Collateral and Asymmetric Information in Lending Markets*

Vasso Ioannidou[†], Nicola Pavanini[‡], Yushi Peng[§]

November 2018

Abstract

We study the benefits and costs of collateral requirements in bank lending markets with asymmetric information. We estimate a structural model of firms' credit demand for secured and unsecured loans, banks' contract offering and pricing, and firm default using detailed credit registry data in a setting where asymmetric information problems in credit markets are pervasive. We provide evidence that collateral mitigates adverse selection and moral hazard. With counterfactual experiments, we quantify how an adverse shock to collateral values propagates to credit supply, credit allocation, interest rates, default, and bank profits and how the severity of adverse selection influences this propagation.

JEL-classification:

Keywords: asymmetric information, structural estimation, credit markets

*We thank for useful suggestions Chiara Fumagalli, Bogdan Stacescu, and seminar participants at the Tilburg Structural Econometrics Group, 2018 Columbia Junior Workshop in New Empirical Finance, CREDIT 2018 Conference (University of Venice), 2018 EARIE Annual Conference, University of Zürich.

[†]Lancaster University and CEPR, v.ioannidou@lancaster.ac.uk

[‡]Tilburg University and CEPR, n.pavanini@tilburguniversity.edu

[§]University of Zürich and SFI, yushi.peng@bf.uzh.ch

1 Introduction

A vast theoretical literature studies the benefits and costs of collateral in debt contracts. On the positive side, collateral is argued to increase borrowers' debt capacity and access to credit, by mitigating both ex ante and ex post agency problems arising from asymmetric information in credit markets. Since Stiglitz and Weiss (1981), the theoretical literature motivated collateral as a screening device to attenuate adverse selection (Bester 1985, Besanko and Thakor 1987*a*), and as a way of reducing various ex post frictions such as moral hazard (Boot and Thakor 1994), costly state verification (Gale and Hellwig 1985), and imperfect contract enforcement (Albuquerque and Hopenhayn 2004).[1] On the negative side, apart from limiting borrowers' use of the pledged assets, collateral is often blamed for amplifying the business cycle, through the so called "collateral channel" (Bernanke and Gertler 1989, Kiyotaki and Moore 1997). In fact, appreciating collateral values during the expansion phase of the business cycle fuels a credit boom, while their subsequent depreciation weakens both the demand and supply of credit, leading to a deeper recession. The collateral channel is viewed as one of main drives of the Great Depression (Bernanke 1983), and as an important factor behind the 2007-2009 financial crisis in the United States (Mian and Sufi 2011, 2014).

The extant empirical literature, provides sharp micro-evidence on the impact of collateral on the demand and supply of credit, analyzing each individually by holding the other constant. Several studies show that increases in exogenous collateral values give firms access to more and cheaper credit for longer maturities (Benmelech, Garmaise and Moskowitz 2005, Benmelech and Bergman 2009), while exogenous drops in collateral values lead to higher loan rates, tighter credit limits and lower monitoring intensity (Cerqueiro, Ongena and Roszbach 2016). The associated changes in credit supply are found to have a significant impact on firm outcomes, such as investment (Chaney, Sraer and Thesmar 2012, Gan 2007) and entrepreneurship (Schmalz, Sraer and Thesmar 2017). Changes in collateral values are also shown to induce similar and contemporaneous changes in households' consumption, which further undermine firms' profits, and hence demand and access to credit (Mian and Sufi 2011, 2014).

We contribute to the empirical literature on collateral by bringing the costs and benefits of collateral into a unified micro-founded structural framework. This approach allows us to test key assumptions and predictions of the theoretical literature that underlie the benefits of collateral, and to study how a shock to collateral values affects both the demand and supply of credit in the presence of asymmetric information frictions. We contribute to the literature on three key dimensions. First, by modelling firms' demand for secured and unsecured credit and subsequent loan repayment, we provide micro-founded evidence of the benefits of collateral under both the ex ante and ex post theories, estimating structural parameters that measure the effectiveness of collateral in mitigating both sets of frictions. Second, by also modelling banks' loan supply of both collateralized and uncollateralized loans, we are able to separately quantify the role of credit demand and credit supply within the collateral channel, accounting for their interaction. We do so

[1] There are many other important theoretical contributions on the role of collateral at mitigating information frictions, including Barro (1976) and Hart and Moore (1994). More specifically, Besanko and Thakor (1987*b*) and Chan and Kanatas (1985) consider ex ante frictions, whereas papers on ex post frictions focus on moral hazard (Boot, Thakor and Udell 1991, Aghion and Bolton 1997, Holmstrom and Tirole 1997), imperfect contract enforcement (Banerjee and Newman 1993, Cooley, Marimon and Quadrini 2004), and costly state verification (Townsend 1979, Williamson 1986, Boyd and Smith 1994).

by simulating a counterfactual scenario where the value of the pledged assets deteriorates, and measure the effect of this shock on borrowers' demand and lenders' pricing of secured and unsecured loans, as well as its impact on firms' default and banks' expected profits. Third, by estimating a micro-founded model with both adverse selection and moral hazard, we can study how the effectiveness of collateral as a screening and monitoring device and the propagation of collateral shocks are influenced by the severity of adverse selection.

We estimate our empirical framework using the detailed credit registry data of Bolivia for the period between March 1999 and December 2003. Besides extensive data availability, especially on collateral values and recovery rates, Bolivia provides a good setting for analysis for two main reasons. First, the Bolivian credit market is characterized by deep informational asymmetries between borrowers and lenders, where the informational inefficiencies highlighted by the extant literature are likely to be important. In fact, even for our sample of mostly large and less risky firms, there is very little reliable information other than what is available through the credit registry. This happens because the vast majority of Bolivian firms do not have audited financial statements, the quality of existing financial statements is poor as many firms engage in tax evasion, and capital markets are not well developed, making the banking sector the principal source of debt finance for most of these firms (Sirtaine, Skamnelos and Frank 2004). This aspect is particularly useful in the context of our model, as it minimizes most differences in the available information between the bank and the econometrician. Second, during the period of analysis the Bolivian credit market did not undergo any deregulation wave or phenomena such as loan sales and securitization. Banks in the sample are operating in steady-state under the traditional "originate and hold" model, allowing us to more closely approximate the bank and borrower incentives modeled in the related literature.

On the demand side, we estimate a structural model of borrowers' demand for credit where firms choose their preferred bank, and conditional on this choice they select a secured or unsecured loan and how much to borrow. We model imperfect competition among lenders allowing banks to be differentiated products and borrowers to have preferences for bank characteristics other than the contract terms offered. We also model borrowers' default on these loans. We let borrowers have heterogeneous preferences for loan interest rates and collateral requirements, and we allow their unobserved heterogeneity in price and collateral sensitivity to be jointly distributed with unobservable borrower characteristics that determine whether they default on their loans. This follows the approach of the empirical literature on testing for asymmetric information (Chiappori and Salanié 2000, Einav, Jenkins and Levin 2012), allowing us to test for the empirical relevance of both the ex ante and ex post channels of collateral, and to quantify adverse selection and moral hazard in this market. The first channel predicts a negative correlation between borrowers' sensitivity to collateral and their default unobservables, which implies that riskier firms have greater disutility from pledging collateral than safer ones, and hence determines the extent to which collateral can mitigate adverse selection. The second channel predicts a negative effect of collateral on default risk, which implies that when firms pledge collateral their incentives to default on a loan are reduced, consistent with collateral mitigating moral hazard. Similar to Crawford, Pavanini and Schivardi (2018), we interpret a positive correlation between borrowers' price sensitivity and their default unobservables as evidence of adverse selection, since riskier borrowers are less price sensitive and thus more likely to take credit. Finally, we interpret a positive causal effect of loan interest rate on default as additional evidence of moral hazard.

On the supply side, we allow banks to offer borrower-specific contracts, in the form of secured and unsecured loans, and compete Bertrand-Nash on interest rates to attract borrowers. We let borrowers have private information about their unobservable (to both the lender and the econometrician) default risk, which implies that each bank offers the same interest rate to observationally equivalent borrowers. Specifying banks' borrower-specific profit functions we derive the equilibrium pricing equations for both secured and unsecured loans for each lender, and use these to back out their marginal costs. We then use the combination of demand, default, and supply models to conduct counterfactual policy experiments, where we simulate how shocks to collateral values or the severity of adverse selection influence the demand and supply of credit and banks' expected profits. This allows us to study the propagation of the collateral channel in the presence of asymmetric information, and to investigate how this propagation varies with the severity of the information frictions.

As mentioned above, we estimate our models using loan-level data from the Bolivian credit register. The credit registry includes detailed contract and repayment information on all loans originated in Bolivia. We have data for the period 1999-2003 and focus on commercial loans granted by commercial banks as in Berger, Frame and Ioannidou (2011). This allows us to keep the set of lenders and borrowers homogenous and focus on a class of loans where collateral is (only) sometimes pledged, as predicted by the theoretical literature. This includes instalment loans and single payment loans, which account for 91% (85%) of the total value (number) of commercial loans in our sample.[2] We avoid modelling the evolution of borrower-lender relationships over time, to minimize the asymmetry of information about borrowers' quality between the econometrician and banks, and therefore focus on firms that take a loan for the first time within our sample period. Crucially, these are the borrowers for which information frictions might be most severe, and collateral requirements might be most effective. One challenge we face is that we only observe the loan a borrower finally chooses, but not the whole set of offers available to the borrower. We therefore need to predict the set of contracts that are available to each borrower as well as the interest rate offered. Exploiting multiple lending relationships that each borrower has, we use fixed effects models and a propensity score matching method to predict the available contracts and the missing interest rates. The advantage of using borrower fixed effects is that it controls for borrowers' information that is observable to banks but not to the econometrician. In the estimation of the structural model, we provide an identification strategy to address potential price endogeneity concerns in both our borrowers' demand and default models.

We find evidence consistent with both the ex ante and ex post theories of collateral, and quantify their empirical relevance. Consistent with the presence of adverse selection, we find a positive and significant correlation of 0.45 between borrowers' price sensitivity and their default unobservables, implying that riskier borrowers are indeed less price sensitive and hence more likely to demand a loan than safer borrowers. In accordance with the ex ante theories that collateral mitigates adverse selection, we find a negative and significant correlation of -0.81 between borrowers' sensitivity to collateral and their default unobservables, which suggests that riskier borrowers tend to have a higher disutility from pledging collateral, and are therefore less likely to demand a secured loan compared to safe borrowers, allowing collateral to serve as a screening device. Furthermore, we find that riskier borrowers have a higher marginal rate of substitution of

[2]We do not include mortgage or credit card loans as they are either always secured or always unsecured.

collateral for price – a key assumption in the ex ante theories, which to the best of our knowledge has never been tested before. Consistent with the presence of moral hazard, we also find a positive and significant effect of loan interest rates on default. Our estimates indicate that a 10% increase in loan interest rates raises the average default probability of a loan by 18.1%. Finally, in accordance with the ex post theories that pledging collateral mitigates moral hazard, we find a negative and significant effect of collateral on default, suggesting that on average posting collateral decreases the probability of default by 88.8%.

We use the estimates of our structural model, together with our supply side framework, for counterfactual policy experiments. We simulate the effects of a 40% drop in collateral values on credit supply, credit allocation, interest rates, and banks' profits.[3] This exercise allows us to study the propagation of the collateral channel across various credit, borrower, and bank outcomes. We find that almost 20% of loans become unprofitable under this scenario, which could imply that those loan applications would now be rejected, while the remaining ones experience a 10% increase on average in interest rates, a 15% average reduction in credit demand, and a 19% decrease in bank profits. We further investigate the role of adverse selection and how its severity influences the propagation of collateral shocks. We find that when adverse selection is more severe, it is easier for lenders to achieve separation of safe and risky borrowers using collateral (i.e., collateral becomes a more effective screening device). As a result, stronger adverse selection mitigates the propagation of the collateral channel, making the increases in interest rates and default in response to a 40% drop in collateral value less pronounced. We also find, however, that when adverse selection is high, banks suffer larger drops in expected profits as the use of collateral for screening reduces their profit margins ex ante.

We contribute mostly to three broad strands of literature. First, we provide new supportive evidence of the ex ante and ex post theories of collateral. Existing work provides reduced form evidence consistent with theoretical predictions of both sets of theories. Consistent with the ex post theories that banks require collateral from observably riskier borrowers, several studies document that the incidence of collateral is positively related to observable borrower risk. Evidence for the ex ante theories is instead scarce, as borrowers' unobservable risk is typically not observable to the econometrician and difficult to disentangle from ex post frictions. One exception is Berger, Frame and Ioannidou (2011), who exploit an information sharing feature of the Bolivian credit registry, using borrowers' historical performance that is unobservable to lenders but observable to the econometricians as a proxy of borrowers' private information. Their findings support both sets of theories and indicate that ex post frictions are empirically dominant. The structural approach in this paper allows us to go beyond testing the two sets of motives for pledging collateral to additionally assessing whether collateral is effective in mitigating the associated frictions.

Some papers use borrower-lender relationships to proxy for ex ante asymmetric information, assuming that the length of a credit relationship implies less asymmetric information and hence less need for collateral (Petersen and Rajan 1994, Berger and Udell 1995, Degryse and Van Cayseele 2000). However, a strong

[3] A 40% drop in collateral values is similar in magnitude to drops in collateral values documented in the literature during economic downturns, such as the burst of the Japanese assets price bubble that caused land prices in Japan to drop by 50% between 1991 and 1993 (Gan 2007), the early 30% drop of the Case-Shiller 20-City Composite Home Price Index in the U.S. during the 2007-2009 financial crisis, and the increase in average repo haircut on seven categories of structured debt from zero to 45% between August 2007 and December 2008 (Gorton 2010).

borrower-lender relationship could also reduce the cost of monitoring and state verification problems, accordingly resulting in less ex post asymmetric information. This approach cannot thus disentangle whether less observed collateral in longer borrower-lender relationships is the result of reduced ex ante or ex post asymmetric information. Other studies have used different ways to identify unobserved risk. For example, Gonas, Highfield and Mullineaux (2004) argue that for large, rated, and exchange listed firms asymmetric information is less severe, and show that those firms are less likely to have secured loans. In Berger, Espinosa-Vega, Frame and Miller (2011), the authors take advantage of the adoption of an information-enhancing loan underwriting technology, after which lower collateral incidence is consistent with the ex ante channel. We contribute to the literature by proposing a micro-founded mechanism to incorporate and test for both the ex ante and ex post theories, and by estimating a structural model that allows us to simultaneously quantify the magnitude of adverse selection and moral hazard, and their effects on credit supply.

Second, we contribute to the empirical literature on the collateral channel. One line of papers in this area focusses on how exogenous variation in collateral values influences credit supply and bargaining power in default by exploiting exogenous variation in commercial zoning regulations (Benmelech, Garmaise and Moskowitz 2005), asset redeployability of airline fleets (Benmelech and Bergman 2008, 2009), and regulatory changes affecting creditor seniority (Cerqueiro, Ongena and Roszbach 2016, 2018). Another line of papers in this area explores the effect of the collateral channel on credit supply and firm outcomes, still exploiting exogenous shocks to collateral values, with applications to firms' investment (Chaney, Sraer and Thesmar 2012, Gan 2007), employment (Ersahin and Irani 2018), and entrepreneurship (Adelino, Schoar and Severino 2015, Corradin and Popov 2015, Kerr, Kerr and Nanda 2015, Schmalz, Sraer and Thesmar 2017). Benmelech and Bergman (2011) study instead how drops in collateral values, arising from negative externalities of bankrupt firms on their non-bankrupt competitors, amplify industry downturns. A more recent line of papers in this area also studies the amplifying role of the housing net worth channel during the recent financial crisis. House price appreciation prior to the financial crisis triggered significant increases in existing homeowners' consumer demand and leverage (Mian and Sufi 2011), while the subsequent collapse in house prices during the financial crisis led to decreases in consumer demand, which in turn weakened further the real economy, especially in the non-tradeable sectors (Mian and Sufi 2014). We are closer to the first line of papers in this area, as we focus on the effect of the collateral channel on firms' debt capacity and access to credit. Our structural approach allows us to trace the impact of shock to collateral values, accounting for feedback effects between banks' and borrowers' behavior. Differently from the papers listed above – that exploit identification strategies aiming to hold credit demand or supply fixed – our structural framework can decompose the collateral channel into its demand and supply effects. Moreover, our approach also allows us to capture spillover effects of a shock to collateral values from secured to unsecured loan rates and demand, a channel previously unexplored by the extant literature.

Last, we also contribute to the recent strand of literature on empirical models of asymmetric information using both reduced form and structural methods (Karlan and Zinman 2009, Adams, Einav and Levin 2009, Einav, Jenkins and Levin 2012). Our modelling approach is closest to Crawford, Pavanini and Schivardi (2018), who focus on the interaction between asymmetric information and imperfect competition in the context of Italian unsecured credit lines. We share a similar identification method by combining credit demand for differentiated products and ex post debt performance. We generalize their approach by considering both secured and unsecured loans, allowing for multi-dimensional bank screening through both interest rates and collateral requirements. More generally, we contribute to the growing literature using structural methods from empirical industrial organization to model financial markets, with applications to deposits (Ho and Ishii 2011, Egan, Hortaçsu and Matvos 2017), corporate loans (Crawford, Pavanini and Schivardi 2018), mortgages (Benetton 2017), insurance (Koijen and Yogo 2016), and investors' demand for assets (Koijen and Yogo 2018).

The paper is organized as follows. Section 2 provides a data description and institutional details. In Section 3 we present the structural model. Section 4 describes the econometric framework, including price prediction and identification strategies. The estimation results are presented in Section 5. Section 6 presents the counterfactuals, and Section 7 concludes.

2 Data and Descriptive Evidence

We make use of the data from Central de Informatión de Riesgos Crediticios (CIRC), the public credit registry of Bolivia, provided by the Bolivia Superintendent of Banks and Financial Entities (SBFE). The SBFE requires all formal (licensed and regulated) financial institutions in Bolivia to record information on their loans. We have access to detailed monthly loan-level information for all corporate loans originated by formal financial institutions in Bolivia from 1999 to 2003.[4] For each loan, we have information on the identity of the bank originating the loan, the date of loan origination, the maturity date, the loan amount, the loan interest rate, the type and value of collateral securing a loan as well as ex-post performance information (i.e., overdue payments or defaults). Borrowers information includes a unique identification number that allows to track borrowers across banks and time, an industry classification code, the region where the loan was originated, the borrowing firms' legal structure, current and past bank lending relationships, the borrowers' internal credit rating with each bank, as well as current and past credit history (i.e., overdue payments or default with any bank).

To reduce information asymmetries in the Bolivian credit markets, the SBFE requires banks to share information in the credit registry with other participating institutions. After written authorization from a prospective borrower, a bank can access the registry and obtain a credit report, which contains information on all outstanding loans of the customer for the previous two months. The report includes information on outstanding exposures and past repayment history (outstanding delinquencies and past defaults). Besides the information shared through the registry, banks have limited reliable information about potential borrowers, as during the sample period there was no other comprehensive private credit bureau operating in the country (De Janvry, Sadoulet, McIntosh, Wydick, Luoto, Gordillo and Schuetz 2003), and the vast majority of firms in Bolivia do not have audited financial statements (Sirtaine, Skamnelos and Frank 2004).

The credit registry includes loans from commercial banks as well as other non-bank financial institutions

[4] We also observe data from January 1998 to December 2003. As the type of credit is available after March 1999, we only use the data from March 1999 to December 2003 for the analyses. The data from January 1998 to February 1999 is used to identify borrow-lending relationship.

(e.g., microfinance institutions, credit unions). To keep the set of lenders and borrowers homogenous in terms of financial structure and regulation, we focus exclusively on commercial loans granted by commercial banks. There are several types of commercial credit contracts in the data, including credit cards, overdrafts, installment loans, discount loans, and credit lines. To give a meaningful role to the ex ante and ex post theories of collateral we focus on loan contracts for which collateral is (only) sometimes pledged, as in Berger, Frame and Ioannidou (2011). This includes installment loans and discount loans, which account for 91% (85%) of the total value (number) of commercial loans in our sample. In order to minimize the information asymmetry on borrowers' private information between the econometrician and banks, we follow the literature on testing for asymmetric information (Chiappori and Salanié 2000) and focus only on the firms that enter the formal credit market for the first time, for which banks have no previous record.[5]

Our empirical analysis is divided into two parts, where we make use of two partially different subsamples. First, the estimation of the structural model is conducted using the data for the first main loan that a new firm obtains during the sample period.[6] Second, as we explain in detail in Section 4, we need to predict interest rates for loan contracts not chosen by borrowers. For this exercise, we increase slightly the sample size to achieve higher statistical power for prediction, and enlarge the sample to loans originated within 6 months from each borrower's first loan origination. This larger sample consists of 2,877 loans granted to 1,421 borrowers among which 561 are new borrowers,[7] whereas the first more restrictive sample includes the 561 first loans chosen by 561 new borrowers.

