Bank Market Power*

Luigi Bocola† Gideon Bornstein‡ Cedomir Malgieri§ Federico Puglisi¶

July 2026

Abstract

Local loan markets are often highly concentrated, and small and medium-sized firms may have few alternatives when borrowing from banks. Do banks enjoy market power in lending, and what are its macroeconomic consequences? We propose a measurement strategy that uses credit-register data to quantify bank market power and develop a general equilibrium model to assess its aggregate effects. Our empirical approach decomposes loan interest rates into firm and bank fixed effects and interprets bank fixed effects as deviations from the law of one price. Using the Italian credit register, we find that bank fixed effects account for 15–20% of the variation in interest rates and are strongly correlated with local market shares. We then develop a general equilibrium model featuring oligopolistic banks subject to leverage constraints, firms with working capital requirements, and segmented bank–firm relationships. We show that the model matches key statistics about market concentration, bank-firm relationships, loan demand elasticities, and the dispersion of interest rates in our data. We use the model to quantify the macroeconomic effects of bank market power and to study how changes in capital requirements affect bank competition and interest rates.

*The views and opinions in this paper are those of the authors and do not necessarily reflect those of the Bank of Italy or the Eurosystem. We thank Markus Brunnermeier, Jose Ignacio Cuesta, Erik Hurst, and seminar participants at Frontiers of Macroeconomics, Stanford, the European Central Bank and the Federal Reserve Bank of Minneapolis. Alexander Gilsa Hansen provided excellent research assistance. All errors are our own. Luigi Bocola thanks the National Science Foundation for generous funding.

†Stanford University, CEPR, NBER.

‡The Wharton School, University of Pennsylvania, NBER.

§W.P. Carey School of Business, Arizona State University.

¶Bank of Italy.

1 Introduction

In many countries, local loan markets are highly concentrated, and firms have few alternatives to bank borrowing when financing their operations. These observations raise two natural questions: How much market power do banks exert in lending, and by how much does bank market power depress economic activity? These questions are central for understanding the cost of credit faced by firms, the transmission of financial shocks, and the role of policies that affect the structure of the banking sector.

In this paper, we contribute to this debate in two ways. First, we develop a methodology to quantify deviations from perfect competition in loan markets using administrative data on loan prices. We decompose the interest rates that firms pay on loans as the sum of two components, a firm and a bank fixed effect. The bank fixed effect is identified from the dispersion in interest rates across lenders for loans taken by the same borrower and, as such, captures deviations from the law of one price. We document, by applying this methodology to the European credit register, that bank fixed effects are economically large. In addition, we show that they are strongly related to standard measures of market power: banks that charge high interest rates are also those with a large share of loans in their local market. The second contribution of the paper is to extend a baseline macro-finance model, in the tradition of Gertler and Karadi (2011) and Gertler and Kiyotaki (2010), to allow for bank market power. We fit the model using estimates of loan demand elasticities and the distribution of bank-firm relationships and show that it predicts well several patterns in the observed distribution of interest rates. We then use the model to assess the economic implications of bank market power and to study how financial policies affect bank competition.

Our empirical analysis combines two complementary sources of administrative data. First, we use AnaCredit, the harmonized euro-area credit register, which provides rich loan-level information on contractual interest rates, loan and borrower characteristics but only covers a relatively short time period. Second, we use the Italian credit register and complementary sources to obtain a longer time series of bank-firm relationships, effective interest rates[1] for different types of loans, and detailed characteristics of borrowers and lenders. Taken together, these sources provide detailed information on prices and quantities, allowing us to study both cross-sectional and time-series variation in the pricing of bank credit.

Our empirical strategy proceeds in two steps. In the first step, we residualize interest rates on short-term lending facilities with respect to the bank's assessment of the borrower's credit risk, including default probabilities and recovery values for any pledged collateral.[2] This allows us to

[1] Unlike in AnaCredit, contractual interest rates are not available in the Italian Credit Register. Effective interest rates are computed using information on interest payments, actual borrowed amounts, and the time span over which credit is outstanding.

