013) assume that all debt contracts are not enforceable, so entrepreneurs can always renege. A common assumption in this literature is that entrepreneurs can start a new relationship using a fraction of the capital they previously had, a setup which leads to an endogenous collateral constraint.[32] The focus of those papers is

[32] In some papers in this literature, switching is not an option. Leaving the relationship entails shutting down the firm and consuming the stealed capital (Albuquerque and Hopenhayn (2004), Kovrijnykh (2013)). In this case, we could interpret the switching cost as the lost ability to operate the firm.

to study how financial frictions can originate due to the limited commitment to repay of the entrepreneur. Usually, in these models, there is no exogenous risk between the time in which the capital is lent to the entrepreneur and the time in which the loan is supposed to be paid back. It is the non-enforceability of debt which limits the borrowing capacity of the firm and leads to misallocation in the economy. Therefore, in these models, default never occurs on-path.

Assumption 2. Suppose instead that we make Assumption 1. The firm and the bank can trade with a traditional non-contingent debt contract, and this contract is enforceable in court.[33] Because debt is non-contingent, if there is risk the firm will be unable to repay in some states,[34] and a critical choice has to be made regarding what happens in such states of insolvency. The key intuition is that in the absence of any type of *bankrupcy costs*, such insolvency is not problematic for the firm and the bank. On the contrary, the predicted outcome of the model would be that the firm would promise to pay back a very large amount, but ex-post it would almost always be insolvent, and debt is renegotiated to a lower amount. In this way, non-contingent debt becomes a contingent instrument. Therefore, if dealing with insolvency is costless, then the firm will never be constrained. Since at least Cooley and Quadrini (2001), the literature has recognized that if instead being insolvent is costly [35] (for example because of renegotiation or monitoring costs), then firms would avoid borrowing extreme amounts, because when default probabilities increase lenders would charge very high lending rates to compensate not only for the amount lost in renegotiation, but also for the deadweight loss. Therefore, the prediction of models with such an insolvency cost is that firms with limited internal funds will be financially constrained.

Assumption 3. Suppose we make assumptions 1 and 2, but we ignore switching costs. Then, we retrive a model with firm dynamics and competitive lenders offering zero-profit loan pricing schedules on a period-by-period basis. We can think of these models as the limiting case of my model when $\psi \to 0$. The literature using this type of setup does not use the tools from dynamic contract theory. The reason is that the choice of future promises becomes redundant when $\psi = 0$, because no promise other than the zero-profit loan pricing schedule is sustainable, and so the optimal contract collapses to choosing in each period the

[33]Such enforceability, in practice, means that if the firm does not repay the debt, then the firm is shut down and the lender can seize and sell the asset of the firm to recover the amount lent. This is the typical assumption made in a larger firm dynamics literature, such as in Cooley and Quadrini (2001), Khan, Senga, and Thomas (2014), Ottonello and Winberry (2020).

[34]This is obvious in a one-period setting: default occurs when output falls below debt. Instead, in recursive settings, the firm can usually rollover debts, so that it can repay its legacy debt using not only its cash-on-hand (output and possible proceeds from selling capital) but also through new borrowing. In this context, default occurs when the firm is unable to raise sufficient funds from new lenders to repay the legacy debt.

[35]Khan, Senga, and Thomas (2014), Ottonello and Winberry (2020) assume that when the firm is insolvent, it must exit the economy, a fraction of the capital is lost, and the lenders recover the remaining fraction. This is an extreme type of insolvency cost. Instead, Cooley and Quadrini (2001) assumes that debt is renegotiated and brought down to the highest possible amount that allows the firm to repay, but this procedure involves a fixed cost for lenders.

best point on the zero-profit loan pricing schedule. More formally, I show in Appendix B.2 that the combination of promise-keeping constraints and incentive-compatibility constraints collapses to a zero-profit lending condition.