Table 1 Panel A provides summary statistics of the loans in each sample. The average annual interest rate is just above 14% for both samples, and secured loans have lower interest rate than unsecured loans on average. Between 30% to 40% of loans are collateralized. The incidence of collateral is higher when the borrower obtains bank credit for the first time (i.e. in the restricted sample). Over 40% of collateralized loans are secured with immovable assets (real estate). The median collateral value to the loan amount is 1.5 in both samples. The average loan maturity is between 16 and 20 months, whereas the average loan amount is around 150,000 USD. Between 50% to 65% of loans are installment loans, while the rest are single payment loans. 6% of all loans and 2% of first loans are classified as having potential problems or as being unsatisfactory or doubtful. Around 60% of borrowers are corporations, while the rest are mainly sole proprietorships or partnerships. Between 13% to 28% of loans are granted to borrowers who had at least one non-performing loan during the sample period after receiving their first loan, and this will be our definition of defaulting borrower throughout the rest of the paper.[8] We summarize in Panel B monthly bank balance sheet information on household deposits, an important piece of data that we will use in our identification strategy later on. Deposits from households are distinguished into savings and demand deposits with a mean of 62 and 60 millions USD respectively. On average, banks pay 1.26 millions USD as interest on deposits, and the average annualized interest rate on savings deposits is 7 percentage points.

[5]A firm is defined as a new entry in the credit market if the first loan of a firm was originated in the time period from March 1999 to December 2003 without any existing loan from January 1998 to February 1999.

[6]If a borrower has more than one loan in the first month, we consider the loan with the largest amount as the main loan.

[7]The remaining 860 non-new borrowers are those that we know borrowed for the first time between January 1998 and February 1999, but for which we don't have information on loan terms during those months, only from February 1999 onwards.

[8]We use this definition in line with Crawford, Pavanini and Schivardi (2018), as new borrowers take some time to reach the default stage.

Table 1: Summary Statistics of Commercial Loans

Variable	N. Obs	Mean	St. Dev.	N. Obs	Mean	St. Dev.
Panel A: Loan Level	**First Loans**			**Loans in First Six Months**		
Interest Rate	561	14.69	2.55	2,871	14.11	2.54
Secured	211	14.21	2.30	842	13.72	2.56
Unsecured	350	14.97	2.65	2,029	14.28	2.51
Collateral	561	0.38	0.49	2,871	0.29	0.46
Immovable	211	0.41	0.49	842	0.45	0.50
Value-to-Loan Ratio	211	2.87	4.88	842	2.82	8.24
Amount	561	156.01	483.02	2,871	139.60	561.83
Maturity	561	20.53	26.58	2,871	16.87	24.11
Installment	561	0.62	0.49	2,871	0.53	0.50
Bad Credit Rating	561	0.02	0.15	2,871	0.06	0.24
Corporation	561	0.61	0.49	2,871	0.60	0.49
Defaulting Borrower	561	0.13	0.34	2,871	0.28	0.45
Panel B: Bank Level						
Saving Deposit	619	62.17	51.78			
Demand Deposit	619	60.10	46.05			
Deposit Interest Expense	619	1.26	1.02			
Saving Deposit Interest Rate	619	6.99	3.32			
Panel C: Loss Rates						
Loss Given Default	283	0.35	0.46			
Loss from Defaulting Borrower	299	0.05	0.18			

Note: This table summarize new borrowers' first loan. Interest Rate is the annual percentage rate, which is divided into two subgroups: interest rate for secured loans (Secured) and unsecured loans (Unsecured). Collateral is a dummy variable taking the value of one if a loan is secured and zero if it is unsecured. Immovable is a dummy variable taking the value of one if the collateral is immovable (real estate) and zero otherwise. Value-to-Loan Ratio is the ratio of collateral value to the loan amount for secured loans only. The loan Maturity is in months, and loan Amount is in 1,000 USD. Installment is a dummy variable taking the value of one if this is an installment loan and zero if it is a single payment loan. Bad Credit Rating is a dummy variable taking the value of zero if the loan has no overdue payments or is not in default and one otherwise. Corporation is a dummy variable taking the value of one if the borrower is a corporation and zero if it is a sole proprietorship or partnership. Defaulting Borrower is a dummy variable taking the value of one for loans that are granted to borrowers who had at least one non-performing loan in the sample and zero otherwise. Saving Deposit, Demand Deposit, Deposit Interest Expense are in millions of USD. Saving Deposit Interest Rate is the annual percentage rate. Loss Given Default is the loss rate of default loans. Loss from Defaulting Borrower is the loss rate from borrowers who had at least one non-performing loan during the sample period after receiving their first loan.

The loans are originated by 12 commercial banks,[9] half of which are foreign owned. Borrowers are located in 8 different regions.[10] As illustrated in Figure 1, the number of banks that are lending to new borrowers varies significantly across regions. More banks are present in urban areas. For example, in La Paz, the country's capital, all 12 banks originated loans to new borrowers, while in more rural areas such as Potosi, only 3 banks originated loans to new borrowers. Each bank is active across different regions. For example, Banco Nacional De Bolivia and Banco De Credito De Bolivia established new lending relationships in almost all regions, while Banco Do Brasil only granted loans to new borrowers in La Paz. This gives us heterogeneity in borrowers' choice sets of banks depending on their location. In particular, we define a lending market as the region-quarter combination where and when each borrower is making its choice of preferred lender and loan, and all banks actively lending in each market as each borrower's potential choice set. In total, we have 105 region-quarter markets in the sample.

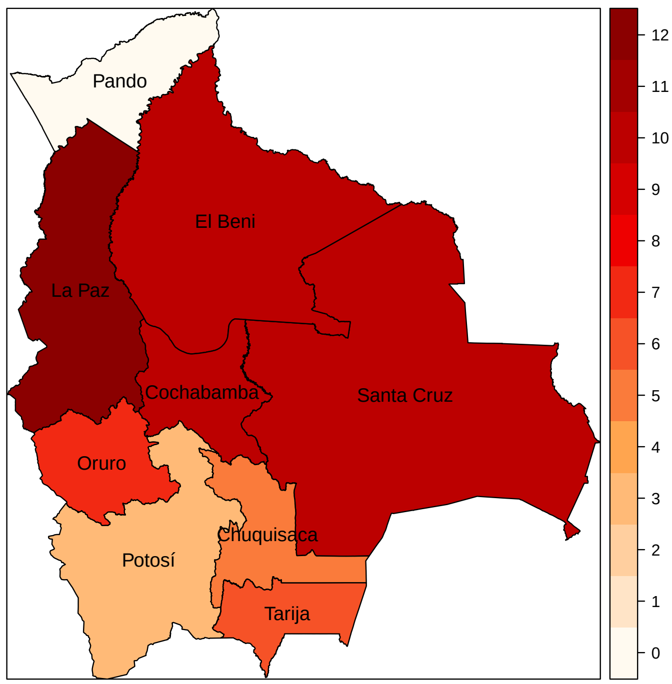


Figure 1: Number of Banks Establishing New Borrow-Lending Relationship across Regions

Note: This figure shows the regions where banks granted loans to new borrowers during 1999 to 2003. The banks are Banco Nacional De Bolivia S. A., Banco Mercantil S. A., Banco De Credito De Bolivia S. A., Banco De La Nacion Argentina S. A., Banco Do Brasil S. A., Banco Industrial S. A., Citibank N.A. Sucursal Bolivia, Banco Santa Cruz S. A., Banco Union S. A., Banco Economico S. A., Banco Solidario S. A., Banco Ganadero S. A.. The regions are Chuquisaca, La Paz, Cochabamba, Oruro, Potosi, Tarija, Santa Cruz and El Beni, U.S.A.

Among the loans granted to new borrowers within the first 6 months, nearly one-third of loans are secured, that is 213. Borrowers compare potential loan offers not only among banks, but also with respect to whether they have to pledge collateral or not. The data suggest that a certain level of discretion exists. For example,

[9]We drop ABN AMRO Bank N.V. as it left the Bolivian market in November of 2000, and before exiting it only granted 4 loans to new borrowers. We also exclude Banco Boliviano Americano S. A. as it failed in May of 1999.

[10]The 8 regions are Chuquisaca, La Paz, Cochabamba, Oruro, Potosi, Tarija, Santa Cruz and El Beni, and foreign markets. El Beni and Cochabamba are considered as one region because there are only five new borrowers in El Beni. No new borrower-lending relationships are observed in the region of Pando.

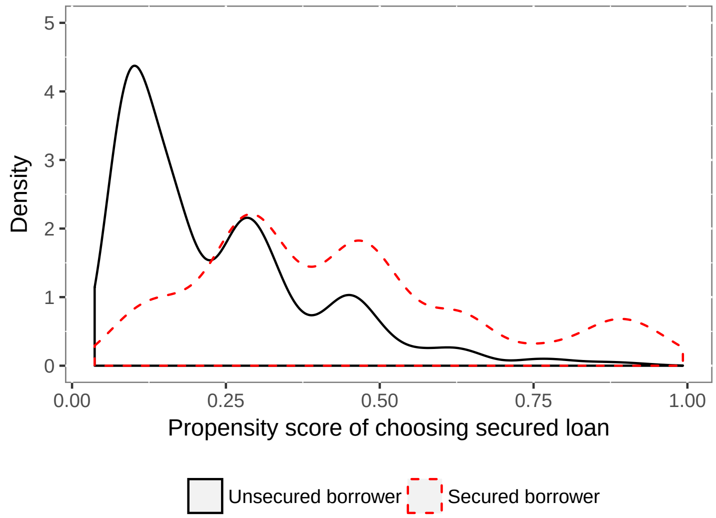


Figure 2: Propensity Score of Choosing A Secured Loan

Note: This figure shows the distributions of the propensity score of choosing a secured loan as opposed to an unsecured loan for borrowers that accepted a secured loan (secured borrower) or an unsecured loan (unsecured borrower). The solid line represents unsecured borrowers, and the dashed line represents secured borrowers. There is a wide range over which the two distributions overlap: A borrower with a propensity score in the overlapping region can become either a secured or an unsecured borrower.

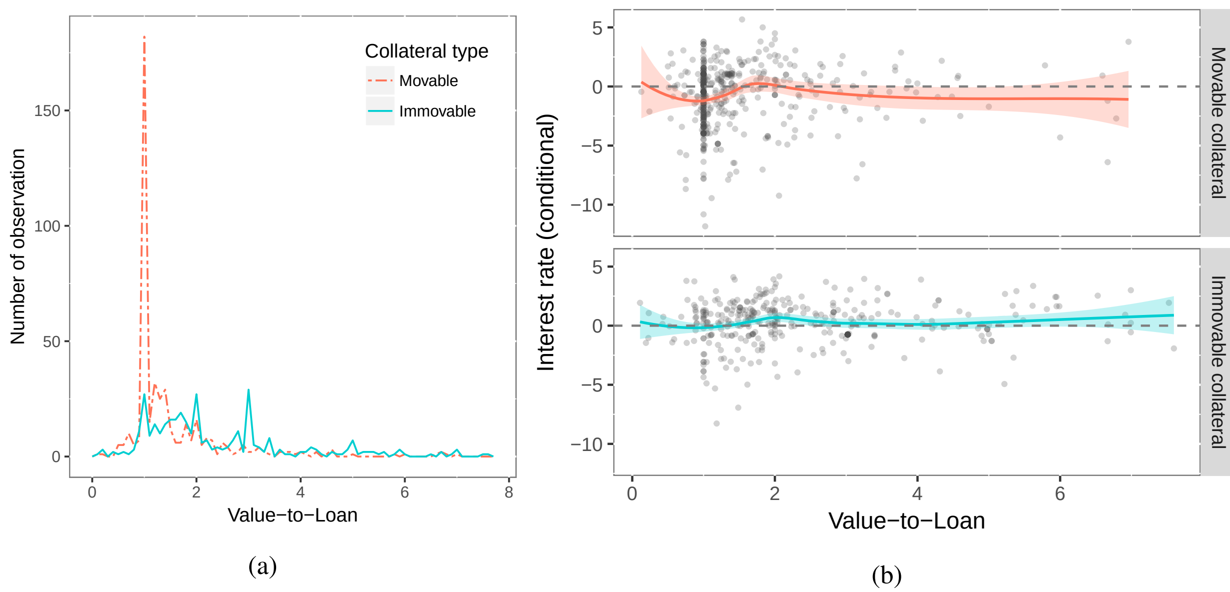


Figure 3: Collateral to Loan Ratio by Types

Note: The two figures illustrate the distribution of collateral to loan value. The collateral to loan ratio is truncated at 8, which means the collateral value is 8 times of the loan amount. There are 7% (2%) of loans with movable (immovable) collateral that have Value-to-Loan ratio above 8. Subfigure (a) depicts the distribution of collateral to loan value for movable and immovable collateral. Subfigure (b) plots the relationship between collateral value-to-loan ratio and interest rate conditional on the loan size, the loan maturity, firms' legal structures and bad credit rating (i.e. the residuals of a linear regression of interest rates on all those variables). Immovable collateral includes real estate. Movable collateral includes long-term deposits, inventory, accounts receivable, bonds, vehicles, tools and equipment, etc. The two fitted lines stand for smoothed conditional means of interest rates and the shadow areas are 95% confidence intervals.

Figure 2 reports the distributions of propensity score for a collateralized loan, that is the probability of taking a secured loan, both for borrowers that take up a secured or an unsecured loan in the data. The two propensity score distributions' overlap in the middle indicates that a wide range of borrowers with similar characteristics are almost equally likely to choose secured or unsecured contracts.[11]

Value-to-loan ratios may be influenced not only by banks' collateral requirements, but also by the borrowers' available collateral. As shown in Figure 3 (a), there is a mass of collateral value-to-loan ratio at one, and for around 60% of the loans the ratio is between 1 and 2. Distinguishing between movable and immovable collateral reveals that if borrowers pledge movable collateral (such as long-term deposits, bonds, inventory, and accounts receivable), the frequency of value-to-loan ratios spikes at one, while if borrowers pledge immovable collateral types (real estate), the value-to-loan ratios exhibit more variation. As immovable collateral is likely to be indivisible, this suggests that variation in value-to-loan ratios may be largely influenced by the type of collateral pledged, which in turn may be influenced not only by the bank collateral requirements, but also by the borrowers' type of available collateral. This is further confirmed in Figure 3 (b), which plots the relationship between interest rates and value-to-loan ratios with smoothly fitted lines and 95% confidence intervals. Conditional on the loan size, the loan maturity, firms' legal structures, and bad credit rating, the value-to-loan ratio has little impact on interest rates for loans secured with immovable collateral, while for loans secured with movable collateral the interest rates are statistically lower from zero at value-to-loan close to one.

It remains an open question whether borrowers in our sample use all of their pleadable assets for the secured loans they take, or if they have any remaining assets that could be pledged if they wanted to take any extra collateralized credit. This is of course an important piece of information for our analysis, because when we simulate a drop in collateral value in our counterfactuals we don't give borrowers the option of pledging additional assets to increase their debt capacity. We justify this assumption with descriptive evidence consistent with borrowers being "collateral constrained". On the one hand, we find that 31% of borrowers who chose as first loan an unsecured one obtain a new unsecured loan within 3 months. On the other hand, we find instead that just 19% of borrowers who chose as first loan a secured one obtain a new secured loan within 3 months. Among this 19%, only 4% use a different collateral type (movable or immovable) compared to the one used for the first loan, while the remaining 96% uses the same collateral type. We just focus on a three months horizon as firms might be acquiring new assets over time, eventually expanding their potential set of pledgeable assets. We interpret this as evidence of firms being collateral constrained, hence almost always using the maximum value of their pleadable assets to take credit. This then allows us to rule out the option for firms to pledge new assets when their pleaded assets drop in value.

One last important piece of information that we derive from our data is banks' recovery rates on unsecured loans in case of default. This data is a key input in the banks' expected profit function that we will define later on. Given that we don't observe directly all defaulting loans' write-offs, mostly due to right censoring in the data, we focus on all delinquent commercial loans in our credit register reaching maturity within

[11] The propensity score is based on bank identity, loan amount and maturity categories, borrower's legal structure, and on whether this is the first loan. In Section 4.1.2 we discuss the propensity score matching in detail. The detailed specification of the propensity score is presented in Table A.2.

the sample period. For these loans, we approximate the recovery rate as one minus the loss rate, which we calculate as the write-off at maturity divided by the amount of the loan. To address the concern that banks usually report write-offs for individual loans with some delay, we only focus on loans that have been persistently classified as non-performing (i.e. past due for at least 30 days) or in default (i.e. in liquidation) for at least 6 months. As shown in Table 1 Panel C, the average loss rate given default is 0.35 and therefore the recovery rate in default is 0.65, in line with estimates in the literature.[12] We define the loss rate from defaulting borrowers, which are those having at least one non-performing loan within the sample period, as the borrower's total amount of write-offs divided by the borrower's total amount of loans granted. This variable will be mechanical smaller than the loss rate given default, as we are simply increasing the size of the denominator from the previous formula by taking into account the borrower's total amount of loans granted. We find in fact an average loss rate from a defaulting borrower of 0.05. We need to construct this variable to match our definition of defaulting borrower that we will later on use in the structural model, where the unit of observation will be an individual loan between a firm and a bank. Accordingly, if on the one hand our defaulting borrower variable is on average actually higher than the default probability over an individual loan, on the other hand this is balanced by the loss rate from defaulting borrower that is on average lower than the loss given default over an individual loan.

3 The Model

3.1 Demand and Default Model

Our modeling approach builds on Crawford, Pavanini and Schivardi (2018). We assume that new borrowers seek credit for an exogenously given amount and maturity combination,[13] and shop around banks that actively lend in their region-quarter looking for the most profitable option. We allow firms to choose not only their preferred bank, but also whether they want to pledge collateral or not, conditional on a bank offering them the option of both a secured and an unsecured loan. Unfortunately, we don't observe firms not taking

[12]The literature suggests that bank loan recovery rates range from 60% to 90%. Loan characteristics, borrower characteristics, and macroeconomic conditions affect the recovery rates. Asarnow and Edwards (1995) use 831 commercial and industrial loans and 89 structured loans made by Citibank over 24 years and find an average recovery of 65% for commercial and industrial loans and 87% for heavily collateralized structured loans. Acharya, Bharath and Srinivasan (2007) report recovery rates of 81.12% for bank loans in the United States for the period from 1982 to 1999. Khieu, Mullineaux and Yi (2012) find the average recovery rate is 84.14% for North American loans in default in the period 1987-2007. Davydenko and Franks (2008) provide information on small firms that defaulted on their bank debt in France, Germany, and the United Kingdom in the years 1996 to 2003. The bank recovery rates are sharply different with median recovery rates of 92% in the United Kingdom, 67% in Germany, and 56% in France.

[13]We will allow firms to choose their preferred loan amount in the counterfactual exercises, as discussed in Section 4.3. However, allowing for endogenous firms' choice of amount and maturity at this stage would substantially complicate the model, as it would require us to assume a set of potential amount and maturity options available to the borrower that we don't observe in the data. Moreover, it would imply that banks could use amount and maturity as additional screening and competitive devices, on top of interest rates and collateral requirements. However, given the non-exclusive nature of these loan contracts, it is less likely that banks would use the loan amount as a screening device, as borrowers can linearize the price schedule by taking multiple loans from various banks. Modeling these margins is challenging and we leave it for future research.

loans, so we are unable to model borrowers' choice of an outside option. More specifically, we let borrower $i = 1, .., I$ in market $m = 1, .., M$, defined as a region-quarter combination, take a loan of type $k = \mathcal{S}, \mathcal{U}$, where $\mathcal{S}$ stands for secured and $\mathcal{U}$ for unsecured, from bank $j = 1, .., J_m$ based on the following indirect utility function, which determines its demand (D):

$$U^D_{ijkm} = \alpha^D_{\mathcal{P}i} P_{ijkm} + \alpha^D_{\mathcal{C}i} C_{ijkm} + X'_{jm} \alpha^D_{\mathcal{X}} + \nu^D_{ijkm}, \tag{1}$$

where P_{ijkm} is the interest rate offered by bank j to borrower i, C_{ijkm} is a dummy indicating whether the loan is secured $\mathcal{S}$ or unsecured $\mathcal{U}$, X_{jm} are bank-market characteristics, and ν^D_{ijkm} are Type 1 Extreme Value distributed shocks. We let $\alpha^D_{\mathcal{P}i}, \alpha^D_{\mathcal{C}i}$ be borrowers' normally distributed heterogeneous preferences for interest rate and collateral, which will depend on borrowers' private information $\varepsilon^D_{\mathcal{P}i}, \varepsilon^D_{\mathcal{C}i}$ (unobserved by banks and the econometrician), and borrowers' observed characteristics Y_i (observed by both banks and the econometrician), as follows:

$$\alpha^D_{\mathcal{P}i} = \overline{\alpha}^D_{\mathcal{P}} + Y'_i \delta_{\mathcal{P}} + \varepsilon^D_{\mathcal{P}i} \quad , \quad \alpha^D_{\mathcal{C}i} = \overline{\alpha}^D_{\mathcal{C}} + Y'_i \delta_{\mathcal{C}} + \varepsilon^D_{\mathcal{C}i} \tag{2}$$

Following the descriptive evidence reported in Section 2, we assume that when choosing a secured loan a firm has no discretion over the type and amount of collateral to pledge, as this is entirely determined by the lender. We model a situation in which the firm presents its pleadable assets to the lender and requires the maximum amount of credit that the lender is willing to grant using those assets as collateral. Hence, we rule out any signaling that the firm might engage in by choosing a specific type and amount of collateral to pledge. We do so to keep the model tractable, and because we don't have data on other potential pleadable assets that each firm might have.