[2] Our benchmark analysis focuses on overdraft accounts because they are comparatively homogeneous contracts: they are callable and do not have an explicit maturity. In the paper we show that our results are quantitatively similar if we were to use other categories of short-term lending.

isolate variation in credit prices that is not directly accounted for by risk or contract composition. Consistent with recent studies that use credit register or supervisory loan data for other countries, e.g., Cavalcanti et al. (2023), Faria-e-Castro et al. (2025a), and Ferreira and Kozeniauskas (2025), we find substantial dispersion in interest rates even after conditioning on these indicators.[3]

In the second step, we regress the residualized interest rates on a firm and a bank fixed effect. The firm component captures systematic differences in borrowing rates across firms, while the bank component captures systematic differences in the interest rates charged by banks on the same firm.[4] The identification of the bank fixed effect comes from firms that borrow from multiple lenders at the same time: once these overlapping lending relationships generate a connected network, we can recover bank fixed effects relative to one another from within-firm comparisons of borrowing rates. In our setting, the bank fixed effects can be estimated from a simple cross-section, since we observe many firms borrowing from multiple lenders at the same time.

We begin by applying this framework to AnaCredit, focusing on overdraft accounts of Italian firms outstanding in the third quarter of 2019. We document that the bank fixed effect accounts for 17% of the dispersion in residualized interest rates—equivalent to roughly half of the share of variance in interest rates accounted for by our indicators of default risk and recovery values. These effects remain economically important even when controlling for factors that might otherwise account for heterogeneity in loan pricing. For example, a possible explanation for the dispersion in interest rates is that overdraft rates depend on other characteristics of the contract such as the committed amount, whether the interest is fixed or variable, the date of inception, and on characteristics of the bank-firm match, such as the length of the relationship. We show that all these factors, while important for driving the cross-sectional dispersion in interest rates, have little impact on the share of variance accounted for by the bank fixed effect.

We then turn to the longer Italian credit register sample to study how these forces evolved over time. We find that the importance of bank fixed effects declined during the 2000s and then rose again after the financial crises of the 2010s, a pattern consistent with a period of increasing competitive pressure followed by post-crisis consolidation in banking.[5] In addition, we show that across provinces banks with larger market shares tend to charge higher interest rates, even after controlling for bank balance-sheet characteristics and time effects. This evidence supports the

[3]In our residualization step, we find that default probability and losses upon default explain 26% of the cross-sectional dispersion in interest rates in overdraft accounts. This is somewhat higher than that found in other studies, which typically pool all types of loans. Using Brazilian data, Cavalcanti et al. (2023) show that observable measures of risk explain about 24% of the variation in loan spreads. In Faria-e-Castro et al. (2025a), default probability and loss-given-default explain (together with other loans' characteristics and sector-time fixed effects) only 19% of the dispersion.

[4]Our definition of a bank is the interaction between a banking group (e.g. *UniCredit*) and a province (e.g. Rome), thus allowing branches of the same group in different locations to have different fixed effects.

[5]See for example De Bonis et al. (2014) for an analysis up to 2014. The European Central Bank has instead documented that euro-area banking markets became more concentrated after the global financial crisis, with about one-third of banking groups disappearing (European Central Bank, 2021).

interpretation of bank fixed effects as capturing economically meaningful differences in market power.

While this evidence points to a significant deviation from the law of one price, it is silent about how large the underlying rents are and about its macroeconomic implications. To make progress on this dimension, we develop a dynamic model that can generate cross-sectional dispersion in interest rates and use it to measure the macroeconomic consequences of market power in banking. Our model builds on the macro-banking framework of Gertler and Kiyotaki (2010) and Gertler and Karadi (2011): banks issue deposits to households, subject to a leverage constraint, and invest these funds in firms that need to borrow to finance working capital. The key innovation relative to the existing literature is to assume that financial services provided by banks are imperfect substitutes, and that firms are randomly matched to a subset of banks. Conditional on these matches, banks interact strategically under oligopolistic competition when setting interest rates.

In equilibrium, the interest rate charged by a bank to a firm is a markup over the bank's marginal cost of funds. As in standard variable-markup models, these markups are heterogeneous across banks and firms. When a firm borrows from many atomistic lenders, the elasticity that each lender faces is related to how substitutable loan services are. When a firm is instead matched with just one lender, the elasticity is related to the speed at which the firm reduces production when facing higher interest rates. To the extent that the former is larger than the latter, our model predicts that loan markups will be larger for banks with large market shares. This also implies that firms matched with only a few banks will typically face higher markups. The marginal cost, in turn, reflects both the rate at which banks borrow and the scarcity of deposits relative to the loans they want to extend. While the former is common across banks, the latter may be heterogeneous because banks' leverage constraints bind differently across institutions.

We discipline the model parameters via indirect inference, matching, among other targets, the average number of banks from which firms borrow, the distribution of bank market shares, and reduced-form estimates of the elasticity of loan demand with respect to interest rates. Specifically, we apply the methodology of Amiti and Weinstein (2018) to the Italian credit register to isolate idiosyncratic changes in a bank's lending rate that are unrelated to loan-demand factors, and we study how loans adjust to such shocks. We estimate a semi-elasticity of -19 in levels and -12 when loans are measured as a fraction of total firm borrowing.[6] These values, together with the distribution of bank-firm relationships and the size distribution of banks, discipline the average markup in our model and how it varies across banks and firms.