Assuming that $\psi = 0$ is probably a good benchmark for the bond market. However, the majority of firms actually borrow from banks, with very sticky and long-term relationships. Furthermore, zero-profit loan pricing implies that lending rates should respond one-for-one to monetary policy rate and firm default risk, while empirical evidence shows that these pass-through are limited within banking relationships, suggesting that some deviations from the competitive lending benchmark are needed to understand these relationships.

Table 37: Summary of Key Assumptions and Related Literature

Debt enforceable?	Bankruptcy costs?	Switching costs?	Model outcome	References
No	irrelevant	0	No borrowing can be sustained.	Not interesting.
No	irrelevant	>0	Switching costs can help sustain some borrowing. In models with limited enforcement, typically firms lose a fraction of capital if they switch, leading to an endogenous collateral constraint.[a]	Cooley, Marimon, and Quadrini (2004) Albuquerque and Hopenhayn (2004), Rampini and Viswanathan (2010) , Kovrijnykh (2013)
Yes	0	irrelevant	The first-best can be achieved. Debt is set very high, and ex-post renegotiated most of the times to prevent insolvency. Debt effectively becomes contingent.	Not interesting.
Yes	>0	0	Lenders provide zero-profit loan pricing schedules each period. Firms without much equity are constrained (e.g., young firms).	Cooley and Quadrini (2001) Khan, Senga, and Thomas (2014) Ottonello and Winberry (2020) . Nested in this paper.
Yes	>0	>0	Lenders break-even in a relationship ex-ante, but not necessarily on each loan. Insurance can emerge (individual loans can look cheap or expensive).	This paper.
Yes	>0	infinite	All firms achieve first-best because we have de-facto complete markets (this conclusion might depend on how we set up the bankruptcy cost)	Not interesting. Nested in this paper.

[a] In some papers in this literature, switching is not an option. Leaving the relationship entails shutting down the firm and consuming the stealed capital (Albuquerque and Hopenhayn (2004), Kovrijnykh (2013)). In this case, we can interpret losing the ability to operate the firm as a switching cost.

B.6 Solution Algorithm

I divide the discussion of the numerical solution into two parts. First, I outline the algorithm I use to solve for the value and policy functions, given the steady-state prices. Second, I outline the solution to the steady-state prices, which is relatively straightforward. Third, I outline the algorithm to solve for the transition to an aggregate shock, which extends the setup in Boppart, Krusell, and Mitman (2018) to an environment with state-contingent contracts.

B.6.1 Value and Policy Functions

Before discussing the solution algorithm, it is useful to recall the optimality conditions of the contract.

FOC-B:

$$\theta\Big[\underbrace{\frac{\partial E_{z'|z}[b^m(s')]}{\partial b'}}_{\text{Improves loan sale value}}+\underbrace{\mu E_{z'|z}[\frac{\partial v^m(s')}{\partial b'}]}_{\text{Worsens firm market cont. value}}\Big]$$
$$-(1-\theta)\Big[\underbrace{E_{z'|z}[\eta(z')\frac{\partial v^m(s')}{\partial b'}]}_{\text{Relax Firm PC}}+\underbrace{E_{z'|z}[q(z')\frac{\partial b^m(s')}{\partial b}]}_{\text{Tightens Bank PC}}\Big]=0 \tag{35}$$

FOC-K:

$$\underbrace{\beta(1-\theta)E_{z'|z}[\frac{\partial W(s',v'(z'))}{\partial k'}]}_{\text{Improves bank value tomorrow}}+\underbrace{\beta\theta\frac{\partial E_{z'|z}[b^m(s')]}{\partial k'}}_{\text{Improves loan resale value}}+\underbrace{\mu\beta\theta E_{z'|z}[\frac{\partial v^m(s')}{\partial k'}]}_{\text{Improves Firm Market value tomorrow}}$$
$$=\underbrace{1+\Phi_{k'}(k,k')}_{\text{Effective cost of capital}}+\beta(1-\theta)\Big[\underbrace{E_{z'|z}[\eta(z')\frac{\partial v^m(s')}{\partial k'}]}_{\text{Tighten Firm-Mkt PC}}-\underbrace{E_{z'|z}[q(z')\Big(\frac{\partial W(s',v'(z'))}{\partial k'}-\frac{\partial b^m(s')}{\partial k'}\Big)]}_{\text{Relax bank PC}}\Big] \tag{36}$$