Similarly to demand, we model borrowers' default (F) as being determined by the following indirect utility function:

$$U^F_{ijkm} = \overline{\alpha}^F + \alpha^F_{\mathcal{P}} P_{ijkm} + \alpha^F_{\mathcal{C}} C_{ijkm} + X'_{jm} \alpha^F_{\mathcal{X}} + Y'_i \alpha^F_{\mathcal{Y}} + \varepsilon^F_i, \tag{3}$$

where ε^F_i represents the borrower's private information component, unobserved by banks and the econometrician, that affects their likelihood of repayment. In the spirit of the empirical literature on testing for the presence of asymmetric information (Chiappori and Salanié 2000, Einav, Jenkins and Levin 2012), we let $\varepsilon^D_{\mathcal{P}i}, \varepsilon^D_{\mathcal{C}i}, \varepsilon^F_i$ be distributed according to the following multivariate normal distribution:

$$\begin{pmatrix} \varepsilon^D_{\mathcal{P}i} \\ \varepsilon^D_{\mathcal{C}i} \\ \varepsilon^F_i \end{pmatrix} = \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\mathcal{P}} & \rho_{\mathcal{PC}} \sigma_{\mathcal{P}} \sigma_{\mathcal{C}} & \rho_{\mathcal{P}F} \sigma_{\mathcal{P}} \\ \rho_{\mathcal{PC}} \sigma_{\mathcal{P}} \sigma_{\mathcal{C}} & \sigma^2_{\mathcal{C}} & \rho_{\mathcal{C}F} \sigma_{\mathcal{C}} \\ \rho_{\mathcal{P}F} \sigma_{\mathcal{P}} & \rho_{\mathcal{C}F} \sigma_{\mathcal{C}} & 1 \end{pmatrix} \right). \tag{4}$$

The demand and default model allows us to disentangle the adverse selection and moral hazard channels. The adverse selection channel is identified through the covariance matrix of unobservables, which captures the relations of unobserved default risk and firms' unobservable preference for interest rate and collateral in demand. The moral hazard channel is identified through the direct impact of interest rate and collateral on default, given that the selection has been accounted through unobservables.

We interpret a positive correlation between unobservables determining price sensitivity and default $\rho_{\mathcal{P}F} > 0$ as evidence of adverse selection, as riskier borrowers have lower price sensitivity (as $\overline{\alpha}^D_{\mathcal{P}} < 0$) and therefore are more likely to take credit. We interpret a negative correlation between unobservables determining collateral sensitivity and default $\rho_{\mathcal{C}F} < 0$ as evidence that collateral can mitigate adverse selection by inducing separation of borrowers of different risk, as risker borrowers have higher disutility from pledging collateral. Moreover, we would expect $\rho_{\mathcal{PC}} < 0$, which implies that borrowers with higher disutility from price (safe ones if $\rho_{\mathcal{P}F} > 0$) have lower disutility from pledging collateral (safe ones if $\rho_{\mathcal{C}F} < 0$). Finding that $\rho_{\mathcal{PC}} < 0$ is also evidence that collateral combined with interest rate can serve as signaling or screening device, because it implies that a price sensitive borrower is more likely to be collateral tolerant. Consequently, safer firms find it more favorable than risky ones to pledge collateral for lower interest rate, and banks can offer a lower interest rate for collateralized loans as the pool of borrowers that self selects into those will be more creditworthy. This would be evidence consistent with the ex ante private information hypothesis that justifies the presence of collateral.

Our model captures moral hazard through two distinct channels. The first is through $\alpha^F_{\mathcal{P}}$. Finding that $\alpha^F_{\mathcal{P}} > 0$ implies that, conditional on selection, a higher interest rate increases the likelihood that a borrower will default, which provides evidence of moral hazard. The coefficient $\alpha^F_{\mathcal{P}}$ can identify the moral hazard channel distinctly from the adverse selection channel, which is captured by the correlations between unobservables, leaving the remaining relationship between loan interest rates and default to capture the ex post moral hazard channel. The second is through $\alpha^F_{\mathcal{C}}$. Finding that $\alpha^F_{\mathcal{C}} < 0$ implies that, after controlling for selection, borrowers pledging collateral are less likely to default, as they have more at stake. This coefficient allows to evaluate whether collateral is effective in mitigating ex post incentive problems.

3.2 Supply

We let banks use the interest rate on secured $\mathcal{S}$ and unsecured $\mathcal{U}$ loans both as a competitive and as a screening device. In particular, we assume that banks compete Bertrand-Nash on interest rates for each individual borrower. We don't model banks' decision to offer either both secured and unsecured loans or one of the two types to each borrower, mostly to keep the model tractable. We do however observe in the data heterogeneity across borrowers in terms of types of loans offered, mostly varying across banks and firms' industries. As described in detail in Section 4, we rely on propensity score matching to determine whether each borrower is offered both types of loans or only one type by each bank.[14] To be more specific, we allow each bank j to set its interest rates on secured $\mathcal{S}$ and unsecured $\mathcal{U}$ loans to maximize its expected profit from a relationship with borrower i as follows:

$$\Pi_{ijm} = \sum_{k \in \{\mathcal{S},\mathcal{U}\}} \mathbb{1}_{ijkm} \Pi_{ijkm}, \tag{5}$$

where $\mathbb{1}_{ijkm}$ indicates the availability of type k loan. Banks can offer both loans, one of them or neither to any borrower. Expected profits from secured and unsecured loans are defined as:

[14] Alternatively, we could assume that all banks offer both secured and unsecured loans to all borrowers. This might however be an inaccurate representation of borrowers' choice sets, which could lead to biased estimates of price and collateral preferences.

$$\begin{aligned}\Pi_{ijkm} &= \left[(1+T_{ijm}P_{ijkm})-MC_{ijkm}\right]Q_{ijkm}\left(1-F_{ijkm}\right)+\left[R_{ijkm}-MC_{ijkm}\right]Q_{ijkm}F_{ijkm}\\ &= \left[(1+T_{ijm}P_{ijkm})\left(1-F_{ijkm}\right)-MC_{ijkm}+R_{ijkm}F_{ijkm}\right]Q_{ijkm},\end{aligned}\tag{6}$$

where T_{ijm} is the term of the loan (in years) determined by the firm demand, P_{ijkm} is the interest rate offered by bank j to borrower i for loan type k, and F_{ijkm} is the expected default probability of the borrower under each loan type. MC_{ijkm} is the marginal cost of the lending relationship with firm i, including cost of capital as well as administrative and screening costs, which can vary across bank, market and loan type. Q_{ijkm} is the expected demand defined as the probability of demand times the size of the loan:

$$Q_{ijkm}=\text{Pr}^{D}_{ijkm}LS_{ijkm},\tag{7}$$

where Pr^{D}_{ijkm} is the probability of demand and LS_{ijkm} is the loan size.[15] R_{ijkm} is the bank's loan recovery rate in default. We assume that:

$$\begin{aligned}R_{ij\mathcal{S}m} &= \min\left\{CV_{ijm},\left(1+T_{ijm}P_{ij\mathcal{S}m}\right)\right\}\\ R_{ij\mathcal{U}m} &= \min\left\{\omega_{ijm},R_{ij\mathcal{S}m}\right\}\end{aligned}\tag{8, 9}$$

where CV_{ijm} is the collateral value to loan amount ratio if the firm would post collateral, and ω_{ijm} is the expected recovery rate for defaulting borrowers.[16] The recovery rate for secured loans depends on the collateral value, but cannot exceed each borrower's total repayment obligation. As the creditor is secured against the collateral, the recovery rate of the secured loan must be at least as high as the unsecured loan. If a bank offers both a secured and an unsecured loan to a borrower, taking the first order conditions of the bank's profit with respect to each interest rate delivers the following equilibrium pricing equations:

$$\begin{aligned}1+T_{ijm}P_{ijkm} &= \frac{MC_{ijkm}}{1-F_{ijkm}-\frac{Q_{ijkm}}{Q_{ijkm,P_k}}F_{ijkm,P_k}}\\ &\quad-\frac{T_{ijm}\left(1-F_{ijkm}\right)\frac{Q_{ijkm}}{Q_{ijkm,P_k}}+R_{ijkm}\left(F_{ijkm}+\frac{Q_{ijkm}}{Q_{ijkm,P_k}}F_{ijkm,P_k}\right)}{1-F_{ijkm}-\frac{Q_{ijkm}}{Q_{ijkm,P_k}}F_{ijkm,P_k}}\\ &\quad+\frac{\left[(1+T_{ijm}P_{ij-km})\left(1-F_{ij-km}\right)-MC_{ij-km}\right]Q_{ij-km,P_k}}{1-F_{ijkm}-\frac{Q_{ijkm}}{Q_{ijkm,P_k}}F_{ijkm,P_k}}.\end{aligned}\tag{10}$$

There are two types of loan, secured and unsecured, i.e., $k\in\{\mathcal{S},\mathcal{U}\}$ and $-k$ is the other loan type. $Q_{ijkm,P_{\mathcal{S}}}$ and $Q_{ijkm,P_{\mathcal{U}}}$ are the derivatives of demand with respect to secured and unsecured interest rates, $F_{ijkm,P_{\mathcal{S}}}$, $F_{ijkm,P_{\mathcal{U}}}$ are the derivatives of default with respect to secured and unsecured interest rates, and $-\frac{Q_{ijkm}}{Q_{ijkm,P_k}}$

[15] These two variables will be defined more in detail respectively in Section 4.2 and Section 4.3.

[16] In our counterfactuals we will actually be using the average recovery rate calculated from the loss rate reported in Table 1, that is uniform across borrowers and banks. We plan to make this more heterogeneous.

is bank j's markup on a loan of type k to firm i. The first term on the right hand side of the equation shows how the *effective marginal costs* influence interest rates, whereas the second term describes the effect of the *effective markup*. We refer to Crawford, Pavanini and Schivardi (2018) for a detailed discussion on how these two terms, and in particular their denominator, capture the interaction of adverse selection and imperfect competition in their effect on loan pricing. We focus instead on two main novel aspects of our pricing first order condition.

The first novelty is that, in the second term on the right hand side of the pricing equation, the value of the collateral directly affects the recovery rate, and hence the interest rate offered. Intuitively, this implies that the higher is the collateral value (and the recovery rate) the lower will be the interest rate, due to the negative sign in front of the second term on the right hand side of the equation. This makes economic sense, as more collateral (or better recovery rate) implies less risk and more profit for the lender in case of default. This effect however depends on the sign and magnitude of the term in parenthesis that R_{ijkm} multiplies, which can be interpreted as follows. The more likely is the firm to default (larger F_{ijkm}) the larger is going to be the price reduction driven by the recovery rate, as the bank now gives more importance to the value of the collateral pledged. However, the stronger is the bank's markup $\frac{Q_{ijkm}}{Q_{ijkm,P_k}}$, which is negative, the smaller is going to be the price reduction driven by the recovery rate, as the bank exercises its market power.

The second new point is that the two interest rates on secured and unsecured loans in each bank-borrower combination are jointly determined and affect each other, as the two types of loans are in direct competition for the same borrowers. This competition effect is captured by the last term on the right hand side of equation (10). It shows that a higher profit for a secured (unsecured) loan is positively associated with the interest rate for the unsecured (secured) loan offered by the same bank to the same borrower. In other words, banks are multi-product firms and internalize their profits from the secured (unsecured) loans when setting the interest rate for the unsecured (secured) loan to borrower i.

Our counterfactual on the collateral channel, where we shock the value of the collateral and hence the value of the recovery rate R_{ijkm}, will therefore rely on the mechanisms highlighted by this first order condition to propagate to the supply response of banks, and consequently to their expected profits, and to borrowers' demand and default.

4 Econometric Model

4.1 Price Prediction

In order to construct the full choice set of each borrower we need to predict all loan contracts available to a borrower and their corresponding interest rates. We make a set of assumptions to determine borrowers' contract availability. First, we include a bank in a borrower's choice set if that bank granted at least one loan in the region-quarter combination where-when the borrower is taking her loan. Second, if a bank has never granted a loan with a similar amount and duration to similar new borrowers of the same type, we assume that the bank is not part of the borrower's choice set. Once we determine each borrower's available choice set, we predict the interest rates of contracts not observed in the data following a three steps procedure. First, we use an OLS regression model with a large set of fixed effects to predict the average interest rate across all loans that each borrower is offered by all banks it borrowed from in each market. Crucially, using multiple loans for each borrower, we are able to recover borrower-specific fixed effects that capture both hard and soft information common to all banks that is used for pricing. Second, as the first step doesn't give us a separate prediction for secured and unsecured loans' interest rates, we use propensity score matching to pair borrowers that are equally likely to take a secured loan from a given bank, and then assign the secured rate of a firm that took a collateralized loan in the data to its matched counterpart that took instead an uncollateralized loan, and vice-versa. Last, we combine these two methods to give the most credible prediction of loan interest rates for secured and unsecured loans for each borrower-bank combination. In what follows, we describe these steps in detail and assess the prediction accuracy of our approach. Note that we only need to predict interest rates to estimate our demand model, whereas we will use actual interest rates to estimate our default model.

4.1.1 Fixed Effects Model

In the first step we predict the average interest rate I_{ijm} across secured and unsecured loans borrowed by firm i from bank j in market m as follow:

$$I_{ijm} = \overline{\beta} + \beta_{\mathcal{A}} A_i + \beta_{\mathcal{M}} M_i + \gamma_{jm} + \lambda_i + \epsilon_{ijm}, \tag{11}$$

where A_i indicates borrower i's loan amount category, and M_i indicates i's maturity category. Both variables are categorized by quantiles.[17] γ_{jm} are bank-market fixed effects, λ_i are borrower fixed effects, and ϵ_{ijm} are prediction errors. By including multiple loans granted to the same borrower within the first six months from its first loan origination, we gain the possibility of identifying borrowers' fixed effects, which are likely to capture, at least to some extent, how the soft and hard information that banks acquire at origination (unobserved by the econometrician) maps into interest rates. Using the estimated coefficients $\widetilde{\beta}$, $\widetilde{\gamma}_{jm}$, $\widetilde{\lambda}_i$ we can predict I_{ijm} for all banks j that are available in market m.

Table 2 shows the results for predicting the average interest rate. In the first column, we report the estimation results for equation (11). The model's adjusted R-square is 0.914, indicating that the explanatory variables explain a large fraction of the variation of the average loan interest rate in the data. To evaluate the accuracy of this model, in the second column of Table 2 we report estimation results of a default model where the residuals from equation (11) along with all other explanatory variables, except for the borrower fixed effects, are included as explanatory variables, and the dependent variable is a dummy equal to one if a borrower has any non-performing loans within our sample period.[18] Crucially, we find that residuals are not statistically nor economically significant, which suggests that our prediction error is not related to borrowers' default

[17] The four loan amount categories are 600$ to 15,000$, 15,001$ to 30,000$, 30,001$ to 90,000$, and 90,009$ to 12,000,000$. The four maturity categories are 1 to 2.9 months, 3 to 5.9 months, 6 to 18 months, 18.1 to 180 months.

[18] This implies that there is no variation in the default dependent variable across loans within a borrower, therefore we cannot include borrower fixed effects.

and hence represents noise in banks' pricing strategy. We interpret this as a sign of the accuracy of our price prediction method.

Table 2: Price Prediction for Average Interest Rate

	Observed Price	**Default**
Price Residual		0.000
		(0.016)
Amount: 15,000$ to 30,000$	−0.101	−0.040
	(0.082)	(0.025)
Amount: 30,000$ to 90,000$	-0.170^{*}	0.004
	(0.088)	(0.024)
Amount: 90,000$ to 12,000,000$	−0.010	−0.028
	(0.105)	(0.025)
Maturity: 3 to 6 months	-0.293^{***}	−0.035
	(0.080)	(0.024)
Maturity: 6 to 18 months	-0.281^{***}	−0.020
	(0.092)	(0.025)
Maturity: 18 to 180 months	0.021	−0.030
	(0.107)	(0.026)
Bank-Market FE	Yes	Yes
Borrower FE	Yes	No
Constant	15.141^{***}	0.030
	(0.527)	(0.224)
Observations	2,871	2,871
R^2	0.967	0.406
Adjusted R^2	0.914	0.264

Note: This table shows the price prediction for average cost. The first column shows the OLS regression result for equation (11). The dependent variable is observed interest rate. Loan amount and maturity categorized by their quantiles. The first category of loan amount (600$ to 15,000$) and maturity (0 to 3 months) are omitted. The second column is to show the price prediction does not miss determinants for default. The price residual means the residuals from equation (11). The dependent variable is the indictor for Non-performing. *p<0.1; **p<0.05; ***p<0.01.

This approach doesn't yet take into account the different interest rates that a bank offers to the same borrower for a secured or an unsecured loan, mostly for reasons of statistical power, as we don't have enough observations to identify firm-secured loan and firm-unsecured loan fixed effects. The predicted average interest rate can thus be thought as the weighted average of interest rate between secured and unsecured loans that bank j has granted to borrower i, where the weight is given by the likelihood that i will take a secured or an unsecured loan. Hence, we rely on propensity score matching to separately predict interest rates for collateralized and uncollateralized loans for each borrower-bank combinations, as described in the next section.

4.1.2 Propensity Score Matching

In the second step we use propensity score matching (PSM) to determine for each firm-bank relationship in each market the probability that the firm will select a secured loan. This probability will be then used to derive from the predicted average interest rate $\widehat{I}_{ijm}$ the predicted loan interest rates for secured and unsecured loans $\widehat{P}_{ij\mathcal{S}m}, \widehat{P}_{ij\mathcal{U}m}$. The matching process works as follows. First, following the criteria suggested by Caliendo and Kopeinig (2008), we select as variables for the PSM the bank identity, the loan amount category, the loan maturity category, the first loan (i.e. whether the loan is the first loan of a new borrower), and the borrower's legal structure (i.e. whether the borrower is a corporation). Second, based on these variables, we use a logistic model to determine the propensity score PSC_{ijm} of borrower i in market m being a "secured" borrower when taking a loan from bank j. Third, we match each firm i that took a secured (unsecured) loan from bank j with another firm with the same propensity score PSC_{ijm} that has instead taken an unsecured (secured) loan from bank j, and assign to each other the secured (unsecured) interest rate $\tau_{ij\mathcal{S}m}$ ($\tau_{ij\mathcal{U}m}$) for the loan we don't observe in the data. When there are more than one match for the same combination of PSM_{ijm} we use random assignment. As a result, for each firm we obtain the interest rate for secured and unsecured loans offered by all banks that are actively lending in the market. Appendix A.1 provides detailed information on the optimal matching algorithm and the selection of the variables.

We restrict the potential matches to be loan contracts provided by the same bank with the same matching variables, which implies that for some borrower type-bank combinations we may not find any secured or unsecured match, and hence assume that either the secured or the unsecured loan is not offered to that borrower. Therefore, the predicted loan contracts are those provided by banks that are actively lending in a region-quarter combination, and those that are offered to borrowers with similar characteristics in the sample.

When both secured and unsecured loans are available and the matching is done, we define the interest rate difference $\mathcal{D}_{ijm}$ as the difference between the matched unsecured interest rate $\tau_{ij\mathcal{U}m}$ and the matched secured interest rate $\tau_{ij\mathcal{S}m}$:

$$\mathcal{D}_{ijm} = \tau_{ij\mathcal{U}m} - \tau_{ij\mathcal{S}m}. \tag{12}$$

In the next step, we use both this interest rate difference $\mathcal{D}_{ijm}$ and the propensity score PSC_{ijm} to derive the predicted interest rates $\widehat{P}_{ij\mathcal{S}m}, \widehat{P}_{ij\mathcal{U}m}$. The reason why we don't use the matched $\tau_{ij\mathcal{U}m}, \tau_{ij\mathcal{S}m}$ as predicted interest rates is that $\widehat{I}_{ijm}$ captures much more heterogeneity across borrowers thank to the firm-specific fixed effects, hence a combination of the two steps is what provides the most accurate prediction, as explained in the next section.