We test the model's predictions by asking whether it can reproduce key features of the observed

[6] These estimates are consistent with recent work. Using U.S. supervisory loan data, Dempsey and Faria-e-Castro (2025) estimate an overall semi-elasticity of loan demand close to -20. Using euro-area credit register data, Altavilla et al. (2022) estimate an elasticity of substitution across banks in the range of 7 to 23. Our estimate of -19 for loan levels is therefore close to the former, while our estimate of -12 for loan shares lies comfortably within the range reported by the latter.

interest-rate distribution. We show that interest rates in our economy admit a log-linear representation in a firm and a bank fixed effect, like the one we estimate in the data. Firms that are matched with few banks end up facing high interest rates in equilibrium, as the banks they borrow from exploit their market power and charge high markups. The bank fixed effect, in turn, is driven by the sum of the markup and of the marginal cost of that specific bank relative to the average across banks. We show that the estimated model is consistent not only with the observed cross-sectional dispersion in interest rates, but also with the decomposition we estimate in the Italian data: about 20% of cross-sectional dispersion is explained by bank fixed effects. Additionally, the bank fixed effects are strongly correlated with bank market shares and the number of competitors faced by each bank.

After validating the model, we turn to counterfactual analysis. First, we ask what would happen if banks were forced to price their loans at marginal cost. This exercise is useful from a measurement perspective because it allows us to translate the deviations from the law of one price observed in the credit register into an output metric. When banks price at marginal cost, there are two key forces operating in the short run: a reduction in the average interest rate paid by firms and a compression of its dispersion. The first effect increases output through the standard working-capital channel, with lower interest rates increasing labor demand and economic activity. The second channel increases output by reducing misallocation, as equally productive firms in our economy pay different interest rates. We show that the overall gains from the elimination of markups are sizable, with output increasing by 1.3% in the short run. Most of these effects are due to the reduction in the average markup rather than due to the misallocation channel.

In the long run, however, part of these gains is eroded. Because banks price at marginal cost, they accumulate less capital over time. As capital helps banks overcome their leverage constraints, the economy with zero markups converges to a steady state in which financial constraints on banks are tighter. Thus, while eliminating bank market power reduces the distortions coming from loan markups, over time it increases the marginal cost of borrowing for banks. This force pushes output downward. Eventually, this mechanism erodes most of the gains from eliminating bank market power, with output increasing by only about 0.5% in the long run. Accounting for credit risk and systemic bank risk may further dampen the long-run output gain.[7]

In a second counterfactual, we use the model to study the interaction between bank capital requirements and bank competition. Changes to bank capital requirements affect banks asymmetrically depending on the tightness of leverage constraints. Because less-capitalized banks are more likely to be constrained in our framework, a uniform relaxation of bank capital requirements has larger effects on them, reshaping market shares and, in turn, the distribution of loan markups. In

[7] Credit is risk-free in the model, as the goal of this paper is to understand the share of interest-rate dispersion that is not explained by credit risk. Therefore, the increase in output implied by making banks price loans at marginal cost should be interpreted as a counterfactual exercise that holds fixed all forces outside the model, including the riskiness of credit.

our quantitative exercise, about half of the decline in interest rates following a relaxation of capital requirements comes from this endogenous change in markups rather than from a direct reduction in marginal costs. The endogenous change in markups speaks to the potential anti-competitive consequences of capital requirements policies that are often discussed by policymakers, and shows how the structure of credit markets can shape the transmission of regulatory interventions. Our analysis should be interpreted as a quantitative evaluation of how market structure shapes some of the costs of capital requirements. Importantly, we focus on one margin: how these policies affect the cost of lending. The analysis abstracts from considering risk in the relationship between lending activity, interest rates and bank capital and is therefore not meant to provide a comprehensive assessment of capital requirements, since the model does not capture their potential benefits in reducing systemic bank risk and making banks more resilient.