FOC- V

$$\underbrace{\frac{\partial W(s',v'(\tilde{z}))}{\partial v'(z')}}_{\text{Bank value loss}}+\underbrace{\mu}_{\text{Firm value improvement}}$$
$$+\underbrace{\eta(z')}_{\text{Relax Firm-Mkt PC}}+\underbrace{g(z')}_{\text{Relax Firm-Def PC}}+\underbrace{q(z')\frac{\partial W(s',v'(z'))}{\partial v'(z')}}_{\text{Tighten Bank PC}}=0$$

First of all, relabel $\frac{\partial W(s',v'(\tilde{z}))}{\partial v'(z')}=-\mu'(z')$.

The **compact FOC-V** is:

$$\mu + \underbrace{\eta(z')}_{\text{Relax Firm-Mkt PC}} + \underbrace{g(z')}_{\text{Relax Firm-Def PC}} = \mu'(z') + \underbrace{q(z')\mu'(z')}_{\text{Tighten Bank PC}}$$

We have four cases.

1. All PC slack: $\mu'(z') = \mu$
2. PC Default Binding: $\mu'(z') = \mu + g(z')$ (we do not care about g in this case since it does not affect any FOC).
3. Firm Market-PC binding: $\mu'(z') = \mu + \eta(z')$, we get an increase in μ.
4. Bank Market-PC binding: $\mu'(z') = \frac{1}{1+q(z')}\mu$, we get a decline in μ.

The high-level solution strategy is the following.

1. **Step 1: Guess**. Guess $W(z,k,v)$. A good starting guess is: $W^{fb}(z,k) - v$.
2. **Step 2: Preliminaries**. Find outside options $v^m(z,b,k)$ and $b^m(z,b,k)$, compute derivatives wrt b, k and v of continuation values.
3. **Step 3: Pre-backward step.** For each $\mu \in (\tau, 1)$, solve for $v'(z')$ such that $\mu'(z') = \mu$. Then, $\forall b', k'$, impose PC and compute the $v'(z')$ that accounts for PC. Then, compute μ' at such value, and find the Lagrange multiplier associated with (z', b', k'), using the four cases described above.
4. **Step 4: Core backward step.** For each (z, μ)
   (a) Solve for $b'(k', \mu)$ using FOC-B
   (b) Solve for k' and v':
      i. First, solve for $v'(z', k', \mu)$ using the results from step 2, accounting for all PC.
      ii. Then, use $b'(k', \mu)$ and $v'(z', k', \mu)$, and the Lagrange multipliers in FOC-K to find the solution for $k'(z, \mu)$.
      iii. Use PK to solve for $v(z, k, \mu)$
      iv. Now invert numerically $v(z, k, \mu)$ to obtain $\mu(z, k, v)$.
5. **Step 5: Reorganize solution at production stage**. At this point, for each (z, k, v) we can:
   - solve for μ and then solve for the policy functions $b'(z, k, \mu)$, $k'(z, k, \mu)$ and $v'(z, k, z', \mu)$

- Use the policy functions above to obtain $W_P(z,k,v)$

6. **Step 6: Liquidation choice.** Find states in which liquidation is optimal, and obtain $W(z,k,v)$. Adjust μ to 1 in states where liquidation is optimal.

7. **Step 7: Check convergence.** Check if the new $W(z,k,v)$ is close enough to the guess. If not, update $W(z,k,v)$ and start again from Step 2.

B.6.2 Steady-State

Once the algorithm outlined in the previous subsection is set up, solving for the steady-state is straightforward. The only price we need to solve for is the wage w_{ss} that guarantees equilibrium in the labor market:[36] $\mathscr{L}^D(w) = L^s$.

Once we have found the wage that clears the labor market, we can solve the firms' problems and find the ergodic firm distribution, which will deliver aggregate output and investment. From these, we can compute aggregate consumption, which is constant. That level of consumption is then the equilibrium consumption rate when $Q = \beta$.