4.1.3 Price of Secured and Unsecured Loans

In the last step we predict the interest rate of secured and unsecured loans by adjusting the predicted average interest rate $\widehat{I}_{ijm}$ depending on the propensity score. Intuitively, if most of the loans used to predict $\widehat{I}_{ijm}$ are secured, then $\widehat{I}_{ijm}$ will be a good predictor for $\widehat{P}_{ij\mathcal{S}m}$, but a bad predictor for $\widehat{P}_{ij\mathcal{U}m}$. The opposite occurs if most of the loans used to predict $\widehat{I}_{ijm}$ are unsecured. The propensity score is what determines the probability that the loans used to predict $\widehat{I}_{ijm}$ are secured. Therefore, for a given average interest rate $\widehat{I}_{ijm}$ and price difference $\mathcal{D}_{ijm}$, the interest rates for secured and unsecured loans are defined as follows:

$$\begin{aligned}
\widehat{P}_{ij\mathcal{S}m} &= \widehat{I}_{ijm} - (1 - PSC_{ijm})\mathcal{D}_{ijm}, \\
\widehat{P}_{ij\mathcal{U}m} &= \widehat{I}_{ijm} + PSC_{ijm}\mathcal{D}_{ijm}.
\end{aligned}$$

Taking a secured loan as an example, this means that if a borrower is very likely to choose a secured loan ($PSC_{ijm} \approx 1$), then also most of the loans used to predict $\widehat{I}_{ijm}$ should be secured ones, and therefore it is reasonable to have that $\widehat{P}_{ij\mathcal{S}m} \approx \widehat{I}_{ijm}$. If on the other hand a borrower is very unlikely to choose a secured loan ($PSC_{ijm} \approx 0$), then most of the loans used to predict $\widehat{I}_{ijm}$ should be unsecured ones, which implies that $\widehat{I}_{ijm} \approx \tau_{ij\mathcal{U}m}$, and therefore it is reasonable to have that $\widehat{P}_{ij\mathcal{S}m} \approx \widehat{I}_{ijm} - \tau_{ij\mathcal{U}m} + \tau_{ij\mathcal{S}m} \approx \tau_{ij\mathcal{S}m}$. A similar argument applies for the case of the unsecured loan interest rate.

If bank j only provides one contract to borrower i, then the average interest rate is just the price of the available contract, and the other contract is not available. Hence:

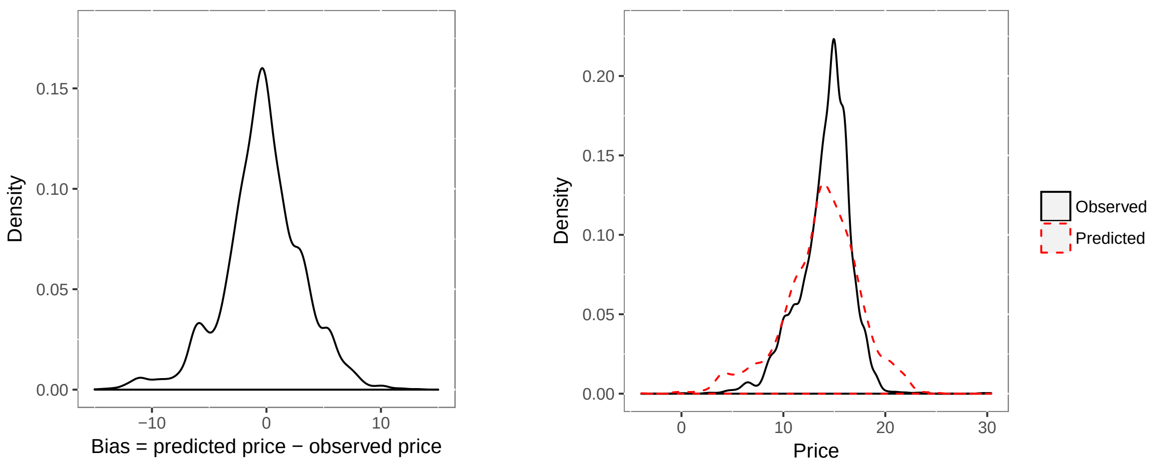
$$\begin{aligned}
\widehat{P}_{ij\mathcal{S}m} &= \widehat{I}_{ijm} \quad \text{if only secured loan is available;} \\
\widehat{P}_{ij\mathcal{U}m} &= \widehat{I}_{ijm} \quad \text{if only unsecured loan is available.}
\end{aligned}$$

If bank j provides neither contract to borrower i, then no contract is available to that firm.

4.1.4 Price Prediction Results

Based on our choice set assumptions and matching procedure, we predict the set of available contracts for each borrower at the time of her first loan's origination. From the benchmark case in which all banks were to offer both types of loans to each borrower, our assumptions and matching end up keeping 42.6% of those contracts as actually available to the borrowers. Among the unavailable contracts, in 83.9% of the cases they are not available as the bank is not actively lending in the borrower's market, and in 14.1% of the cases as the bank does not offer the amount and maturity combination required by the borrower. The median *secured borrower* (i.e. borrower that chose a secured loan in the data) has 5 secured and 6 unsecured loans available, while the median *unsecured borrower* (i.e. borrower that chose an unsecured loan in the data) has 4 secured and 7 unsecured loans available. Among the available contracts, in 11% of the cases a bank only provides a secured loan to a borrower, in 37.8% of the cases only an unsecured one, and in 51.2% of the cases it offers both types of loans. Our propensity score matching allows for different contract availability between secured and unsecured borrowers, which implies that banks can screen borrowers both with contract terms and contract availability. More detailed information on the contract availability is presented in Appendix A.2.

In order to assess the accuracy of our price prediction, we compare actual and predicted interest rates for the contracts that we observe in the data. Figure 4 (a) shows the distribution of the prediction bias, measured as the difference between predicted and observed interest rates. The prediction biases are concentrated around zero with mild deviations. Similarly, Figure 4 (b) shows the distribution of observed and predicted interest rates. Although the predicted prices have a higher standard deviation, the two distributions have a very large overlap.[19]



(a) Prediction Bias (b) Observed Prices and Predicted Prices

Figure 4: Price Prediction Accuracy

Note: Subfigure (a) depicts the distribution of bias of price prediction. The bias is defined as the predicted price minus the observed price. Subfigure (b) shows the distributions of observed prices (black solid line) and predicted prices (red dashed line). The total number of observation is 2,871.

4.2 Demand and Default

We estimate the model by simulated maximum likelihood, using a mixed logit for the demand model and a probit for the default model. Starting from the former, we define the probability that borrower $i = 1, .., I$ in market $m = 1, .., M$ takes a type $k = \mathcal{S}, \mathcal{U}$ loan from bank $j = 1, .., J_m$ as follows:

[19] In Appendix A.3 we present another price prediction method that only uses fixed effects, which has similar prediction accuracy but is less flexible in terms of contract availability.

$$
\begin{aligned}
\Pr^D_{ijkm} =& \int\int \frac{\exp\left(\alpha^D_{\mathcal{P}i}P_{ijkm} + \alpha^D_{\mathcal{C}i}C_{ijkm} + X'_{jm}\alpha^D_{\mathcal{X}}\right)}{\sum_{\jmath=i}^{J_m}\sum_{\ell=\mathcal{S}}^{\mathcal{U}}\exp\left(\alpha^D_{\mathcal{P}i}P_{i\jmath\ell m} + \alpha^D_{\mathcal{C}i}C_{i\jmath\ell m} + X'_{\jmath m}\alpha^D_{\mathcal{X}}\right)} f(\varepsilon^D_{\mathcal{P}i}, \varepsilon^D_{\mathcal{C}i}) d\varepsilon^D_{\mathcal{P}i} d\varepsilon^D_{\mathcal{C}i} \\
\approx& \frac{1}{S}\sum_{s=1}^{S} \underbrace{\frac{\exp\left(\alpha^D_{\mathcal{P}is}P_{ijm} + \alpha^D_{\mathcal{C}is}C_{ijm} + X'_{jm}\alpha^D_{\mathcal{X}}\right)}{\sum_{\jmath=i}^{J_m}\sum_{\ell=\mathcal{S}}^{\mathcal{U}}\exp\left(\alpha^D_{\mathcal{P}is}P_{i\jmath\ell m} + \alpha^D_{\mathcal{C}is}C_{i\jmath\ell m} + X'_{\jmath m}\alpha^D_{\mathcal{X}}\right)}}_{\Pr^D_{isjkm}},
\end{aligned}
\tag{13}
$$

where we approximate the integral in the first row using Monte Carlo simulations with $S = 100$ Halton draws, and index each draw by s. The simulation draws enter the random coefficients on interest rate and collateral as in equation (2):

$$
\begin{aligned}
\alpha^D_{\mathcal{P}is} &= \overline{\alpha}^D_{\mathcal{P}} + Y'_i\delta_{\mathcal{P}} + \varepsilon^D_{\mathcal{P}is}, \\
\alpha^D_{\mathcal{C}is} &= \overline{\alpha}^D_{\mathcal{C}} + Y'_i\delta_{\mathcal{C}} + \varepsilon^D_{\mathcal{C}is},
\end{aligned}
$$

where, following the conditional distribution of the multivariate normal:

$$
\begin{aligned}
\varepsilon^D_{\mathcal{P}is} &= \sigma_{\mathcal{P}}\zeta^D_{\mathcal{P}is}, \\
\varepsilon^D_{\mathcal{C}is} &= \frac{\sigma_{\mathcal{C}}}{\sigma_{\mathcal{P}}}\rho_{\mathcal{PC}}\varepsilon^D_{\mathcal{P}is} + \sqrt{(1-\rho^2_{\mathcal{PC}})\sigma^2_{\mathcal{C}}}\zeta^D_{\mathcal{C}is} = \sigma_{\mathcal{C}}\rho_{\mathcal{PC}}\zeta^D_{\mathcal{P}is} + \sqrt{(1-\rho^2_{\mathcal{PC}})}\sigma_{\mathcal{C}}\zeta^D_{\mathcal{C}is},
\end{aligned}
\tag{14}
$$

with $\zeta^D_{\mathcal{P}is}, \zeta^D_{\mathcal{C}is} \sim N(0,1)$. Conditional on taking a specific loan from the most preferred bank, which is determined by $\varepsilon^D_{\mathcal{P}i}$ and $\varepsilon^D_{\mathcal{C}i}$, we model each borrower's default probability, that is the probability that the utility from defaulting is positive, as:

$$
\begin{aligned}
\Pr^F_{ijkm} =& \int\int \Phi_{\varepsilon^F_i|\varepsilon^D_{\mathcal{P}i},\varepsilon^D_{\mathcal{C}i}}\left(\frac{\overline{\alpha}^F + \alpha^F_{\mathcal{P}}P_{ijkm} + \alpha^F_{\mathcal{C}}C_{ijkm} + X'_{jm}\alpha^F_{\mathcal{X}} + Y'_i\alpha^F_{\mathcal{Y}} + \widetilde{\mu}_{Fi}}{\widetilde{\sigma}_F}\right) f(\varepsilon^D_{\mathcal{P}i}, \varepsilon^D_{\mathcal{C}i}) d\varepsilon^D_{\mathcal{P}i} d\varepsilon^D_{\mathcal{C}i} \\
\approx& \frac{1}{S}\sum_{s=1}^{S} \underbrace{\Phi_{\varepsilon^F_i|\varepsilon^D_{\mathcal{P}i},\varepsilon^D_{\mathcal{C}i}}\left(\frac{\overline{\alpha}^F + \alpha^F_{\mathcal{P}}P_{ijkm} + \alpha^F_{\mathcal{C}}C_{ijkm} + X'_{jm}\alpha^F_{\mathcal{X}} + Y'_i\alpha^F_{\mathcal{Y}} + \widetilde{\mu}_{Fis}}{\widetilde{\sigma}_F}\right)}_{\Pr^F_{isjkm}},
\end{aligned}
\tag{15}
$$

where $\varepsilon^F_i|\varepsilon^D_{\mathcal{P}i}, \varepsilon^D_{\mathcal{C}i} \sim N(\widetilde{\mu}_{Fi}, \widetilde{\sigma}_F)$. Following the conditional distribution of the multivariate normal, we have that:

$$
\begin{aligned}
\widetilde{\mu}_{Fis} &= (A'_F B_F^{-1} C_F)', \\
\widetilde{\sigma}_F &= 1 - A'_F B_F^{-1} A_F,
\end{aligned}
\tag{16}
$$

with:

$$A_F = \begin{pmatrix} \rho_{\mathcal{P}F}\sigma_{\mathcal{P}} \\ \rho_{\mathcal{C}F}\sigma_{\mathcal{C}} \end{pmatrix} \quad , \quad B_F = \begin{pmatrix} \sigma_{\mathcal{P}}^2 & \rho_{\mathcal{PC}}\sigma_{\mathcal{P}}\sigma_{\mathcal{C}} \\ \rho_{\mathcal{PC}}\sigma_{\mathcal{P}}\sigma_{\mathcal{C}} & \sigma_{\mathcal{C}}^2 \end{pmatrix} \quad , \quad C_F = \begin{pmatrix} \varepsilon_{\mathcal{P}is}^D \\ \varepsilon_{\mathcal{C}is}^D \end{pmatrix}. \tag{17}$$

Solving the matrix multiplication we get:

$$\begin{aligned} \widetilde{\mu}_{Fis} &= \frac{\rho_{\mathcal{P}F} - \rho_{\mathcal{C}F}\rho_{\mathcal{PC}}}{\sigma_{\mathcal{P}}(1-\rho_{\mathcal{PC}}^2)}\varepsilon_{\mathcal{P}is}^D + \frac{\rho_{\mathcal{C}F} - \rho_{\mathcal{P}F}\rho_{\mathcal{PC}}}{\sigma_{\mathcal{C}}(1-\rho_{\mathcal{PC}}^2)}\varepsilon_{\mathcal{C}is}^D \\ &= \frac{\rho_{\mathcal{P}F} - \rho_{\mathcal{C}F}\rho_{\mathcal{PC}}}{1-\rho_{\mathcal{PC}}^2}\zeta_{\mathcal{P}is}^D + \frac{\rho_{\mathcal{C}F} - \rho_{\mathcal{P}F}\rho_{\mathcal{PC}}}{1-\rho_{\mathcal{PC}}^2}\left(\rho_{\mathcal{PC}}\zeta_{\mathcal{P}is}^D + \sqrt{(1-\rho_{\mathcal{PC}}^2)}\zeta_{\mathcal{C}is}^D\right). \end{aligned} \tag{18}$$

$$\widetilde{\sigma}_F = 1 - \frac{\rho_{\mathcal{P}F}^2 + \rho_{\mathcal{C}F}^2 - 2\rho_{\mathcal{P}F}\rho_{\mathcal{C}F}\rho_{\mathcal{PC}}}{1-\rho_{\mathcal{PC}}^2} \tag{19}$$

We use these probabilities to estimate all the parameters $\theta = \{\alpha^D, \alpha^F, \Sigma\}$ jointly by maximum simulated likelihood, where $\alpha^D = \{\overline{\alpha}_{\mathcal{P}}^D, \delta_{\mathcal{P}}, \overline{\alpha}_{\mathcal{C}}^D, \delta_{\mathcal{C}}, \alpha_{\mathcal{X}}^D\}$, $\alpha^F = \{\overline{\alpha}^F, \alpha_{\mathcal{P}}^F, \alpha_{\mathcal{C}}^F, \alpha_{\mathcal{X}}^F, \alpha_{\mathcal{Y}}^F\}$, and $\Sigma = \{\sigma_{\mathcal{P}}, \sigma_{\mathcal{C}}, \rho_{\mathcal{PC}}, \rho_{\mathcal{P}F}, \rho_{\mathcal{C}F}\}$. We use the following log likelihood function:

$$\mathcal{L}(\theta) = \sum_i \log\left\{\frac{1}{S}\sum_{s=1}^{S}\left[\left(\prod_j\prod_k \left(\Pr_{isjkm}^D\right)^{d_{ijkm}}\right)\left(\left(\Pr_{isjkm}^F\right)^{f_{ijkm}}\left(1-\Pr_{isjkm}^F\right)^{1-f_{ijkm}}\right)\right]\right\}, \tag{20}$$

where d_{ijkm} takes the value of one if the borrower chooses a bank-loan combination j with loan type k, and zero otherwise, and f_{ijkm} takes the value of one if the borrower defaults, and zero otherwise.

4.3 Loan Amount

In the demand model we assume that loan amount and maturity are exogenously determined, depending on firms' financing needs. If the exogenous amount assumption can be justified for the demand estimation, it can become problematic when simulating counterfactual scenarios, especially because we don't allow borrowers to choose the outside option of not taking a loan, which would make aggregate credit demand invariant across scenarios. To overcome this limitation, we model separately the loan size LS_{ijkm} (i.e. total amount granted) that firm i borrows from bank j in market m when choosing contract k as follow:

$$\log(LS_{ijkm}) = \overline{\zeta} + \zeta_{\mathcal{P}}P_{ijkm} + \zeta_{\mathcal{C}}C_{ijkm} + X_{jm}'\zeta_{\mathcal{X}} + Y_i'\zeta_{\mathcal{Y}} + v_{ijkm}, \tag{21}$$

where P_{ijkm} is the interest rate, C_{ijkm} is the collateral dummy, and X_{jm} and Y_i include the same variables as in demand and default utility except for loan amount categories. v_{ijkm} is an IID normally distributed error term. This model will allow us to have variation in credit demand in the counterfactual scenarios, as it will enter banks' profit functions.

Table 3: First Stage Results

	Predicted Price	**Observed Price**
Saving deposits	0.065***	0.110***
	(0.017)	(0.026)
Saving to demand deposit ratio	0.046***	0.182***
	(0.015)	(0.048)
Deposit interest expense	0.659***	0.438***
	(0.065)	(0.095)
Saving deposit interest rate		0.165***
		(0.032)
Interest rate in other markets		0.072**
		(0.034)
Loan Controls	Yes	Yes
Amount FE	Yes	Yes
Maturity FE	Yes	Yes
Bank FE	Yes	Yes
Region FE	Yes	Yes
Industry FE	Yes	Yes
Constant	11.542***	11.545***
	(0.290)	(0.687)
Observations	23,900	2,592
R^2	0.151	0.533
Adjusted R^2	0.149	0.523

Note: This table shows the first stage results for prices. In the first column, the dependent variable is predicted price. In the second column, the dependent variable is the price we observe in the sample. The instrumental variables are the amount of saving deposits, the saving deposits amount to demand deposits amount ratio, interest rate of deposits, interest rate of saving deposits. Loan Controls includes Collateral, Installment, Corporation, Bad Credit Rating. The number of observation is less than the total number of predicted prices and observed prices because there are some missing values in the instrumental variables. *p<0.1; **p<0.05; ***p<0.01.

4.4 Identification

Since we do not know the precise actuarial model that banks use to determine the interest rate for each borrower, a natural concern is that the loan interest rate, both predicted (used in the demand model) and observed (used in the default model), may be endogenously related to unobservables that influence borrowers' demand and default. If this is the case, our estimates of the price sensitivity in both the demand and the default models are likely to be biased. To address this potential endogeneity concern, we use the control function approach suggested by Train (2009), motivated by the fact that both demand and default are nonlinear models.[20] This method consists of two steps. In the first stage, we regress the predicted and actual interest rates on the same set of observables that we use in the demand and default models, plus a set of instrumental variables. In the second stage, we include the residuals from each pricing regression as control variables in the demand and default models to control for any unobserved factors correlated with prices, thus allowing the identifying variation left over in prices to be orthogonal to demand and default unobservables.

In line with Crawford, Pavanini and Schivardi (2018), we use two partially overlapping sets of instruments for demand and default, as they need to satisfy different exclusion restrictions. For both the demand and the default models we include proxies for banks' funding sources and costs from household deposits data, such as the total amount of saving deposits, the ratio of savings to demand deposits, and deposit interest rates. Columns (1) and (2) in Table 3 present the first-stage results for predicted and observed loan interest rates, showing that these instruments are relevant for both measures of loan interest rates, with positive coefficients as expected. We believe this set of instruments fulfills the exclusion restriction, as household deposit markets represent a different segment of banking activity compared to corporate loans, therefore any change in its conditions is likely to be correlated with loan rates, but uncorrelated with unobserved determinants of firms' choice of bank and of their likelihood of default. Additionally, for the default model's first-stage, we include as instruments the saving deposits interest rates as well as loan interest rates charged by the same bank in the same quarter in other regions. This latter instrument, in the spirit of Hausman and Taylor (1981), can violate the exclusion restriction in the demand model, but is unlikely to violate it for the default model, as the loan interest rates in other markets are unlikely to affect borrowers' ex post behavior.

5 Results

5.1 Estimates

We use data on each borrower's choice of her first loan to estimate the demand, default, and loan amount models. Table 4 presents the estimation results of our structural model. The first two columns refer to the demand equation, with the first and second columns reporting respectively the estimates of the mean component of the random coefficient on price and collateral. The third column refers to the default equation.

[20]We implement this control function approach also in the loan amount model, using the same instruments as in the demand model.

The bottom panel shows the covariance matrix of the unobservables. Both the demand and the default equations are estimated using maximum simulated likelihood. The fourth column reports OLS regression results for the loan amount model, which are mainly used in the counterfactual analyses for the supply-side model.

In the demand equation, we control for bank-fixed effects, and allow the random coefficients (RC) on prices and collateral to depend on unobserved heterogeneity. The mean utilities from interest rate and collateral in the demand model are reported in the "Constant" row of Table 4, in the first two columns. We find that on average borrowers get disutility from higher interest rates and from pledging collateral. The mean own price and collateral elasticities suggest that a 10% increase in interest rate reduces the own probability of demand by 1.1%, and requiring collateral reduces the own probability of demand by 15.8%. The last column shows that the interest rate have a negative impact on loan amount: a 1 percentage point increase in interest rate decreases the loan amount by 21.9%. Therefore, in our counterfactuals we allow demand to adjust to price changes through both an extensive margin (demand probability) and an intensive margin (loan amount).