Related literature. Several recent papers use credit register data to document salient features of the cross-sectional distribution of loan rates paid by firms. This includes Cavalcanti et al. (2023) for Brazil, Altavilla et al. (2024) for the euro area, Faria-e-Castro et al. (2025b) and Amiti et al. (2026) for the United States and Ferreira and Kozeniauskas (2025) for Portugal. A common finding in this literature is that there is sizable dispersion in borrowing rates across firms even after controlling for observable indicators of credit risk. Among these papers, the one closest to ours empirically is Ferreira and Kozeniauskas (2025), who also estimate "AKM-style" regressions for observed loan rates with firm and bank fixed effects identified from firms borrowing from multiple lenders. Consistent with our analysis, they find that bank-side heterogeneity accounts for 10-15% of the observed dispersion in interest rates.[8]

Our contribution to this literature is threefold. First, we explicitly connect dispersion in bank fixed effects to deviations from the law of one price in loan markets, and show that this statistic is informative about the degree of bank market power.[9] Second, we show that a model with sticky firm-bank relationships and variable bank markups can account well for the cross-sectional dispersion in interest rates observed in the data. Third, we combine this measurement with our model to quantify the macroeconomic costs of monopolistic rents in loan markets.

A second strand of the literature studies the macroeconomic effects of bank dynamics through the lens of dynamic general equilibrium models, building on the foundational frameworks of Gertler and Kiyotaki (2010) and Gertler and Karadi (2011). More recently, the high concentration of the banking industry has led researchers to incorporate bank market power into these frameworks. Jamilov and Monacelli (2025) develop a model of heterogeneous financial intermediaries with market power in deposit markets, while Ulate (2021) and Wang et al. (2022) study how bank market

[8]This figure is the closest one to our approach, in which the bank fixed effect is identified exclusively from the dispersion in lending rates applied to the same borrower at the same time.

[9]Our analysis corroborates this result by documenting that the bank fixed effects correlate with concentration measures in loan markets.

power shapes the transmission of monetary policy. Dempsey (2025) studies how heterogeneous banks compete both with one another and with the bond market, and Aguirregabiria et al. (2025), Maingi (2026), and Bordeu et al. (2026) analyze how bank market structure shapes the spatial allocation of credit. Most closely related to our paper, Corbae and d'Erasmo (2021), Villa (2025), Morelli and Herkenhoff (2024) develop quantitative models of banking industry dynamics with oligopolistic competition to study the welfare cost of market power, business cycles and the effects of macroprudential policies. We contribute to this literature by developing a model of oligopolistic competition among non-atomistic banks and showing that, given plausible values of loan-demand elasticities and concentration measures, it matches key features of the cross-sectional dispersion in interest rates observed in the credit register.

An influential view considers monopolistic rents important for supporting long-term relationships between banks and firms, which in turn can foster lending to small and young firms whose prospects are uncertain and hard to monitor. In Petersen and Rajan (1995) banks are more willing to finance young and credit-constrained firms in concentrated markets because they better can internalize the future rents generated by a lending relationship. There is substantial evidence documenting the importance of relationship lending in credit markets, see for example the empirical analysis of Crawford et al. (2018) or the recent quantitative analysis in Beraldi (2025). It is important to point out that our counterfactuals only assess the efficiency gains coming from eliminating bank markups—not the indirect effects that weaker bank-firm relationships would have on bank lending. In this respect, our approach closely mirrors that of Hsieh and Klenow (2009), Baqaee and Farhi (2020), and Edmond et al. (2023), who assess the efficiency losses due to monopolistic rents in product markets without considering the effects that eliminating those rents would have on other margins, e.g., innovation activity.[10]

Finally, our paper relates to a growing literature on market power across different markets and its welfare consequences. In an influential paper, De Loecker et al. (2020) document a substantial increase in markups in the United States since 1980, while Edmond et al. (2023) and De Loecker et al. (2021) study the welfare costs of markup distortions. In labor markets, Yeh et al. (2022) measure markdowns and Berger et al. (2022) quantify the welfare cost of labor market power. In banking, Drechsler et al. (2017) and Begenau and Stafford (2018) study market power in deposit markets, while Benetton et al. (2025) analyze market power in mortgage lending. Our paper instead focuses on loan markets, where we quantify rents associated with lender market power and assess their macroeconomic costs.

Layout. The remainder of the paper is organized as follows. Section 2 presents the data and the empirical evidence. Section 3 describes the general equilibrium model. Section 4 presents the

[10]Our work is therefore related to a literature measuring the welfare cost of markdowns and labor market power (Berger et al., 2022).

quantitative analysis, including the estimation of the model and the main counterfactual exercises. Section 5 concludes.

2 Empirical Evidence

This section presents the main empirical findings of the paper. Section 2.1 describes data sources, definitions and sample selection. Section 2.2 describes the empirical model that we use to decompose interest rates into a firm fixed effect and a bank fixed effect, and presents the results of this decomposition. Section 2.3 analyzes further the bank fixed effects, showing that they correlate with measures of market power.