B.6.3 Aggregate Shock: Extension of Boppart, Krusell, and Mitman (2018)

Solving for the response to an aggregate shock, such as a monetary policy rate hike, in an economy with heterogeneous agents is notoriously challenging, as the state-space is infinitely dimensional. The frontier methods rely on a sequence-space approach, such as in Boppart, Krusell, and Mitman (2018) and, more recently, in Auclert et al. (2021). This relies on studying the perfect foresight transition back to the steady-state of an economy that gets hit by to a one-time unexpected aggregate shock ("MIT shock").[37]

What makes my framework unique compared to the contexts where these methods have been applied is that banks and firms write contracts that are state-contingent. This creates the following tension. If the ex-ante probability of the arrival of the aggregate shock is positive, then the economy will not be in steady-state since it is continuously hit by shocks, and we need to keep track of the infinitely-dimensional state of the economy (this could be done with approximations, as in Krusell and Smith (1998)). Working with MIT shocks allows to

[36] In practice, as discussed in the calibration in Section 6.1.2, I follow Kaplan, Moll, and Violante (2018) and Ottonello and Winberry (2020), and calibrate at each step of the estimation process ψ_L such that the employment rate in steady-state is 60% ($L^s = 0.6$).

[37] In contemporaneous work, Auclert, Rognlie, Straub, and Tapák (2024) provide an extension of the SSJ method to accomodate complete markets with respect to aggregate risk. In their approach, tractability is preserved leveraging the result that marginal utilities of all agents would move in proportion after an aggregate shock. However, this approach goes all the way to assuming complete markets - or simple forms of incomplete markets -, and cannot accomodate my environment where market incompleteness arises endogenously from partecipation constraints. Relying on the framework in Boppart, Krusell, and Mitman (2018), while computationally less efficient than SSJ, guarantees the flexibility needed to accomodate state-contingent contracts with limited commitment.

compute exact responses and preserve tractability, but raises a conceptual issue in my context: if the MIT shock is truly unexpected (ex-ante zero probability), do the agents actually contract upon it? This aspect is related to the point risen in Mukoyama (2021) that "MIT shocks imply incomplete markets". I provide a unifying solution to these two challenges by assuming that aggregate shocks happen with a probability that tends to zero. This enables to compute state-contingent contracts and preserve the tractability of MIT shocks.

First, because the probability is non-zero, the choices of contracts for that state are well-defined. The promised value contingent on both idiosyncratic and aggregate shocks $v'(z', \varepsilon')$ is solved following Equation (37), which is derived from the problem with aggregate shocks described in Section B.4. The probability of the aggregate shock cancels out in this equation, and therefore we can solve for $v'(z', \varepsilon')$ even if the probability of the aggregate shock tends to zero.[38]

Second, tractability is preserved because when these shocks happen with a probability that tends to zero, the economy will almost surely remain in steady-state and, more importantly, at such limit the value and policy functions in the steady state do not depend on the transition that follows the arrival of an MIT shock (except for the state-contingent value promises for the realization of the aggregate shock). It is straightforward to show that the values and policies for d, b and k in the problem with aggregate shocks described in Section B.4 converge to the steady-state problem when the aggregate shock arrives with probability that tends to zero.

I now outline the steps in my algorithm. If needed, recall that the optimal contract problem is described in detail in Appendix B.4. All the steps of the algorithm are identical to the shooting algorithm as in Boppart, Krusell, and Mitman (2018), except Step 3, which is new and I will describe in greater detail.

We assume that the economy transitions back to steady-state in T periods.

1. **Step 1: Guess.** Guess sequence for consumption path $\{C_t\}_{t \in \{0,T\}}$.

2. **Step 2: Backward iteration.** Solve backward for prices, and policy and value functions starting at $T-1$.

3. **NEW Step 3: State-contingent promises.**

   (a) solve for $v'(z', \varepsilon')$, the promised values conditional on the realization of the aggregate shock ε', using (37). This has to be done for each state (s, v) and for each realization of the idiosyncratic shock z'.