Since we have no information on borrowers that do not demand a bank loan, we cannot control for loan and borrower characteristics, as these are constant across borrowers' options in their choice set and their effect on demand would therefore not be identified. We have experimented interacting price and collateral with the borrowers' variables we have (legal status and rating), but found no statistically significant effect.

In the default equation, we include bank, loan amount, maturity, region, and borrower-industry fixed effects. We find that the loan interest rate has a positive and significant effect on default, while collateral has a negative and significant effect. The results suggest that on average a 10% increase in the interest rate increases the probability of default by 18.1%, while posting collateral decreases the probability of default by 88.8%. Consistent with Stiglitz and Weiss (1981), the price effect implies that, conditional on selection, a higher interest rates makes borrowers less likely to repay their loan. The collateral result instead is consistent with collateral mitigating the ex post incentive problem. When borrowers pledge collateral they are more likely to repay, given that they have more at stake in the loan. Consistent with the ex post theories of collateral, this result indicates that collateral is an effective tool in mitigating moral hazard and other ex post problems that facilitate or encourage defaults.

The bottom panel of Table 4 shows the covariance matrix for unobservable shocks. The positive and significant correlation between price sensitivity and borrowers' unobserved riskiness $\rho_{\mathcal{P}F}$ suggests that firms with high unobservable default risk are less price sensitive and more likely to take credit, which we interpret as evidence of adverse selection. On the other hand, the negative and significant correlation between collateral sensitivity and borrowers' unobserved riskiness $\rho_{\mathcal{C}F}$ suggests that riskier firms are less likely to demand credit if collateral is required, which we interpret as evidence that collateral can mitigate adverse selection and induce separation of borrowers of different risk. Moreover, the negative correlation between price and collateral sensitivities $\rho_{\mathcal{P}\mathcal{C}}$ implies that firms with higher disutility from interest rate have instead lower disutility from collateral, as illustrated by the red linear model smoothed line in Figure 5 (a). This implies that borrowers with higher unobservable risk are more price tolerant as well as collateral sensitive, suggesting that safe borrowers prefer a secured loan with low interest rate, while risky borrowers prefer an

Table 4: Structural Estimation Results

	MSL			**OLS**
	Demand		**Default**	**Amount**
	Price RC	Collateral RC		
Constant	-0.159^{***}	-0.175^{***}	-2.498^{***}	12.272^{***}
	(0.013)	(0.012)	(0.027)	(0.502)
Price			0.526^{***}	-0.219^{***}
			(0.013)	(0.025)
Collateral			-0.196^{***}	0.327^{***}
			(0.021)	(0.123)
Price residual	0.413^{***}		1.139^{***}	0.020
	(0.010)		(0.014)	(0.034)
Installment			0.103^{***}	-0.224
			(0.022)	(0.190)
Corporation			0.235^{***}	0.377^{***}
			(0.022)	(0.120)
Bad Credit Rating			1.136^{***}	1.160^{***}
			(0.081)	(0.380)
Bank FE	Yes		Yes	Yes
Amount FE	No		Yes	No
Maturity FE	No		Yes	Yes
Industry FE	No		Yes	Yes
Region FE	No		Yes	Yes
Observations	561		561	561
	$\sigma_{\mathcal{P}} = 0.148^{***}$			
	(0.012)			
Covariance matrix	$\rho_{\mathcal{PC}} = -0.240^{***}$	$\sigma_{\mathcal{C}} = 0.209^{***}$		
	(0.014)	(0.014)		
	$\rho_{\mathcal{P}F} = 0.452^{***}$	$\rho_{\mathcal{C}F} = -0.810^{***}$	$\sigma_F = 1$	
	(0.018)	(0.012)		

Note: This table presents the structural estimation results. The first two columns are for demand, and the third column is for default. The last column is for loan amount, where the dependent variable is the logarithm of loan amount. There are two random coefficients (RC) in demand: price (1st column) and collateral (2nd column), which contain constant and a normally distributed random terms. In demand part, the variable Price stands for predicted price, while in default part, Price stands for observed price. Price and Price residual are normalized at 95 percentile of predicted price (i.e., 18 percentage points per year) in the demand and default model. *p<0.1; **p<0.05; ***p<0.01.

unsecured loan with high interest rate.

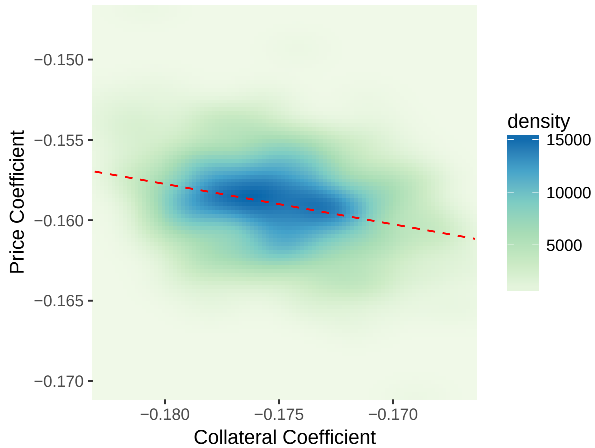
Figure 5 gives a graphical interpretation of these results. Subfigure (a) reports the joint distribution of the price and collateral coefficients, where the center corresponds to two mean utilities, and the two random coefficients are negatively correlated as indicated by the red dashed line. Subfigure (b) shows the relationship between borrowers' preferences for price and collateral and their unobserved riskiness levels. As conditional on taking a specific loan the unobserved risk ε_i^F is normally distributed with idiosyncratic mean $\tilde{\mu}_{Fi}$, we use as measure of unobserved risk the estimate of this mean as of equation (18), which is distributed with mean 0.00 and standard deviation 0.01. A standard deviation increases in our measure of unobserved risk $\tilde{\mu}_{Fi}$ increases the probability of default by 3.5% on average.

Risky borrowers are in red while safe borrowers are in green. The riskier a borrower is, the further away it locates from the center towards the top-left corner. That is, riskier borrowers have lower price disutility and higher collateral disutility. The opposite holds for safe borrowers, they are closer to the bottom-right corner, as they have lower collateral disutility and higher price disutility. Hence, this figure demonstrates that it is possible for banks to screen borrowers using collateral. Notice that the collateral coefficient to price coefficient ratio corresponds to the borrower's marginal rate of substitution of collateral for price $MRS_{\mathcal{C},\mathcal{P}}$. As illustrated in the figure, riskier borrowers have higher $MRS_{\mathcal{C},\mathcal{P}}$, as assumed by the theoretical literature that motivates collateral as a screening device of unobserved borrower risk. Therefore, by setting the interest rates on secured and unsecured contracts, banks can make the interest rate benefit of choosing a secured loan compared to choosing an unsecured loan high enough for safe borrowers but too low for risky borrowers, inducing a separating equilibrium. Hence, safe borrowers will be more likely to choose a secured loan with low interest rate, while risky borrowers will be more likely to choose an unsecured loan with a high interest rate, just as what Figure 5 (b) shows.

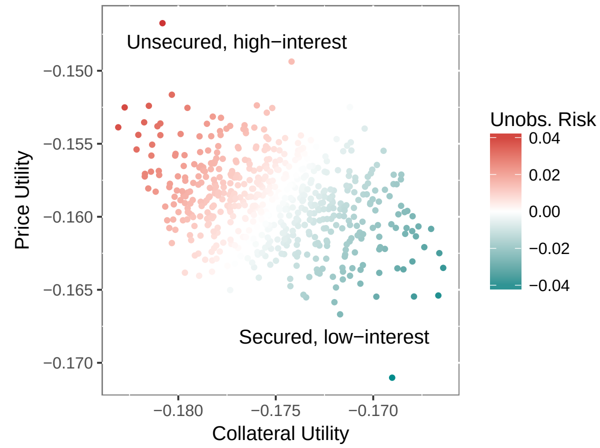
These results confirm the existence of both ex ante and ex post asymmetric information frictions and show that collateral can reduce both kinds of frictions. Furthermore, it provides empirical evidence that risky borrowers have a higher marginal rate of substitution of collateral for price, a fundamental assumption in the ex ante theories of collateral (Bester 1985, Chan and Thakor 1987), which to the best of our knowledge has never been tested before. Exploiting the variation in borrowers' preferences, lenders can use interest rate and collateral to affect borrowers' choices, implement screening, reduce credit rationing, and increase social welfare.

5.2 Model Fit

We use the estimates of the demand and default models to calculate predicted credit demand $\widehat{Q}_{ij\mathcal{S}m}$, $\widehat{Q}_{ij\mathcal{U}m}$, default probabilities $\widehat{F}_{ij\mathcal{S}m}$, $\widehat{F}_{ij\mathcal{U}m}$, and their derivatives with respect to interest rates. Credit demand is defined as demand probability times the loan amount. Based on equation (10), we solve the first order conditions to back out the marginal costs for secured and unsecured loans:



(a)



(b)

Figure 5: Random Coefficients of Price and Collateral

Note: These figures plot model estimated price and collateral coefficients for all firms. Subfigure (a) plots the joint density of price and collateral coefficients. The red dashed line is the linear model fitted line which captures the correlation between price and collateral coefficients. Subfigure (b) plots each observation explicitly. *Unobs. Risk* is the estimated unobserved risk. High unobserved risk firms are in red and low unobserved risk firms are in green. The dash line goes through the origin of the coordinate system with a slope 0.9

$$\widehat{MC}_{ij\mathcal{S}m} = \frac{1}{\widehat{Q}_{ij\mathcal{S}m,P_\mathcal{U}}\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{S}} - \widehat{Q}_{ij\mathcal{S}m,P_\mathcal{S}}\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{U}}}(B\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{S}} - A\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{U}}), \tag{22}$$

$$\widehat{MC}_{ij\mathcal{U}m} = \frac{1}{\widehat{Q}_{ij\mathcal{S}m,P_\mathcal{U}}\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{S}} - \widehat{Q}_{ij\mathcal{S}m,P_\mathcal{S}}\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{U}}}(A\widehat{Q}_{ij\mathcal{S}m,P_\mathcal{U}} - B\widehat{Q}_{ij\mathcal{S}m,P_\mathcal{S}}), \tag{23}$$

where:

$$\begin{aligned} A \;=\; & \left[(1+T_{ijm}\widetilde{P}_{ij\mathcal{S}m})(1-\widehat{F}_{ij\mathcal{S}m}) + \widehat{R}_{ij\mathcal{S}m}\widehat{F}_{ij\mathcal{S}m}\right]\widehat{Q}_{ij\mathcal{S}m,P_\mathcal{S}} \\ & + \left[T_{ijm}(1-\widehat{F}_{ij\mathcal{S}m}) - (1+T_{ijm}\widetilde{P}_{ij\mathcal{S}m})\widehat{F}_{ij\mathcal{S}m,P_\mathcal{S}} + \widehat{R}_{ij\mathcal{S}m}\widehat{F}_{ij\mathcal{S}m,P_\mathcal{S}}\right] \\ & + \left[(1+T_{ijm}\widetilde{P}_{ij\mathcal{U}m})(1-\widehat{F}_{ij\mathcal{U}m}) + \widehat{R}_{ij\mathcal{U}m}\widehat{F}_{ij\mathcal{U}m}\right]\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{S}}, \end{aligned} \tag{24}$$

$$\begin{aligned} B \;=\; & \left[(1+T_{ijm}\widetilde{P}_{ij\mathcal{U}m})(1-\widehat{F}_{ij\mathcal{U}m}) + \widehat{R}_{ij\mathcal{U}m}\widehat{F}_{ij\mathcal{U}m}\right]\widehat{Q}_{ij\mathcal{U}m,P_\mathcal{U}} \\ & + \left[T_{ijm}(1-\widehat{F}_{ij\mathcal{U}m}) - (1+T_{ijm}\widetilde{P}_{ij\mathcal{U}m})\widehat{F}_{ij\mathcal{U}m,P_\mathcal{U}} + \widehat{R}_{ij\mathcal{U}m}\widehat{F}_{ij\mathcal{U}m,P_\mathcal{U}}\right] \\ & + \left[(1+T_{ijm}\widetilde{P}_{ij\mathcal{S}m})(1-\widehat{F}_{ij\mathcal{S}m}) + \widehat{R}_{ij\mathcal{S}m}\widehat{F}_{ij\mathcal{U}m}\right]\widehat{Q}_{ij\mathcal{S}m,P_\mathcal{U}}. \end{aligned} \tag{25}$$

If only one type $k \in \{\mathcal{S}, \mathcal{U}\}$ is offered, then the marginal costs implied by our model estimates are:

$$
\begin{aligned}
\widehat{MC}_{ijkm} &= (1+T_{ijm}\widetilde{P}_{ijkm})\left(1-\widehat{F}_{ijkm}-\widehat{F}_{ijkm,P_k}\frac{\widehat{Q}_{ijkm}}{\widehat{Q}_{ijkm,P_k}}\right) \\
&\quad +T_{ijm}(1-\widehat{F}_{ijkm})\frac{\widehat{Q}_{ijkm}}{\widehat{Q}_{ijkm,P_k}}+\widehat{R}_{ijkm}\left(\widehat{F}_{ijkm}+\frac{\widehat{Q}_{ijkm}}{\widehat{Q}_{ijkm,P_k}}\widehat{F}_{ijkm,P_k}\right), \qquad (26)
\end{aligned}
$$

where the recovery rates $\widehat{R}_{ij\mathcal{S}m}$ and $\widehat{R}_{ij\mathcal{U}m}$ are defined in equations (8) and (9). Note that these depend on the collateral value CV_{ijm}, which is observable for secured borrowers, but not for unsecured borrowers. Hence, for each unsecured borrower, we take the collateral value of their respective matched secured borrower found using propensity score matching.

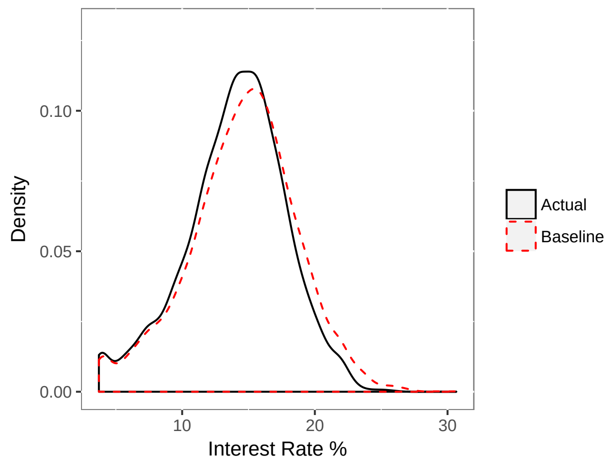
Table 5: Descriptives of Model Fit

	N. Obs	**Mean**	**Median**	**Std. Dev**	**% Neg.Profits**
Actual Interest Rate	4,564	13.97	14.26	3.81	3.92
Baseline Interest Rate	4,623	14.62	14.85	4.14	2.67
Actual Default	4,564	0.13	0.11	0.10	3.92
Baseline Default	4,623	0.14	0.11	0.11	2.67
Actual Demand	4,564	10,939	5,407	20,533	3.92
Baseline Demand	4,623	10,194	4,797	19,890	2.67
Actual Profit	4,564	807	187	2,150	3.92
Baseline Profit	4,623	806	187	2,159	2.67
Marginal Cost	4,564	1.19	1.09	0.25	3.92
Marginal Cost/Repayment	4,564	0.93	0.95	0.06	3.92

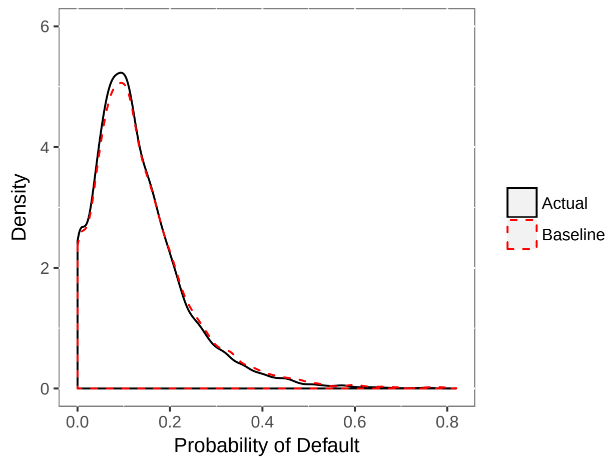
Note: This table summarizes the model fit results. For each variable we report both descriptive statistics from the data (Actual) and the model predicted equilibrium (Baseline). Interest Rate is in percentage points, Default is a probability, Demand is the product of demand probability and loan amount in USD. Profit is in USD. Four contracts that have positive profit but unrealistic prices are excluded.

Table 5 reports the descriptive statistics in terms of prices, default, demand and profits for our baseline model. The first row of each section ("Actual") reports the interest rates, demand, default, and profits observed in the data. The second row ("Baseline") shows the same equilibrium outcomes as predicted by our model in the baseline scenario. For each of these rows, we report the mean, median and standard deviations of the loan contracts that have non-negative profits. In fact, given that our model doesn't allow for borrower rejection, in a few cases of unprofitable borrowers, the equilibrium price is pushed to a very high level in order to minimize the borrower's demand probability and loan amount. Hence, we don't report the descriptive statistics for those cases and just summarize the share of contracts that have negative profits in the last column, under each scenario. In general, the model predicted equilibrium is very close to the actual observed outcomes, which is illustrated in the distribution of actual and model fit outcomes in Figure 6.

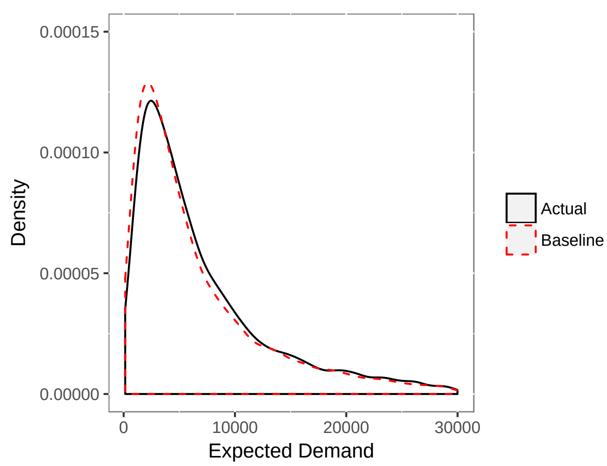
We analyze banks' marginal costs and profit margins, variables usually unobserved in the data, backing them



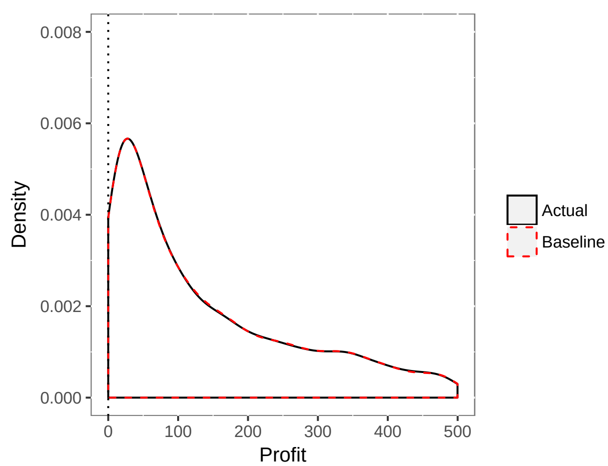
(a) Equilibrium Price



(b) Probability of Defaul



(c) Expected Demand



(d) Profit

Figure 6: Model Fit

Note: This figure shows the distribution of interest rate, demand, default, and profit for not rejected contracts. Expected demand is trimmed at 30,000$, which represents 94% of all contracts. Profit is trimmed at 500$, which represents 70% of all contracts.

Table 6: Marginal Cost

	Marginal Cost			Marginal Cost/Repayment		
	(1)	(2)	(3)	(4)	(5)	(6)
Collateral	0.070***	0.061***	0.049***	0.005***	0.008***	0.012***
	(0.007)	(0.007)	(0.007)	(0.002)	(0.002)	(0.002)
Low Quality		0.118***	0.108***		−0.041***	−0.037***
		(0.007)	(0.007)		(0.002)	(0.002)
Both			0.074***			−0.026***
			(0.008)			(0.002)
Constant	1.134***	1.078***	1.032***	0.935***	0.950***	0.967***
	(0.038)	(0.037)	(0.037)	(0.004)	(0.008)	(0.008)
Bank-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,564	4,564	4,564	4,564	4,564	4,564
R^2	0.112	0.163	0.179	0.094	0.229	0.267

Note: This table shows the OLS regression results of model implied marginal costs and the marginal cost to repayment on different loan and firm characteristics. The dependent variable in Column (1) to (3) is the model implied marginal cost $\widehat{MC}_{ijkm}$, and in Column (4) to (6) is the marginal cost to the repayment obligation $\widehat{MC}_{ijkm}/(1+T_{ijkm}\widetilde{P}_{ijkm})$. *Collateral* is an indicator which equals one for secured loan. *Low Quality* indicates firms with probability of default based on observed characteristics above median. *Both* equals one if a loan belongs to a pair of secured and unsecured loans that are offered by one bank to the same borrower. *p<0.1; **p<0.05; ***p<0.01.

out from our model's first-order condition. These model-implied marginal costs capture the overall cost of lending an extra dollar, including among other things funding, screening, and monitoring costs. We can then calculate how profitable an extra dollar lent is by looking at the ratio of marginal cost to the repayment obligation, as the ratio suggests that the bank can extract high margins from lending. Table 6 presents regression results of model implied marginal costs on loan and borrower characteristics. The dependent variable in the first three columns are the model implied-marginal costs, $\widehat{MC}_{ijkm}$, while in the last three columns are the marginal costs divided by the repayment amount $1 + T_{ijkm}\widetilde{P}_{ijkm}$. As shown in Columns (1) and (4), secured loans have higher marginal costs both in absolute and relative terms, indicating that secured loans are more costly and less profitable than unsecured ones. The marginal cost of lending one dollar with collateral is 0.07 dollar higher than that of unsecured loans, equivalent to 5.9% of the average marginal cost. On average, the marginal cost represents 93.5% of the repayment for unsecured loans, while it represents 94% for secured ones. For low quality firms, whose probability of default based on observed risk is above the median, banks have higher marginal costs to lend. However, the marginal costs to total repayment obligation is lower, meaning that banks can charge higher interest rates for borrowers with a high probability of default, and hence the bank's profit margin is higher. In Columns (3) and (6), we add another variable that indicates whether banks provide both secured and unsecured contracts to a firm. That is, banks are using collateral for screening. We find that screening is costly, as providing both types of contracts implies higher marginal costs for banks, but yields higher profit margins. This result highlights the information rent banks obtain by screening, captured by the last term of equation (10).