2.1 Data

We perform our analysis using two main datasets. Our baseline analysis uses AnaCredit, the harmonized credit register for the euro area. This register provides detailed information on all loans issued by credit institutions in the euro area above €25,000, starting from 2018-M9.[11] We then complement this analysis with information from the Italian Credit Register. While this register is less detailed than AnaCredit (it provides information at the firm rather than loan level), it gives us a longer time-series and provides information on small banks.

AnaCredit covers all loans above €25,000 euros extended by credit institutions in the euro area and records, at the loan level, the contractually agreed interest rate and interest rate type, the outstanding and committed amounts, and key loan characteristics including loan type, loan purpose, date of inception, maturity, and allocated collateral value. AnaCredit also reports a bank assessment of the firm probability of default, defined as the one-year probability that a debtor will fail to meet its obligations. This figure is primarily aligned with the Internal Ratings-Based (IRB) approach for credit risk, as established in the Capital Requirements Regulation (CRR) and is only reported when the filing institution has opted for an IRB approach on the specific loan issued. AnaCredit also directly reports firm-level information on total revenues, assets and number of employees. One limitation of AnaCredit is that it does not include some small banks with a national market share below 2%, as these credit institutions do not have to report their loans to AnaCredit.

The second dataset, covering the 2005–2019 period, is built by matching several administrative sources. The core comes from the Italian Credit Register and the TAXIA archive on interest

[11] Not all banks participate in the program. National Central Banks can let financial institutions (temporarily) opt out from reporting information if the total outstanding amount of loans is less than (4%) 2% of the national figure.

rates, supplemented with balance sheet data on borrowers from Cerved and on lenders from the Supervisory Files (COREP) on banks and banking groups.

The Italian Credit Register, managed by the Bank of Italy, is a comprehensive database recording the debts of households and firms toward the banking and financial system. It collects information submitted by virtually all banks and credit intermediaries in Italy. It contains detailed quarterly information on credit issued by banks and other credit intermediaries above a minimum threshold of €30,000, covering both disbursed and undisbursed credit.[12] Information is reported at the firm-bank level, and all loans that a firm has taken from a specific bank are aggregated into three categories: *rischi a revoca* (revolving loans), *rischi autoliquidanti* (loans backed by account receivables), and *rischi a termine* (fixed-term loans). Bank of Italy provides intermediaries with information on their customers' total indebtedness, loan types, and repayment behavior. Intermediaries can also query the register about prospective borrowers to assess creditworthiness.

TAXIA complements the Credit Register with information on interest rates for borrowers with at least €75,000 in overall granted or disbursed credit by institution. It is reported by all but the smallest banks, and its coverage is highly representative: reporting banks account for roughly 80 percent of outstanding credit. For each of the three loan categories in the credit register, TAXIA records the actual interest amount paid by debtors in each quarter and the computational principal, as computed by the bank, on which the effective interest rate is applied to compute the actual interest paid by debtors to a given bank. Effective interest rates are hence computed as the ratio between interest paid and the relevant computational principal for each quarter. The interest paid is recorded net of commissions and fees.

CEBI (Centrale dei Bilanci) is a proprietary database compiled by an external risk-analytics firm, CERVED, providing detailed firms' balance sheet and income statement information at the yearly frequency and expected default probability at 1, 3 and 5 years horizons. Its coverage is substantial: total credit to Cerved firms represents approximately three-quarters of loans by Italian banks to the nonfinancial corporate sector. Balance sheet information on lenders comes from the Supervisory Files on banks and banking groups, containing virtually all information present on Balance Sheet and Income Statements as well as Prudential reporting documents (COREP).

All data sources are matched based on the year-quarter of observation, the credit institution unique charter code (ABI), and the firm unique VAT code. For our baseline analysis, we will focus on overdraft accounts[13] from AnaCredit. As AnaCredit is only available since 2019, we then complement our analysis with overdraft accounts sourced from the Italian Credit Register matched with TAXIA. Overdraft accounts are the vast majority of revolving credit for Italian firms, and they are conventionally superior for analyzing bank loan rate setting policies compared to other

[12] When a borrower is classified as a bad debtor, based on an assessment of their overall financial situation rather than isolated late payments, the reporting threshold drops to €250.

[13] Overdraft accounts are defined as revolving credit facilities that allow account holders to withdraw more money than is available in their current or cash management accounts, up to a pre-approved limit.

instruments because they have fairly standardized basic features, are short-term, and often have variable rates or frequently updated fixed rates, unlike long-term fixed loans. These features are especially important when using the Italian credit register because, unlike AnaCredit, all term loans issued are reported in aggregated form at the firm level without information on maturity, date of inception, or reference rate used at the time of inception. As a result, dispersion in interest rates for term loans would be confounded by variation in loan characteristics.