   (b) solve for the distribution of firms (s, v) on impact .

[38] The probability of the aggregate shock could affect μ, the slope of $W(s, v)$. At the limit when the aggregate shock arrives with zero probability, we can solve for the steady-state μ separately since the value function $W(s, v)$ does not depend on the outcomes following the realization of a shock arriving with zero probability.

- The distribution over $s=(z,b,k)$ is at steady-state.
- The distribution over v is already out-of steady-state at time 0, and can be solved for using the initial steady-state distribution over (s,v) in combination with the promised values found in Step 3 (a).

Remark: this steps requires knowing values functions and prices at $t=0$, which are solved for in Step 2, given the guess in Step 1.

4. **Step 4: Solve forward.** Solve forward for the aggregates by iterating the distribution of firms, using the policies found in Step 2, and starting from the distribution found in Step 3.

5. **Step 5: Update guess.** Restart from (1) using the new $\{C_t\}_{t\in 0,T}$ as guess (or a weighted average with the old guess).

When markets are incomplete with respect to the aggregate shock (e.g. only non-contingent bonds are available), then the financial positions at $t=0$ are known. For example, in Ottonello and Winberry (2020), upon arrival of the shock the distribution of firms is still at steady-state. But in a model with state-contingent contracts, part of the state is the promised value v, which is state-contingent. If we do not properly account for contracting on aggregate shocks, then it is unclear which value v should be used when such shock hits. One alternative could be to use to set $v'(s',\varepsilon')=v'_{ss}(z')$, that is to keep the promised values that were promised for that idiosyncratic contingency but assuming that no aggregate shock had hit. However, this would make little economic sense, and would break the result that at the limit when $\psi^{f,b}\to 0$ the response of my economy to the aggregate shock would be exactly as the response of the competitive economy.

FOC for v'. The optimality condition that pins down $v'(z',\varepsilon')$ in Step 3 for the contingency of the aggregate shock is:

$$\mu+\underbrace{\eta(z',\varepsilon')}_{\text{Relax Firm-Mkt PC}}+\underbrace{g(z',\varepsilon')}_{\text{Relax Firm-Def PC}}=\mu'(z',\varepsilon')+\underbrace{q(z',\varepsilon')\mu'(z',\varepsilon')}_{\text{Tighten Bank PC}} \tag{37}$$

where $\mu'(z',\varepsilon')=-\frac{\partial W^{t=0}(s',v'(z'))}{\partial v'(z')}$, with $W^{t=0}$ being the value function right after the arrival of the aggregate shock at $t=0$, which has been computed in Step 1.

Equation (37) also gives an important economic insight into the nature of insurance against aggregate shocks. Because both firms and banks are owned by the same representative household, the SDF cancels out. Intuitively, there cannot be gains transfering money from the bank to the firm in a recession simply because the SDF is high, since also the shareholders of the bank value dollars highly in that state. However, this insurance can be valuable if financial frictions make firms particularly exposed to the aggregate shock, raising their shadow value of equity. Intuitively, if $\mu'(z',\varepsilon') < \mu(z')$ for most z', this means firms are more constrained when the aggregate shock hits, given the same idiosyncratic state.

B.7 Numerical Solution

To provide deeper insight into the functioning of the model, I display the solution for debt and capital for a firm with a particular state (z,k), varying the level of μ (which is isomorphic to varying v). As μ approaches 1, the policy for capital approaches the first-best solution (which abstracts from financing frictions), instead, low levels of μ correspond to a high shadow value of equity and a substantial gap from the unconstrained solution. In equilibrium, low levels of μ are also associated with higher debt b'.