6 Counterfactuals

We conduct three counterfactual policy experiments to quantify the credit demand and supply responses to a shock to collateral value, and understand the role of asymmetric information within the collateral channel. First, we simulate a 40% drop in collateral value and quantify the changes in lenders' profits and interest rates, and in borrowers' demand and default. Second, we increase the extent of adverse selection and document how the effectiveness of collateral changes. Last, we introduce simultaneously both a shock to collateral value and an increase in adverse selection, to show how the extent of the agency problem can mitigate the collateral channel.

6.1 Collateral Value Shock

We use the estimates of our demand and default models, together with our supply-side framework, to understand how a shock to collateral values propagates to credit supply, credit allocation, interest rates, and banks' profits. Through this exercise we aim to separately identify credit demand and supply responses within the collateral channel. We first simulate a scenario where the collateral value CV_{ijm} drops by 40%. This is similar to the magnitude of various collateral value shocks documented in the literature, such as the burst of the Japanese assets price bubble that caused land prices in Japan dropped by 50% between 1991 and 1993 (Gan 2007), the nearly 30% drop of the Case-Shiller 20-City Composite Home Price Index in the U.S.

during the 2007-2009 financial crisis, and the rise in average repo haircut on seven categories of structured debt from zero in August 2007 to 45% in December 2008 (Gorton 2010). This gives rise to various effects through our model. First, it affects directly banks' profits from secured loans through the level of collateral, implying that banks will change their equilibrium interest rate, which will in turn affect demand and default. Also banks' profits from unsecured loans are affected, as some borrowers might now change their choice between a secured and an unsecured loan, which will in turn imply a change in equilibrium interest rates also for uncollateralized loans. This highlights how our model is able to capture spillover effects of the collateral channel from secured to unsecured loans, a novel result compared to the existing literature.

Assuming that the bank's marginal costs of lending to each firm remains constant in the counterfactual scenario, we simulate a 40% drop in collateral value and find the new equilibrium in terms of interest rate P_{ijkm}, probability of default F_{ijkm}, expected demand Q_{ijkm}, and banks' expected profit Π_{ijkm}. Table 7 summarizes the new equilibrium after the collateral value shock compared with the baseline model. We find that a 40% decrease in collateral value generates on average an 11.3% and 8.6% increase in the interest rates of secured and unsecured loans that have non-negative profits, respectively. Overall, the interest rate increases by 9.9%, namely 1.3 percentage points. The probability of default increases by 6.8% on average. The expected demand and profit drop significantly, especially for secured loans, with a 17.9% and a 23.2% decrease, respectively. Crucially, we find that the share of non-profitable offered loans goes from around 3% (1.1% for secured and 3.9% for unsecured loans) in the baseline case to 19.5% (14.8% for secured and 23.2% for unsecured loans). Although we observe a large increase in interest rate and drop in demand and profit for secured loans, a larger portion of unsecured loans become unprofitable when the collateral value drops. Credit supply shrinks as borrowers face higher financing costs and less loan offers. The distribution of the percentage changes in the new equilibrium are depicted in Figure 7. These results are qualitatively in line with the findings in Cerqueiro, Ongena and Roszbach (2016), who investigate how a legal change in Sweden reduces the collateral value by 13% for outstanding loans, generating a 0.2 percentage points decrease in interest rate, an 11% decrease in internal credit limit, and 12 percentage points more delinquent borrowers.

Table 7: Collateral Value Shock

	Percentage Change								**% Neg. Profit**
	Δ Interest Rate		**Δ Default**		**Δ Demand**		**Δ Profit**		
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	
Secured	11.30	18.51	7.81	13.06	-17.91	28.01	-23.20	32.50	14.83
Unsecured	8.63	19.19	5.96	12.54	-11.90	28.39	-15.91	33.35	23.20
Total	9.88	18.92	6.83	12.81	-14.71	28.37	-19.33	33.15	19.49

Note: This table summarizes the average percentage change in equilibrium price (*Interest Rate*), probability of default (*Default*), expected demand (*Demand*) and banks' profit (*Profit*) of non-negative profit loans after collateral value drop by 40% compared with the baseline model. *% Neg. Profit* is the percentage of loans with negative profit. Each row stands for the summary for secured loan, unsecured loan, and the both kinds.

These results quantify the relevance of various components of the mechanism at play in our model, following

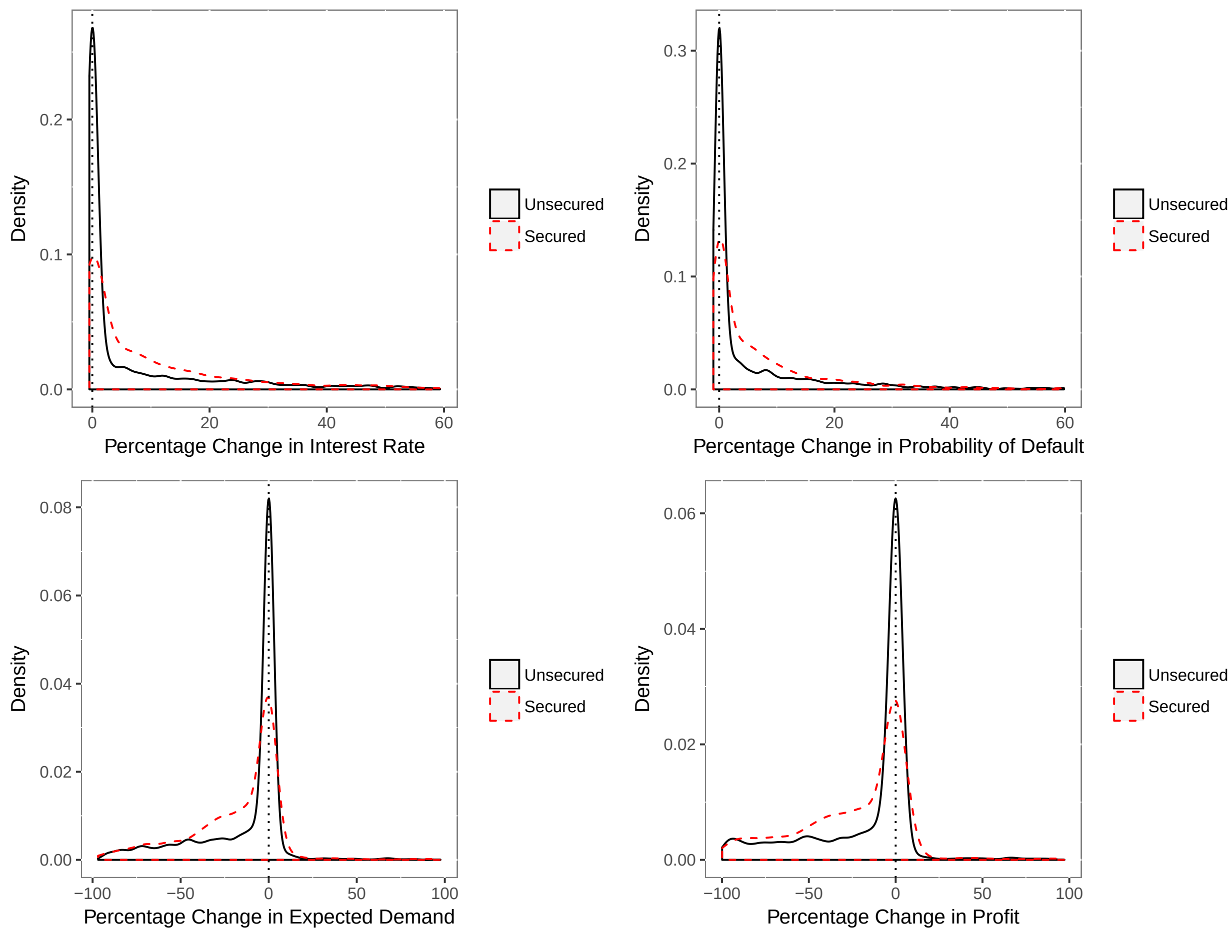


Figure 7: Collateral Value Shock

Note: This figure shows the distribution of the percentage changes in interest rate, probability of default, expected demand and banks' profit for non-negative profit loans after collateral value drop by 40 % compared with the baseline model.

up on the discussion at the end of Section 3.2. A shock to collateral value directly impacts lenders' profits through the recovery rate term. Banks respond to this shock increasing the interest rate on both secured and unsecured loans, as they can use both margins to make up for this potential profit loss. The heterogeneity in these price responses is likely to be driven by both the average borrowers' default rate (F_{ijkm}) and banks' markup terms, as can be seen in equation (10). As expected, borrowers respond to this by reducing their credit demand and increasing their likelihood of default, through the moral hazard channel α_1^F. Another driver of the larger increase in interest rates for secured loans compared to unsecured ones is adverse selection, because safe borrowers are the most price sensitive ones, and the larger price increase might induce them to switch to unsecured loans, worsening the pool of borrowers choosing collateralized loans. In other words, the increase in interest rates for unsecured loans is also determined by the riskiness of the marginal borrowers who switch away from secured loans.

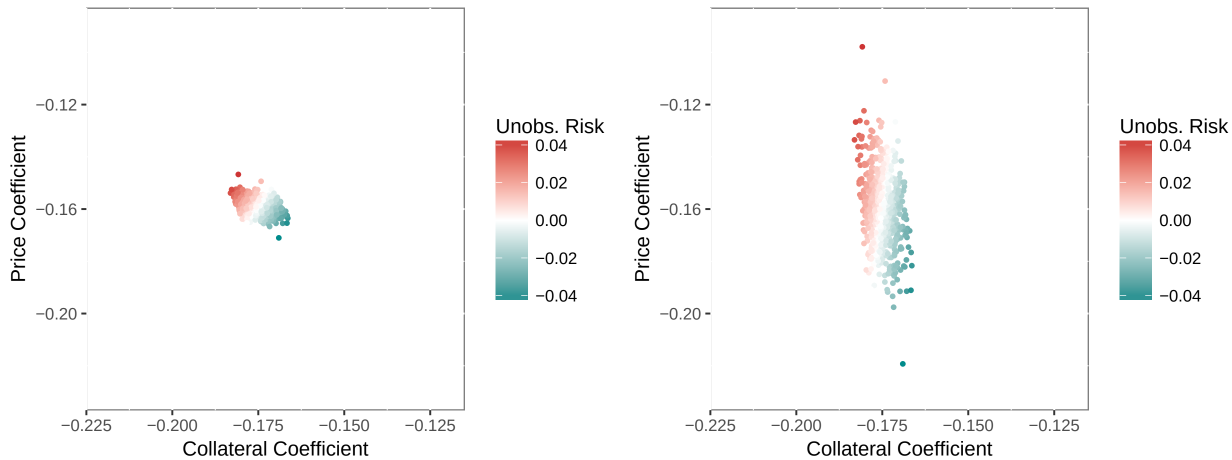
An interesting implication of this shock to collateral value is that the difference in interest rates between secured and unsecured loans is now reduced. In the baseline case, the average interest rate of unsecured loans is 1.1 percentage points higher than the secured loans, with a 95 confidence interval between 0.8 and 1.3 percentage points, while in the collateral value shock scenario, the average interest rate difference drops to 0.7 percentage points with a 95 confidence interval between 0.4 and 1 percentage points. This implies that one of the consequences of the collateral channel is making it less profitable to induce separation between secured and unsecured loans, which can potentially increase the extent of agency costs.

6.2 Adverse Selection Shock

The second counterfactual exercise investigates how the extent of adverse selection affects the effectiveness of collateral as a screening mechanism. In our model, the existence of adverse selection is captured by the positive correlation between unobserved borrowers' riskiness and their price random coefficients ($\rho_{\mathcal{P}F}$). However, if we want to make adverse selection more severe while keeping borrowers' unobserved risk constant, we can only do it by increasing the standard deviation of borrowers' preferences for interest rates ($\sigma_{\mathcal{P}}$). In fact, as shown in equation (18), changing $\rho_{\mathcal{P}F}$ would also affect $\widetilde{\mu}_{Fi}$. Hence, we increase five times the standard deviations of the price random coefficient $\sigma_{\mathcal{P}}$.[21] Figure 8 gives a graphical intuition for this counterfactual, showing how the change in $\sigma_{\mathcal{P}}$ from the baseline (left figure) to the counterfactual (right figure) leads to a more dispersed distribution of borrowers in their preference space, while holding unobserved risk fixed. This can be interpreted as an increase in adverse selection, as riskier borrowers are now even more likely to take credit (less price sensitive), hence moving further towards the top of the figure, while the opposite happens for safe borrowers.

The new equilibrium after the adverse selection shock is summarized in Table 8 and Figure 9. We find that the secured and the unsecured loan change in the opposite direction for all outcomes. When adverse selection increases, the interest rate of secured loans decreases by 0.37%, while it increases by 0.06% for

[21] We arbitrarily chose a fivefold increase as we don't have a benchmark for what a reasonable increase in adverse selection could be. Note that smaller or larger increases in this parameter would just lead to smaller or larger changes in the relevant outcomes of our model.



(a) Baseline Model

(b) Adverse Selection Shock

Figure 8: Price and Collateral Coefficients in the Preference Shock

Note: This figure shows the price and collateral coefficients for each firm in the market. Subfigure (a) is the Baseline model, which is the same as Figure 5 (b). Subfigure (b) is the distribution of price and collateral coefficients when $\sigma_{\mathcal{P}}$ is five times larger than the baseline model.

Table 8: Adverse Selection Shock

	Percentage Change								**% Neg. Profit**
	Δ Interest Rate		**Δ Default**		**Δ Demand**		**Δ Profit**		
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	
Secured	-0.37	0.86	-0.26	0.80	-0.21	4.02	-0.91	3.88	1.14
Unsecured	0.06	0.73	0.12	2.11	-1.11	4.04	-0.70	3.65	3.93
Total	-0.14	0.82	-0.05	1.67	-0.71	4.05	-0.79	3.76	2.69

Note: This table summarizes the average percentage change in equilibrium price (*Interest Rate*), probability of default (*Default*), expected demand (*Demand*) and banks' profit (*Profit*) of non-negative profit loans after $\sigma_{\mathcal{P}}$ increased by five times compared with the baseline model. *% Neg. Profit* is the percentage of loans with negative profit. Each row stands for the summary for secured loan, unsecured loan, and the both kinds.

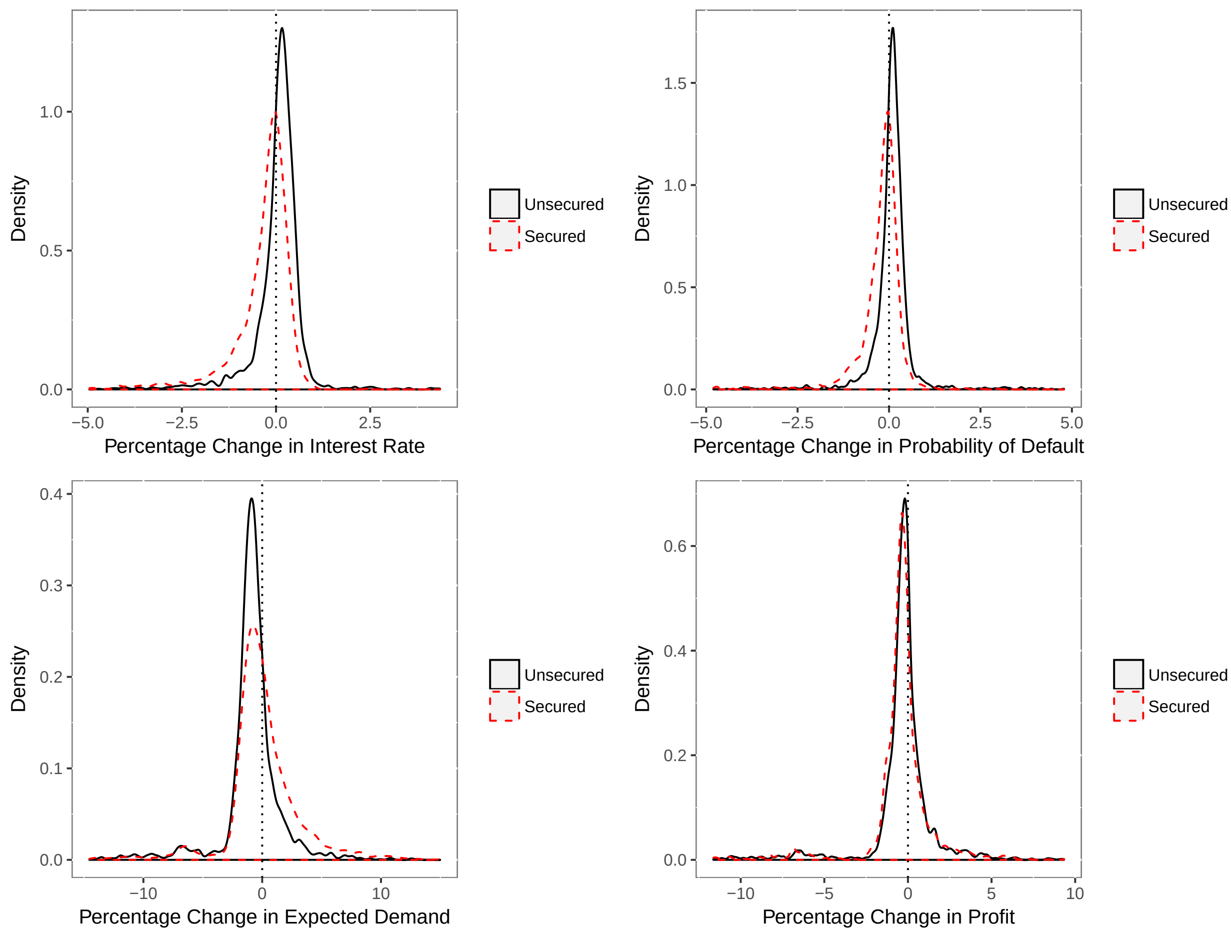


Figure 9: Preference Shock

Note: This figure shows the distribution of the percentage changes in interest rate, probability of default, expected demand and banks' profit for non-negative profit loans after $\sigma_{\mathcal{P}}$ increased by five times compared with the baseline model.

unsecured loans. On average, higher adverse selection lowers the overall interest rate by 0.14%. The probability of default changes accordingly in the same direction as their respective interest rate change. In the new equilibrium, the pool of borrowers attracted by secured loans are less risky than before with a 0.26% decrease in probability of default, while unsecured loans have a riskier pool of borrowers with 0.12% increase in probability of default. On average, expected demand decreases and banks' profits drop by 0.8%. The percentage of negative profit loans is as in the baseline model.

This counterfactual shows that when adverse selection becomes more severe it is easier for lenders to achieve separation of safe and risky borrowers in secured and unsecured loans, because of the larger polarization in their $MRS_{\mathcal{C},\mathcal{P}}$. This implies that the average risk of borrowers choosing unsecured loans increases, whereas that of borrowers choosing secured loans decreases, as reflected by the rise in interest rates for the former and the drop for the latter ones. The decline in default rates for secured loans is caused both by this selection effect and by a reduction in moral hazard through the direct effect of a lower price on default. The opposite happens for default rates of unsecured loans. This simulation exercise shows how collateral becomes a more effective instrument the larger is the extent of adverse selection. It causes an increase in borrowers' surplus, through lower interest rates, but a decrease in lenders' expected profits.

6.3 Collateral Value and Adverse Selection Shock

In the last counterfactual we investigate whether the extent of adverse selection can mitigate the propagation of a shock to collateral value. Therefore, we combine the two previous simulation exercises into one, by simultaneously dropping the collateral value by 40% and raising adverse selection with a fivefold increase in $\sigma_{\mathcal{P}}$. We summarize in Table 9 and Figure 10 the new equilibrium results compared with the baseline case (i.e. no collateral value nor adverse selection shock) across the four relevant outcomes: interest rates, default, demand, and banks' profits.