Loan categories. AnaCredit is broader than the current-account credit used in our benchmark analysis: within its reporting perimeter, it records loan-level information for the reportable corporate credit instruments of credit institutions, including not only overdrafts but also other credit instruments. We use this broader coverage only to construct comparable short-term credit categories and to run robustness checks across them. The benchmark AKM decomposition remains focused on the category we label *liquidity credit*, namely overdrafts on current accounts reported with the residual AnaCredit purpose "other purposes." The remaining categories below are kept separate from the benchmark sample: they are used to describe the composition of short-term credit and to verify that the main results are not driven by the overdraft sample alone. This separation into homogeneous categories is important because different loan types differ in maturity, collateral structure, risk, and pricing, so pooling all instruments would mix economically distinct forms of credit.[14]

Overall we therefore built five loan categories from the intersection of the AnaCredit instrument-type and purpose attributes, together with a maturity restriction for non-overdraft instruments. The first two categories are *overdrafts for other purposes* and *overdrafts for working capital purposes*. In the AnaCredit framework, overdrafts are debit balances on current accounts: funds provided to debtors through a debit balance on an account that is primarily designed to allow credit balances, although ordinary debit balances may also arise. An overdraft is therefore distinct from other revolving credit because it must be associated with a current account. Overdrafts may arise either on current accounts with an agreed credit limit, where the limit is attached directly to the current account, or on current accounts without a contractually agreed limit, for example in the case of unauthorised debits. Within overdrafts, we distinguish the two categories using the AnaCredit purpose attribute. *Overdrafts for working capital purposes* are overdrafts reported with purpose "working capital facility," which refers to financing the cash-flow management of the borrower, excluding pure import or export financing. *Overdrafts for other purposes* are overdrafts reported with purpose "other purposes," the residual purpose code used when the instrument is not assigned to one of the named AnaCredit purpose categories; the AnaCredit manual also specifies this purpose for debit balances on current accounts without an agreed credit limit.

The third category is *trade receivables*. We keep this category separate because AnaCredit treats

[14]See, for instance, Schivardi et al. (2022) and Ivashina et al. (2022).

trade receivables as a distinct instrument type rather than as overdrafts or generic loans. This category covers loans granted on the basis of bills or other documents that give the creditor the right to receive proceeds from transactions involving the sale of goods or the provision of services. It includes factoring and similar transactions, such as acceptances, outright purchases of trade receivables, forfaiting, and the discounting of invoices, bills of exchange, commercial papers, and other claims, provided that the credit institution purchases the receivables, with or without recourse to the debtor. By contrast, financing in which receivables are merely pledged as protection, while the borrower retains the receivables and books the advance as a loan, is not reported in AnaCredit as trade receivables.

The fourth and fifth categories are *short-term other loans for other purposes* and *short-term other loans for working capital purposes*. Both are restricted to instruments with AnaCredit instrument type "other loans" and contractual maturity below one year. In AnaCredit, "other loans" is a residual, non-revolving instrument type: it covers loans that are not classified as deposits, overdrafts, credit card debt, revolving credit, credit lines, reverse repurchase agreements, trade receivables, or financial leases. In particular, instruments entirely disbursed in one instalment and not falling into the other instrument types are classified as "other loans." We impose the maturity restriction to keep these categories comparable to the short-term liquidity instruments used in the benchmark analysis. The two categories differ only by purpose: *short-term other loans for other purposes* have the residual purpose "other purposes," while *short-term other loans for working capital purposes* have purpose "working capital facility." Thus, the fourth category is residual both in instrument type and in purpose, whereas the fifth identifies short-term, non-revolving loans used for the borrower's cash-flow and working-capital management.

Sample selection. Our baseline sample consists of loans issued by banks and to firms headquartered in Italy. We exclude micro firms (fewer than five employees), firms with missing information on key variables, and outliers. Specifically, we drop observations in which loan utilization exceeds total firm assets or in which operating income is very negative. We winsorize interest rates at the top and bottom 2.5 percent and loan utilization at 1 percent to limit the influence of extreme observations.

Summary statistics. Table 1 reports summary statistics for the main sample constructed using AnaCredit data for 2019:Q3. The sample contains 119,597 firm observations. The average firm in the sample employs 58 workers and holds assets of about 31 million euros. The median firm is substantially smaller, with 15 employees and assets of 2.3 million euros, reflecting the prevalence of small and medium-sized enterprises in the Italian economy.