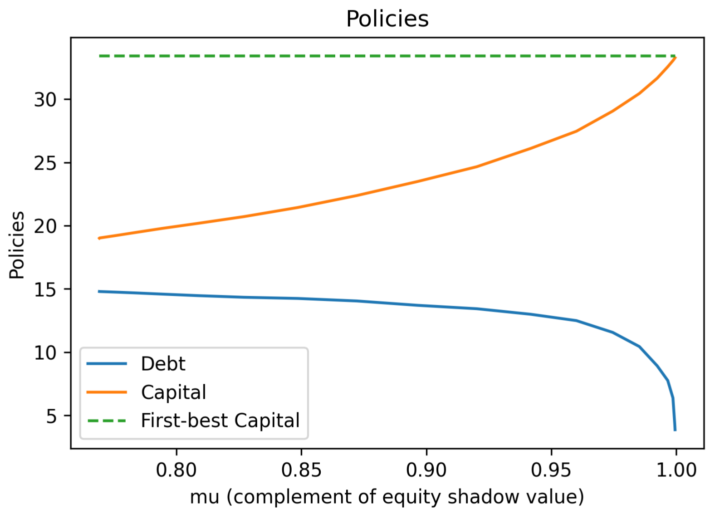


Figure 12: Example of the policy functions for debt and capital depending on firm equity shadow value μ.

B.8 Quantitative Results: insurance vs evergreening

To evaluate the quantitative performance of evergreening, I run my empirical specifications on data simulated from a version of the model with zero switching costs, thus eliminating the insurance mechanism, but where firm bankruptcy is resolved through loan evergreening, as described in Section 5.4. The model with evergreening delivers pass-through roughly consistent with the data, but performs very poorly in delivering history-dependence, as documented in Table 38.

	Data	Model Risk-Sharing	Model Evergreening
	$Spread_t$	$Spread_t$	$Spread_t$
PD_t	0.258***	0.571***	0.209***
	(0.068)	(0.009)	(0.002)
PD_0	0.750***	0.327***	0.047***
	(0.057)	(0.008)	(0.001)

Table 38: PD_0 in the data is the probability of default at the onset of the relationship. Data are organized at the firm-bank-month level. First column: data. Second column: estimated model with positive switching costs. Third column: model with zero switching costs and with resolution of bankruptcy through loan evergreening.

B.9 Model Estimation

To estimate the model, I use data from both the Credit Registry and Orbis. The Credit Registry has a much larger sample, covering the universe of corporate borrowers, and has detailed information on lending conditions, firm default risk, and relationships with banks. Therefore, I use this data to measure the moments on the financial side of the firm. On the real side, the credit registry has only self-reported information on sales and employment. Therefore, I use data from Orbis, which covers a smaller sample of 91,011 firm-year observations, but has accurate information for all the relevant variables on the real-side of the firm: capital, sales, profits, employment and dividends.

B.9.1 Elasticity of Default Risk to Sales

To estimate the elasticity of firm default risk to sales, I first organize data at the firm-year level. I compute the average probability of default of a firm assessed during the year by its lenders. Then, I regress changes in default risk from the previous year against changes in the. I obtain a negative elasticity detailed in Table 12. As discussed more thoroughly in Section 6.1.2, this moment is informative on when firms are more constrained, and support the presence of capital adjustment costs. Absent these costs, default risk tends to move positively with sales, since increases in productivity lead to both higher sales and larger investment demand, raising firm leverage and default risk.

B.9.2 TFP Process Estimation

To estimate the TFP process of firms, I rely on Orbis data. I measure TFP by combining data on labor, capital and sales, using the production function of the model: $y = zk^{\alpha}l^{\nu}$, implying $log(z) = log(y) - \alpha log(k) - \nu log(l)$, using sales for y, fixed assets for k and number of employees for l, and the calibrated values of α and ν.

Then, I estimate the TFP process as follows:

$$log(z_{ft}) = \rho_z log(z_{f,t-1}) + \varepsilon_{ft} \tag{38}$$

Equation (38) directly estimates the persistence of the TFP process ρ_z. To estimate its volatility, σ_z, I compute the standard deviation of the residuals ε_{ft}. When estimating (38), I follow the standard practice of instrumenting $log(z_{f,t-1})$ using its lag, which corrects for temporary shocks. Without this correction, we would estimate an even larger σ_z, and a smaller ρ_z.