When the collateral value shock occurs in a market with high adverse selection, we observe that interest rates increase on average. However, compared with the baseline adverse selection case, the interest rate for secured loans increases less (10.7% vs. 11.3%), whereas the interest rate for unsecured loans increases almost by the same amount (8.5% vs. 8.6%). Overall, interest rates increase by less with high adverse selection in response to a shock to collateral value. Similarly, on average the probability of default increases less when adverse selection is more serious. This implies that if a shock to collateral value hurts the effectiveness of collateral screening, higher adverse selection works in the opposite direction, making the propagation of the collateral channel less severe than in the baseline case, in terms of interest rates and default. This however is mostly driven by the lower price and default increase for secured loans, while instead the larger default increase for unsecured loans reduces overall demand and banks' profits by more than in the baseline case.

Table 9: Collateral Value and Adverse Selection Shock

	Percentage Change								
	Δ Interest Rate		**Δ Default**		**Δ Demand**		**Δ Profit**		**% Neg. Profit**
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	
Secured	10.71	18.32	7.22	13.17	-18.62	30.62	-24.51	34.03	14.88
Unsecured	8.49	18.83	6.18	14.00	-13.49	33.11	-16.96	37.49	23.24
Total	9.53	18.63	6.67	13.62	-15.89	32.07	-20.50	36.10	19.54

Note: This table summarizes the mean and standard deviation of percentage changes in equilibrium price (*Interest Rate*), probability of default (*Default*), expected demand (*Demand*) and banks' profit (*Profit*) of non-negative profit loans after collateral value dropped by 40% and $\sigma_{\mathcal{P}}$ increased by five times compared with the baseline model. *% Neg. Profit* is the percentage of loans with negative profit. Each row stands for the summary for secured loan, unsecured loan, and the both kinds.

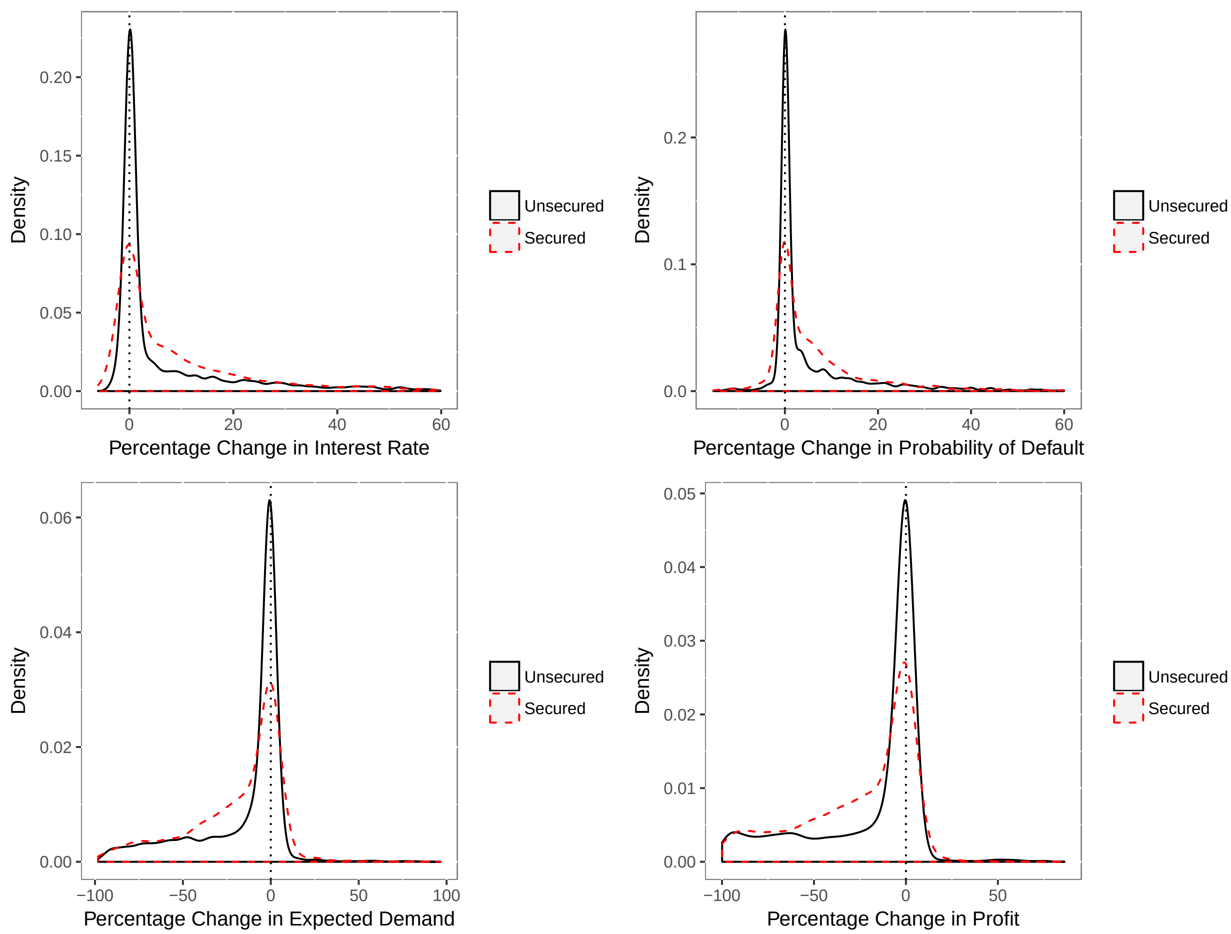


Figure 10: Preference Shock and Collateral Value Shock

Note: This figure shows the distribution of the percentage changes in interest rate, probability of default, expected demand and banks' profit for non-negative profit loans after a 40% drop in collateral value and $\sigma_{\mathcal{P}}$ increased by five times compared with the baseline model.

6.4 Summary of Counterfactuals

We provide additional evidence of the main mechanisms driving the results in our counterfactuals, further investigating how collateral values and the level of adverse selection affect the effectiveness of collateral as a screening device. We estimate a simple regression model using our baseline and counterfactual results to understand the relationship between borrowers' likelihood of choosing a secured loan, given by the corresponding estimated demand probabilities, and their unobserved riskiness, defined as our estimate of $\widetilde{\mu}_{Fi}$ from equation (18). We take as unit of observation each bank-firm combination for which a lender offers both a secured and an unsecured loan, and use as dependent variable in an OLS regression the probability of choosing a secured loan from each bank, conditional on having chosen that specific bank. We estimate this model for our baseline case and for the three counterfactuals we run, and summarize the results in Table 10. We include the interest rate of secured loans, as well as fixed effects for bank, loan amount and maturity, industry, and quarter.

In the baseline model, that is the first column on Table 10, we find that the probability of choosing a secured loan is negatively related to borrowers' unobserved risk, which implies that safe borrowers are more likely to choose a secured loan. In particular, one standard deviation increase in a borrower's unobserved risk leads to a 0.3 percentage points decrease in her probability of choosing a secured loan. This is consistent with collateral mitigating adverse selection problems by inducing separation of borrowers of different risk. However, once we shock the collateral value, the screening effect of collateral vanishes, as can be seen in the second column of Table 10. This reinforces the conclusion stated at the end of Section 6.1, as the drop in collateral value decreases lenders' profits from secured loans, which in turn decreases their incentive to differentiate between safe and risky borrowers using collateral. Moreover, from the borrowers' perspective, the collateral value shock increases significantly the interest rate on secured loans, which decreases safe borrowers' demand for secured loans.

On the contrary, an increase in adverse selection leaves the screening effect of collateral roughly unchanged, as reflected by the negative coefficient on unobserved risk in the third column of Table 10 compared to the baseline case. Moreover, the demand for secured loans also increases on average. This is in line with the findings in Section 6.2, as higher adverse selection makes it easier for banks to separate borrowers of different risk using collateral, which results in lower average risk of borrowers choosing secured loans and therefore lower prices, and the other way around for borrowers choosing unsecured loans. The last column in Table 10 refers to the third counterfactual, the collateral value shock with high adverse selection. Differently from the results in the second column, under this scenario the higher level of adverse selection counterbalances the collateral channel, allowing collateral to still be an effective screening device. We find in fact a negative relationship between borrowers' unobserved riskiness and likelihood to choose a secured loan, and the magnitude of this effect is somewhere between the results in the second and third column. This effect is however not statistically significant. This suggests that higher adverse selection mitigates the drop in screening effectiveness of collateral caused by a negative shock to collateral value. However, the cost for banks to maintain screening is higher, as shown by the larger drop in profits in Table 9. This implies that drops in collateral values are more destabilizing to banks profitability in markets with high adverse selection.

Table 10: The Effectiveness of the Collateral as a Screening Device

	Prob. Secured			
	Baseline	CV	AS	CV & AS
Unobserved Risk	-0.206^{**}	-0.116	-0.198^{***}	-0.165
	(0.089)	(0.124)	(0.066)	(0.124)
Interest Rate	-0.002^{***}	-0.002^{***}	0.006^{***}	-0.007^{***}
	(0.001)	(0.0003)	(0.001)	(0.001)
Constant	0.472^{***}	0.491^{***}	0.452^{***}	0.491^{***}
	(0.008)	(0.011)	(0.006)	(0.011)
Bank FE	Yes	Yes	Yes	Yes
Amount FE	Yes	Yes	Yes	Yes
Maturity FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	1,607	1,607	1,607	1,607
R^2	0.072	0.097	0.066	0.072

Note: This table summarizes OLS regression results. The unit of observation is a borrower-bank combination, conditional on the bank offering both secured and unsecured loans to the borrower. Interest Rate is the interest rate of the secured loan. The dependent variable is conditional probability of choosing the secured contract from the pair of contracts provided by the same bank. The explanatory variable *Unobserved Risk* is the simulated unobserved risk. The four columns correspond to baseline model, collateral value shock, adverse selection shock, collateral value and adverse selection shock. *p<0.1; **p<0.05; ***p<0.01.

7 Conclusion

In this paper we study the benefits and costs of collateral requirements in bank lending markets with asymmetric information. We develop a structural model of firms' credit demand for secured and unsecured loans, banks' contract offering and pricing, and firm default using detailed credit registry data on corporate loans and borrowers' performance from Bolivia, a country where asymmetric information problems in credit markets are pervasive. We make three important contributions to the literature.

First, by modelling borrowers' demand for secured and unsecured credit, we provide micro-founded evidence of the benefits of collateral pledging, estimating structural parameters that measure both the ex ante and ex post reduction in agency costs that collateral determines. We provide evidence supporting both the ex ante and ex post theories of collateral. Consistent with the ex ante theories, we find a negative and significant correlation of -0.81 between borrowers' sensitivity to collateral and their default unobservables, which suggests that borrowers with high default risk tend to have high disutility from pledging collateral, and are therefore less likely to demand a secured loan compared to safe borrowers. Furthermore, we provide empirical evidence that riskier borrowers have a higher marginal rate of substitution of collateral for price, a key assumption in the ex ante theories which, to the best of our knowledge, has never been tested before. Consistent with the ex post theories, we find a negative and significant causal effect of collateral on default, suggesting that on average posting collateral decreases the probability of default by 88.8%.

Second, by modelling also lenders' supply of both collateralized and uncollateralized loans, we are able to separately quantify the role of credit demand and supply within the collateral channel, accounting for their interaction. We simulate the effects of a 40% drop in collateral value on credit supply, credit allocation, interest rates, and banks' profits. We find that almost 20% of loans would become unprofitable under this scenario, while the remaining ones would experience a 10% increase on average in interest rates, a 15% reduction in average demand, and a 19% decrease in banks' profits.

Third, we can study how the use of collateral and the propagation of collateral shocks is influenced by asymmetric information frictions. We find that when adverse selection becomes more severe it is easier for lenders to achieve separation of safe and risky borrowers in secured and unsecured loans. As a result, stronger adverse selection mitigates the propagation of the collateral channel, making the increases in loan interest rates and default in response to a shock to collateral value less pronounced. We also find, however, that when adverse selection is high, banks suffer larger drops in expected profits as the use of collateral for screening reduces their profit margins ex ante.

Overall, our results indicate that collateral has a large impact on firms' access and terms of credit. Swings in collateral values have a large effect on the fraction of borrowers that are able to obtain credit, as well as on the amount and terms of credit, by altering banks' expected profitability and equilibrium loan interest rates. Our work opens the floor for various other potential directions of research. First, our approach could be extended to quantify not only how the severity of adverse selection, but also how the severity of moral hazard can influence the propagation of shocks to collateral values. This would have important implications for policymakers, who could then prioritize their interventions on the key friction. Second, the current analysis holds banks' marginal cost of funds constant. Additional counterfactual experiments could allow this to change, providing insights on monetary policy's role on the transmission of shocks to collateral values. Third, our model could be extended by allowing loan maturity to be another screening and monitoring dimension, substituting or complementing the use of collateral. Last, this framework could be used to investigate how policy interventions aimed at improving lenders' recovery rates could mitigate the negative effects of a shock to collateral value. We regard all of these as promising directions of future research.

References

Acharya, V. V., Bharath, S. T. and Srinivasan, A. (2007), 'Does industry-wide distress affect defaulted firms? evidence from creditor recoveries', *Journal of Financial Economics* **85**(3), 787–821.

Adams, W., Einav, L. and Levin, J. (2009), 'Liquidity constraints and imperfect information in subprime lending', *American Economic Review* **99**(1), 49–84.

Adelino, M., Schoar, A. and Severino, F. (2015), 'House prices, collateral, and self-employment', *Journal of Financial Economics* **117**, 288–306.

Aghion, P. and Bolton, P. (1997), 'A theory of trickle-down growth and development', *Review of Economic Studies* **64**(2), 151–172.

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

Asarnow, E. and Edwards, D. (1995), 'Measuring loss on defaulted bank loans: A 24-year study', *The Journal of Commercial Lending* **77**(7), 11–23.

Banerjee, A. and Newman, A. (1993), 'Occupational choice and the process of development', *Journal of Political Economy* **101**(2), 274–298.

Barro, R. J. (1976), 'The loan market, collateral, and rates of interest', *Journal of Money, Credit and Banking* **8**(4), 439–456.

Benetton, M. (2017), Leverage regulation and market structure: An empirical model of the uk mortgage market. Working Paper.

Benmelech, E. and Bergman, N. K. (2008), 'Liquidation values and the credibility of financial contract renegotiation: Evidence from u.s. airlines', *Quarterly Journal of Economics* **123**(4), 1635–1677.

Benmelech, E. and Bergman, N. K. (2009), 'Collateral pricing', *Journal of Financial Economics* **91**, 339–360.

Benmelech, E. and Bergman, N. K. (2011), 'Bankruptcy and the collateral channel', *Journal of Finance* **LXVI**(2), 337–378.

Benmelech, E., Garmaise, M. J. and Moskowitz, T. J. (2005), 'Do liquidation values affect financial contracts? evidence from commercial loan contracts and zoning regulation', *Quarterly Journal of Economics* **120**(3), 1121–1154.

Berger, A. N., Espinosa-Vega, M. A., Frame, W. S. and Miller, N. H. (2011), 'Why do borrowers pledge collateral? new empirical evidence on the role of asymmetric information', *Journal of Financial Intermediation* **20**(1), 55–70.

Berger, A. N., Frame, W. S. and Ioannidou, V. (2011), 'Tests of ex ante versus ex post theories of collateral using private and public information', *Journal of Financial Economics* **100**(1), 85–97.

Berger, A. N. and Udell, G. F. (1995), 'Relationship lending and lines of credit in small firm finance', *Journal of Business* **68**(3), 351–381.

Bernanke, B. (1983), 'Nonmonetary effects of the financial crisis in the propagation of the great depression', *American Economic Review* **73**(3), 257–276.

Bernanke, B. and Gertler, M. (1989), 'Agency costs, net worth, and business fluctuations', *American Economic Review* **79**(1), 14–31.

Besanko, D. and Thakor, A. V. (1987*a*), 'Collateral and rationing: Sorting equilibria in monopolistic and competitive credit markets', *International economic review* **28**(3), 671–689.

Besanko, D. and Thakor, A. V. (1987*b*), 'Competitive equilibrium in the credit market under asymmetric information', *Journal of Economic Theory* **42**(1), 167–182.

Bester, H. (1985), 'Screening vs. rationing in credit markets with imperfect information', *American Economic Review* **75**(4), 850–855.

Boot, A. W. A. and Thakor, A. V. (1994), 'Moral hazard and secured lending in an infinitely repeated credit market game', *International Economic Review* **35**(4), 899–920.

Boot, A. W. A., Thakor, A. V. and Udell, G. F. (1991), 'Secured lending and default risk: Equilibrium analysis, policy implications and empirical results', *Economic Journal* **101**(406), 458–472.

Boyd, J. H. and Smith, B. D. (1994), 'How good are standard debt contracts? stochastic versus nonstochastic monitoring in a costly state verification environment', *Journal of Business* **67**(4), 539–561.

Caliendo, M. and Kopeinig, S. (2008), 'Some practical guidance for the implementation of propensity score matching', *Journal of Economic Surveys* **22**(1), 31–72.

Cerqueiro, G., Ongena, S. and Roszbach, K. (2016), 'Collateralization, bank loan rates, and monitoring', *Journal of Finance* **LXXI**(3), 1295–1322.

Cerqueiro, G., Ongena, S. and Roszbach, K. (2018), Collateral damaged? priority structure, credit supply, and firm performance. Working Paper.

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

Chan, Y.-S. and Thakor, A. V. (1987), 'Collateral and competitive equilibria with moral hazard and private information', *Journal of Finance* **42**(2), 345–363.

Chaney, T., Sraer, D. and Thesmar, D. (2012), 'The collateral channel: How real estate shocks affect corporate investment', *American Economic Review* **102**(6), 2381–2409.

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

Cooley, T., Marimon, R. and Quadrini, V. (2004), ‘Aggregate consequences of limited contract enforceability’, *Journal of Political Economy* **112**(4), 817–847.

Corradin, S. and Popov, A. (2015), ‘House prices, home equity borrowing, and entrepreneurship’, *Review of Financial Studies* **28**(8), 2399–2428.

Crawford, G. S., Pavanini, N. and Schivardi, F. (2018), ‘Asymmetric information and imperfect competition in lending markets’, *American Economic Review* **108**(7), 1659–1701.

Davydenko, S. A. and Franks, J. R. (2008), ‘Do bankruptcy codes matter? a study of defaults in france, germany, and the uk’, *The Journal of Finance* **63**(2), 565–608.

De Janvry, A., Sadoulet, E., McIntosh, C., Wydick, B., Luoto, J., Gordillo, G. and Schuetz, G. (2003), ‘Credit bureaus and the rural microfinance sector: Peru, guatemala, and bolivia’, *Collaborative Research Support Program. The University of California and The FAO Office for Latin America* .

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

Egan, M., Hortaçsu, A. and Matvos, G. (2017), ‘Deposit competition and financial fragility: Evidence from the us banking sector’, *American Economic Review* **107**(1), 169–216.

Einav, L., Jenkins, M. and Levin, J. (2012), ‘Contract pricing in consumer credit markets’, *Econometrica* **80**(4), 1387–1432.

Ersahin, N. and Irani, R. M. (2018), Collateral shocks and corporate employment. Working Paper.

Gale, D. and Hellwig, M. (1985), ‘Incentive-compatible debt contracts: The one-period problem’, *Review of Economic Studies* **52**(4), 647–663.

Gan, J. (2007), ‘Collateral, debt capacity, and corporate investment: Evidence from a natural experiment’, *Journal of Financial Economics* **85**(3), 709–734.

Gonas, J. S., Highfield, M. J. and Mullineaux, D. J. (2004), ‘When are commercial loans secured?’, *Financial Review* **39**(1), 79–99.

Gorton, G. B. (2010), *Slapped by the invisible hand: The panic of 2007*, Oxford University Press.

Hart, O. and Moore, J. (1994), ‘A theory of debt based on the inalienability of human capital’, *Quarterly Journal of Economics* **109**(4), 841–879.

Hausman, J. A. and Taylor, W. E. (1981), ‘Panel data and unobservable individual effects’, *Econometrica* **49**(6), 1377–1398.

Ho, K. and Ishii, J. (2011), ‘Location and competition in retail banking’, *International Journal of Industrial Organization* **29**(5), 537–546.

Holmstrom, B. and Tirole, J. (1997), ‘Financial intermediation, loanable funds, and the real sector’, *Quarterly Journal of Economics* **122**(3), 663–691.

Karlan, D. and Zinman, J. (2009), ‘Observing unobservables: Identifying information asymmetries with a consumer credit field experiment’, *Econometrica* **77**(6), 1993–2008.

Kerr, S., Kerr, W. R. and Nanda, R. (2015), House money and entrepreneurship. Working Paper.

Khieu, H. D., Mullineaux, D. J. and Yi, H.-C. (2012), ‘The determinants of bank loan recovery rates’, *Journal of Banking & Finance* **36**(4), 923–933.

Kiyotaki, N. and Moore, J. (1997), ‘Credit cycles’, *Journal of Political Economy* **105**(2), 211–248.

Koijen, R. S. J. and Yogo, M. (2016), ‘Shadow insurance’, *Econometrica* **84**(3), 1265–1287.