A key feature of the sample is that most firms borrow from a small number of banks. The

Table 1: Summary Statistics

	Obs.	**Mean**	**SD**	**p5**	**p25**	**p50**	**p75**	**p95**
Panel A: Firm Characteristics								
Employment	119,597	58.33	2,015.89	6	9	15	28	123
Total assets (€000)	119,597	31,428	750,000	307	957	2,314	6,522	48,619
Revenue (€000)	119,597	26,120	570,000	412	1,079	2,466	6,852	48,757
Total credit (€000)	119,597	2,658	20,799	39	145	404	1,295	8,431
No. bank relationships	119,597	2.09	1.50	1	1	2	3	5
No. loans with same bank	119,597	3.41	15.67	1	1	2	3	10.00
Panel B: Short-Term Credit (Firm Level)								
Total (€000)	90,478	1,349	7,260	20	98	260	799	4,832
Liquidity credit	90,478	196	1,802	0	0	15	100	620
Working-capital credit	90,478	104	839	0	0	0	0	400
Receivable-based	90,478	442	3,116	0	0	50	285	1,645
Term loans	90,478	363	3,721	0	0	0	25	1,200
Trade financing	90,478	126	1,261	0	0	0	0	352
Other	90,478	118	1,786	0	0	0	0	400
Panel C: Liquidity Credit (Loan Level)								
Interest rate (%)	109,387	6.14	4.78	0.55	2.75	5.00	8.15	14.62
Prob. of default	109,387	0.04	0.12	0.00	0.01	0.01	0.03	0.12
No. bank relationships	109,387	1.98	1.13	1.00	1.00	2.00	3.00	4.00
Age of relationship	109,387	2.96	1.42	1.00	2.00	3.00	4.00	5.00
Maturity	10,145	3.33	5.17	0.81	1.70	3.00	4.24	5.50
Collateral value (%)	109,387	53.65	48.37	0	0	87.78	100.00	100.00
Credit limit (€000)	109,387	249.19	1,500.22	0.13	23.98	54.75	180.00	800.00
Amount outstanding (€000)	109,387	121.55	645.35	0.05	6.41	28.03	85.95	412.20

Notes: The sample consists of loans outstanding in 2019:Q3 for firms headquartered in Italy excluding micro firms (fewer than five employees), issued by IRB Banks (those banks using the Internal Ratings-Based (IRB) approach are the only ones reporting the Probability of default in Anacredit). Interest rates and loan utilization are winsorized at ±2.5% and ±1%, respectively. Monetary values are in euros.

average number of banking relationships per firm is 2.1.[15] Since larger firms tend to have more bank relationships, the average number of banking relationships per firm is higher if firms are weighted by their size or their total borrowing. This is important because, as we will discuss later, the bank fixed effect is identified from firms that borrow from more than one bank.

Panel B reports summary statistics for firms' short-term credit by loan categories. We focus on loans with maturity of less than one year, so that 76% of firms in our sample have at least one of these short-term loans with maturity of less than one year. The first row reports total short-term credit at the firm level, while the following rows break this amount down into its main components: liquidity credit, working-capital credit, receivable-based credit, term loans, and trade financing.

Panel C reports summary statistics for the loan-level sample of loans that falls in the category of liquidity credit, which will be the main unit of observation in our analysis. For each loan, we report the interest rate, the borrower's probability of default, the number of bank relationships

[15]See Kosekova et al. (2025) for an analysis of firm-bank relationships across European countries.

for liquidity credit loans, the age of the bank-firm relationship, loan maturity, collateral value as a share of the credit commitment, the credit limit, and the amount outstanding. These variables summarize both the pricing and contractual terms of liquidity credit, as well as key dimensions of borrower risk and bank-firm relationships. The average interest rate is 6%, with substantial dispersion: the standard deviation is 4.78 percentage points, the 5th percentile is 0.55% and the 95th percentile is 14.62%. This dispersion is the empirical phenomenon we seek to explain and decompose.

2.2 AKM decomposition

This section studies the driving forces of interest rate dispersion across bank–firm pairs. We proceed in two steps. We begin by quantifying how much of the observed cross-sectional dispersion in loan rates can be explained by the most natural competitive determinant of pricing, namely default risk and recovery values in the event of default. We then turn to a statistical decomposition of the residual variation into firm and bank components. This decomposition allows us to quantify the dispersion in interest rates across banks lending simultaneously to the *same* firm, and as such it will be informative about deviations from the law of one price in loan markets.