Koijen, R. S. J. and Yogo, M. (2018), An equilibrium model of institutional demand and asset prices. forthcoming at *Journal of Political Economy*.

Mian, A. and Sufi, A. (2011), ‘House prices, home equity-based borrowing, and the us household leverage crisis’, *American Economic Review* **101**(5), 2132–2156.

Mian, A. and Sufi, A. (2014), ‘What explains the 2007-2009 drop in employment?’, *Econometrica* **82**(6), 2197–2223.

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

Schmalz, M., Sraer, D. A. and Thesmar, D. (2017), ‘Housing collateral and entrepreneurship’, *Journal of Finance* **LXII**(1), 99–132.

Sirtaine, S., Skamnelos, I. and Frank, S. (2004), ‘Bolivia: Challenges in the corporate and banking sectors’, *World Bank, Washington, DC* .

Stiglitz, J. and Weiss, A. (1981), ‘Credit rationing in markets with imperfect information’, *American Economic Review* **71**(3), 393–410.

Townsend, R. M. (1979), ‘Optimal contracts and competitive markets with costly state verification’, *Journal of Economic Theory* **21**, 265–293.

Train, K. E. (2009), *Discrete Choice Methods with Simulation*, Cambridge University Press.

Williamson, S. D. (1986), ‘Costly monitoring, financial intermediation, and equilibrium credit rationing’, *Journal of Monetary Economics* **18**, 159–179.

A Appendix For Online Publication

A.1 The Optimal Matching Algorithm

This section explains the process of determining the optimal matching algorithm. We first provide a simple example of matching. And then, we show how to determine the optimal matching algorithm based on the performance of matching observed unsecured loan (untreated) to observed secured loan (treated).

A Simple Example for Matching For example, if we only observe four firms obtained a loan with the same amount and maturity level in any region and time as shown in Table A.1. The four observed loans (in bold) are originated by two banks. Firm 1 had a secured loan with interest rate 14 percentage points per year from Bank A; Firm 2 and 3 had an unsecured loan with interest rate 16 p.p. and 15 p.p. respectively from Bank A; Firm 4 had an unsecured loan with interest rate 18 p.p. from Bank B. For Firm 1, to find the matches of unsecured loan by Bank A, we only focus on Firm 2 and 3. If the Firm 1's best match is Firm 3, then the interest rate of unsecured loan for Firm 1 by Bank A is 15 p.p.. As no secured loan is observed from Bank B, the secured loan from Bank B is missing. The match for unsecured loan from Bank B is Firm 4. Applying the same method, we can find the matches for all firms. For Firm 2 and 3, the (only) match for secured loan provided by Bank A is Firm 1; the (only) match for unsecured loan provided by Bank B is Firm 4. For Firm 4, if the closest match for unsecured loan provided by Bank A is Firm 2, then the matched interest rate is 16 p.p.. In the end, we obtain matched interest rate of secured and unsecured loan for the specific amount and maturity level provided by all banks in the market.

Table A.1: An Example for Matching

	Bank A		Bank B	
	Secured	Unsecured	Secured	Unsecured
Firm 1	**14**	15	-	18
Firm 2	14	**16**	-	18
Firm 3	14	**15**	-	18
Firm 4	14	16	-	**18**

Select variables for propensity score. Following Caliendo and Kopeinig (2008), two criteria are used to select variables for propensity score. First, the variables must be statistically significant for predicting the propensity score. Second, the variables are chosen to maximize the rate of correct prediction. For observed secured loans, if the estimated propensity score is larger than the fraction of secured loans in the sample, then the observation is classified as a correct prediction. We need to maximize the number of correct prediction.

At the beginning, the variable set only contains bank ID. Next, Relevant variables are added into the variable set. The added variables are kept only if they are statistically significant and can improve the number of correct prediction. The following variables are kept: Bank ID, amount category, maturity category, first loan, corporation. The propensity score generated by these variables give rise to 580 correct predictions out of 842 secured loans.

Table A.2: Matching Results 1

Variable	Unmatched vs. Matched	Mean Treated	Mean Control	%Reduction %bias	%Reduction |bias|	t-Test t	t-Test p>|t|
Bank 3	U	.09609	.10374	-2.6		-0.62	0.536
	M	.10815	.10815	0.0	100.0	0.00	1.000
Bank 5	U	.172	.12979	11.8		2.95	0.003
	M	.20444	.20444	0.0	100.0	0.00	1.000
Bank 7	U	.06524	.00246	35.2		10.94	0.000
	M	.01481	.01481	0.0	100.0	-0.00	1.000
Bank 8	U	.0427	.00246	27.3		8.39	0.000
	M	.01037	.01037	0.0	100.0	-0.00	1.000
Bank 9	U	.13879	.24336	-26.8		-6.28	0.000
	M	.16593	.16593	0.0	100.0	0.00	1.000
Bank 10	U	.0866	.04966	14.7		3.79	0.000
	M	.08741	.08741	0.0	100.0	-0.00	1.000
Bank 11	U	.09727	.15388	-17.1		-4.03	0.000
	M	.11704	.11704	0.0	100.0	0.00	1.000
Bank 14	U	.09253	.09145	0.4		0.09	0.927
	M	.0963	.0963	0.0	100.0	0.00	1.000
Bank 16	U	.03677	.03933	-1.3		-0.32	0.746
	M	.03259	.03259	0.0	100.0	0.00	1.000
Bank 17	U	.00119	.00295	-3.9		-0.87	0.382
	M	0	0	0.0	100.0	.	.
Bank 18	U	.07117	.03933	14.0		3.62	0.000
	M	.04148	.04148	0.0	100.0	0.00	1.000
Amount: 15,000$ to 30,000$	U	.20403	.24926	-10.8		-2.60	0.009
	M	.20148	.20148	0.0	100.0	-0.00	1.000
Amount: 30,000$ to 90,000$	U	.23369	.24385	-2.4		-0.58	0.562
	M	.22963	.22963	0.0	100.0	-0.00	1.000
Amount: 90,000$ to 12,000,000$	U	.33926	.21239	28.7		7.22	0.000
	M	.33926	.33926	0.0	100.0	0.00	1.000
Maturity: 3 to 6 months	U	.21352	.26008	-11.0		-2.64	0.008
	M	.21333	.21333	0.0	100.0	0.00	1.000
Maturity: 6 to 18 months	U	.16845	.2822	-27.5		-6.47	0.000
	M	.15407	.15407	0.0	100.0	-0.00	1.000
Maturity: 18 to 180 months	U	.40095	.18732	48.2		12.36	0.000
	M	.42222	.42222	0.0	100.0	-0.00	1.000
First Loan	U	.60142	.75959	-34.4		-8.65	0.000
	M	.62963	.62963	0.0	100.0	0.00	1.000
Corporation	U	.58363	.60226	-3.8		-0.93	0.354
	M	.62815	.62815	0.0	100.0	0.00	1.000

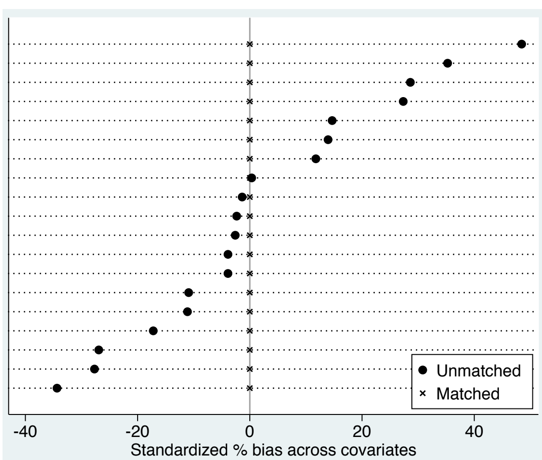
Table A.3: Matching Results 2

Sample	Pseudo-R2	LR χ^2	$p > \chi^2$	Mean Bias	Med Bias	Rubin's B	Rubin's R
Unmatched	0.176	610.97	0.000	17.0	13.9	104.0*	1.79
Matched	-0.000	-0.00	1.000	0.0	0.0	0.0	1.00

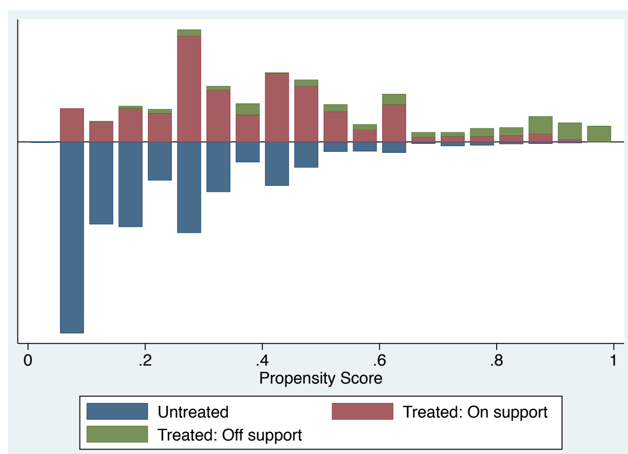
Note: This table summarizes the statistics before (Unmatched) and after (Matched) matching. A rule of thumb for a good match is to have mean and median bias below 3% to 5%, Rubin's B below 25% and Rubin's R between 0.5 and 2.

Choose the algorithm. To ensure that after matching the covariates are as close as possible between secured and unsecured loans, we set the radius very close to zero such that the matched loan must share the same characteristics as the loan to be matched (exact matching). If there are more than one loan that have the same characteristics, we randomly choose one loan as the match. This gives balanced covariates after matching. There are 168 secured loan could not be matched.

Table A.2 presents the matching covariates before and after matching. Table A.3 summarize the statistics before and after matching. Through matching, the differences between the covariates of secured and unsecured loans are completely removed, as the percentage of bias is zero for all covariates. This is also illustrated in Figure A.1 (a). Figure A.1 (b) shows the propensity score distribution of secured loan (Treated) and unsecured loan (Untreated). Treated off support indicates the unmatched secured loan.



(a)



(b)

Figure A.1: Propensity Score Matching Performance

A.2 Price Prediction Results

Contract Availability: Table A.4 shows the number of secured and unsecured loan contracts that are predicted to be available to secured and unsecured borrowers, where secured borrowers are those that chose a collateralized loan in the data, and unsecured borrowers are those who chose an unsecured loan. Our sample includes 2,871 loan contracts (842 secured and 2,029 unsecured), 561 new borrowers, and 12 banks. The maximum number of potential contracts is therefore $2,871 \times 2 \times 12 = 68,904$. For secured and unsecured contracts, the first column is the total number of contracts to be predicted (T), the second column is the number of available contracts (A), the third column is the number of unavailable contracts (U) i.e., either the bank is not active in the market or it never offered similar loan contracts, and the last column is the percentage of available contracts (% A). Secured borrowers are more likely to be offered a secured loan than unsecured borrowers (38% vs 34%), while unsecured borrowers are more likely to be offered an unsecured loan (52% vs 45%).

Table A.4: Summary of Price Prediction by Propensity Score Matching

	Secured loan				**Unsecured loan**			
	T	A	U	% A	T	A	U	% A
Secured borrowers	10,104	3,859	6,245	38	10,104	4,539	5,565	45
Unsecured borrowers	24,348	8,202	16,146	34	24,348	12,735	11,613	52
All borrowers	34,452	12,061	22,391	35	34,452	17,274	17,178	50

Note: This table summarizes the number of secured and unsecured loans that are available for borrowers. The first column is the total number of contracts to be predicted (T), the second column is the number of available contracts (A), the third column is the number of unavailable contracts (U), and the last column is the percentage of available contracts (% A).

Table A.5 shows the availability of the pair of contracts offered by banks to firms. Our matching method allows for the possibility that a bank provides only secured or only unsecured loans to each firm. It also allows banks not to offer any contract to some borrower, either because the bank is not active in the borrower's market (in 83.9% of the cases), or because the bank does not offer the type of loan required by the borrower in terms of amount and maturity (in 14.1% of the cases). Our propensity score matching allows for different contract availability between secured and unsecured borrowers, which means that banks can screen differently secured and unsecured borrowers not only with contract terms, but also with contract availability.

Table A.5: Contract Availability of Secured and Unsecured Contracts

	Both	**Only Secured**	**Only Unsecured**	**Neither**
Secured borrowers	3,087	772	1,452	4,793
Unsecured borrowers	6,849	1,353	5,886	10,260
All borrowers	9,936	2,125	7,338	15,053

Note: This table shows the availability of the two contracts offered by a bank to a borrower. Both means bank borrower offers both secured and unsecured loans. Only Secured (Only Unsecured) means a bank offers only secured (unsecured) loan to a borrower. Neither means a bank offers neither contracts to a borrower. The numbers in the table are the number of bank-firm pair that belongs to the four categories.

A.3 Fixed Effects Method

This section presents another way of price prediction. With propensity score matching, a bank may provide only secured loan or only unsecured loan to a borrower. In this section, we just use fixed effects to ensure that a borrower receive both offers from a bank at the same time.

We start with the estimation for secured borrowers and unsecured borrowers separately.

$$
\begin{aligned}
P_{ijm}^{\mathcal{S}} &= \beta_0^{\mathcal{S}} + \beta_1^{\mathcal{S}} A_i + \beta_2^{\mathcal{S}} M_i + \gamma_{jm}^{\mathcal{S}} + \lambda_i^{\mathcal{S}} + \epsilon_{ijm}^{\mathcal{S}}, \quad \text{if } i \text{ is a secured borrower} \qquad (27)\\
P_{ijm}^{\mathcal{U}} &= \beta_0^{\mathcal{U}} + \beta_1^{\mathcal{U}} A_i + \beta_2^{\mathcal{U}} M_i + \gamma_{jm}^{\mathcal{U}} + \lambda_i^{\mathcal{U}} + \epsilon_{ijm}^{\mathcal{U}}, \quad \text{if } i \text{ is an unsecured borrower} \qquad (28)
\end{aligned}
$$

For borrowers that chose secured loan and unsecured loan, running regressions (27) and (28) respectively, we obtain estimated coefficients ($\hat{\beta}^{\mathcal{S}}$, $\hat{\beta}^{\mathcal{U}}$), bank-market fixed effects ($\hat{\gamma}_{jm}^{\mathcal{S}}$, $\hat{\gamma}_{jm}^{\mathcal{U}}$), and borrower fixed effects ($\hat{\lambda}_i^{\mathcal{S}}$, $\hat{\lambda}_i^{\mathcal{U}}$). For a given firm that borrows from a given bank in the market, market fixed effect and borrower fixed effect should be very close. Therefore, if the fixed effect of borrower or market in one contract is missing, we can use the estimated fixed effect in the other contract as a proxy. For example, market fixed effect is determined as follows.

$$
\hat{\gamma}_{jm}^{\mathcal{S}} = \begin{cases} \hat{\gamma}_{jm}^{\mathcal{S}} & \text{if } \hat{\gamma}_{jm}^{\mathcal{S}} \text{ exists} \\ \hat{\gamma}_{jm}^{\mathcal{U}} & \text{if } \hat{\gamma}_{jm}^{\mathcal{S}} \text{ does not exist but } \hat{\gamma}_{jm}^{\mathcal{U}} \text{ exists} \end{cases} \qquad (29)
$$

$$
\hat{\gamma}_{jm}^{\mathcal{U}} = \begin{cases} \hat{\gamma}_{jm}^{\mathcal{U}} & \text{if } \hat{\gamma}_{jm}^{\mathcal{U}} \text{ exists} \\ \hat{\gamma}_{jm}^{\mathcal{S}} & \text{if } \hat{\gamma}_{jm}^{\mathcal{U}} \text{ does not exist but } \hat{\gamma}_{jm}^{\mathcal{S}} \text{ exists} \end{cases} \qquad (30)
$$

By doing this, we ensure that the $\hat{\gamma}_{jm}^{\mathcal{S}}$ and $\hat{\gamma}_{jm}^{\mathcal{U}}$ always exist at the same time. Similar method applies for borrower fixed effect $\hat{\lambda}_i^{\mathcal{S}}$ and $\hat{\lambda}_i^{\mathcal{U}}$.[22] In the end, we predict $P_{ijm}^{\mathcal{S}}$ and $P_{ijm}^{\mathcal{U}}$ for all i and j in the market according to (27) and (28). Notice that $P_{ijm}^{\mathcal{S}}$ and $P_{ijm}^{\mathcal{U}}$ exist if$\hat{\gamma}_{jm}^{\mathcal{S}}$ and $\hat{\gamma}_{jm}^{\mathcal{U}}$ exist. That is, the bank j is active in region m at time t, which is the same assumption as in propensity score matching method. However, no restrictions on loan characteristics are imposed. Therefore, the choice set is larger in this case.

Similar to Table A.4 and A.5, Table A.6 and A.7 show the contract availability by using fixed effect price prediction. Table A.8 compares the predicted number of contracts by the two methods to the observed sample. Compared with the Fixed Effect (FE) method, the prediction from the Propensity Score Matching (PSM) method is closer to the observed sample, with similar market shares and secured loan proportions. Table A.9 shows the predicted and observed bank market shares in percentage for each region. The observed and predicted market shares are very similar, suggesting that the loan offers predicted by our price prediction method are close to reality.

[22] Assign zero to $\hat{\lambda}_i^s$ and $\hat{\lambda}_i^u$ if the borrower fixed effects are missing

Table A.6: Summary of Price Prediction by Fixed Effect Methods

	Secured loan				**Unsecured loan**			
	T	A	U	% A	T	A	U	% A
Secured borrowers	10,104	6,048	4,056	60	10,104	6,048	4,056	60
Unsecured borrowers	24,348	15,776	8,572	65	24,348	15,776	8,572	65
All borrowers	34,452	21,824	12,628	63	34,452	21,824	12,628	63

Note: This table summarizes the number of secured and unsecured loans that are available for borrowers using fixed effect method. The first column is the total number of contracts to be predicted (T), the second column is the number of available contracts (A), the third column is the number of unavailable contracts (U), and the last column is the percentage of available contracts (% A).

Table A.7: Contract Availability of Secured and Unsecured Contracts: Fixed Effect Method

	Both	Only secured	Only unsecured	Neither
Secured borrowers	6,048	0	0	4,056
Unsecured borrowers	15,776	0	0	8,572
All borrowers	21,824	0	0	12,628

Note: This table shows the availability of the two contracts offered by a bank to a borrower using fixed effect method. Both means bank borrower offers both secured and unsecured loans. Only Secured (Only Unsecured) means a bank offers only secured (unsecured) loan to a borrower. Neither means a bank offers neither contracts to a borrower. The numbers in the table are the number of bank-firm pair that belongs to the four categories.

Table A.8: Price Prediction by Banks

Bank ID	1	3	5	7	8	9	10	11	14	16	17	18
1. Sample												
Loan	366	292	409	60	41	612	174	395	264	111	7	140
% Loan	13	10	14	2	1	21	6	14	9	4	0	5
Secured	83	81	145	55	36	117	73	82	78	31	1	60
% Secured	23	28	35	92	88	19	42	21	30	28	14	43
2. PSM												
Loan	3,799	3,153	3,729	1,077	427	4,306	2,170	3,157	3,655	1,985	12	1,865
% Loan	13	11	13	4	1	15	7	11	12	7	0	6
Secured	1,235	1,026	1,715	902	372	1,839	948	1,090	1,614	718	1	601
% Secured	33	33	46	84	87	43	44	35	44	36	8	32
3. FE												
Loan	5,180	4,772	4,692	2,678	1,534	5,090	3,398	4,332	4,536	3,728	144	3,564
% Loan	12	11	11	6	4	12	8	10	10	9	0	8
Secured	2,590	2,386	2,346	1,339	767	2,545	1,699	2,166	2,268	1,864	72	1,782
% Secured	50	50	50	50	50	50	50	50	50	50	50	50

Note: This table compares the loan offered by banks in the sample, the propensity score matching (PSM) method, and fixed effect (FE) method. For each bank the number of loan provided by that bank (Loan), the market share of that bank (% Loan), the number of secured loan provided by that bank (Secured), and the percentage of secured loan provided by that bank (% Secured) are presented.

Bank Name	Observe vs Predicted	Cochabamba	Santa Cruz	La Paz	Tarija	Chuquisaca	Oruro	Potosi	Foreign
Banco Nacional de Bolivia	O	18	10	12	10	17	32	30	
	P	15	9	12	10	14	38	21	
Banco Mercantil	O	3	10	15	17		12	5	
	P	3	10	13	19		12	6	
Banco de Crédito de Bolivia	O	13	6	14	37	62	12	65	
	P	14	7	17	34	62	12	74	
Banco de la Nación Argentina	O		3	3					
	P		3	3					
Banco do Brasil	O			4					
	P			4					
Banco Industrial	O	25	17	25	34	19	28		
	P	24	17	24	34	21	24		
Citibank	O	13	4	8					
	P	14	4	9					
Banco Santa Cruz	O	13	21	5		1	11		100
	P	14	22	4		2	10		100
Banco de la Unión	O	8	12	10	3	2	2		
	P	7	13	10	3	1	2		
Banco Económico	O	6	6	2					
	P	5	6	2					
Banco Solidario	O	0		0			3		
	P	0		0			1		
Banco Ganadero	O	2	10	2					
	P	2	10	1					

Table A.9: Observed and Predicted Bank Market Share