We think of the law of one price as reflecting a combination of two conditions: loan services offered by different banks are perfect substitutes, and lending markets are perfectly competitive. The fact that a given firm borrows from two different banks at the same point in time, and at two different interest rates, suggests that the loan services offered by these banks are not perfect substitutes. Moreover, interest rate dispersion at the bank-province level implies that the same bank may charge different interest rates across local loan markets, that is, across provinces. As we show in Section 2.3, the estimated bank-province fixed effects are correlated with measures of bank market power.

Default risk. The most natural explanation for interest rate dispersion is default risk. Even under competitive loan pricing, interest rates would vary across loans because they reflect differences in the probability of default and in the recovery value conditional on default. Our first step therefore quantifies how much of the cross-sectional dispersion in loan rates can be accounted for by these two forces. To do so, we estimate:

$$r_{lib,t} = \alpha + \delta'_{pd}\, pd_{ib,t} + \delta'_{col}\, collateral_{lib,t} + \hat{r}_{lib,t}, \tag{1}$$

where $r_{lib,t}$ denotes the interest rate on loan l of firm i with bank b at time t, $pd_{ib,t}$ and $collateral_{lib,t}$ are fixed effects based on the deciles of the distribution of reported default probabilities and collateral values. Importantly, the default probability is observed at the firm–bank level, since it is reported by each bank for each borrowing firm. We also use information on the reported value of pledged collateral, which affects recovery values in the event of default.

We estimate equation (1) in the cross section of loans and find that default risk and collateral jointly explain 26 percent of the cross-sectional variation in interest rates. Thus, while these variables have substantial predictive power, the majority of the observed dispersion remains unexplained. One potential concern is that bank-specific measures of default probability may introduce bias into the estimated role of default risk. To address this issue, we replicate the analysis using external credit ratings from CERVED and find that these ratings have even less predictive power than the bank-reported probability of default. This finding is in line with the evidence in Cavalcanti et al. (2023), Ferreira and Kozeniauskas (2025), Amiti et al. (2026), and Faria-e-Castro et al. (2025b), who document for other countries that a large fraction of interest rate dispersion cannot be explained by default risk.

AKM decomposition. Having established that a substantial share of interest rate dispersion is not explained by default risk and collateral, we next decompose the remaining variation into firm and bank components. Specifically, we estimate a two-way fixed effects regression in the spirit of Abowd et al. (1999):

$$\hat{r}_{lib,t} = \alpha_{b,t} + \alpha_{i,t} + \varepsilon_{lib,t}, \tag{2}$$

where $\alpha_{b,t}$ is the bank fixed effect and $\alpha_{i,t}$ is the firm fixed effect. We define a bank as a bank–province pair, where the province is the province in which the firm is headquartered.[16] By doing so, we are allowing the same banks to systematically charge different interest rates in different markets.

Intuitively, the firm fixed effect is identified from the average interest rate paid by a given firm, while the bank fixed effect is identified from differences in the rates charged by different banks to the same firm. For this purpose, we restrict the estimation of equation (2) to firms that have at least two simultaneous lending relationships with different banks. In contrast to the standard worker–firm setting, where identification relies on workers switching employers over time, here we observe within-period borrowing relationships between the same firm and multiple banks, which provides direct cross-sectional variation for identifying bank fixed effects.

Because these fixed effects are identified only up to a constant, estimation is typically restricted to connected sets. A connected set is a subset of firms and banks that are linked through observed borrowing relationships, so that fixed effects can be compared across all units in the set. In our application, different provinces form disconnected sets because we allow for bank–province fixed effects and assign province information based on the province in which firms are headquartered. We therefore take the largest connected set within each province and normalize the fixed effects

[16] Provinces are administrative units roughly comparable to a U.S. county. According to the Italian Antitrust Authority, the relevant market in banking for antitrust purposes is the province (Guiso et al., 2012). Italy is divided into 110 provinces.

within province so that either the average firm fixed effect or the average bank fixed effect is equal to zero. We discuss below how results may be sensitive to this normalization.

The main object of interest is the relative importance of firm and bank fixed effects in explaining residual interest rate dispersion. In particular, we focus on:

$$\frac{\text{Var}(\alpha_{b,t})}{\text{Var}(\hat{r}_{lib,t})}, \tag{3}$$

that is, the share of the variance of residual interest rates accounted for by bank fixed effects. Estimating this statistic from the fitted fixed effects may, however, lead to biased estimates because of limited-mobility bias (Andrews et al., 2008; Bonhomme et al., 2023; Kline, 2024). The intuition is that, when bank fixed effects are imprecisely estimated, they may absorb idiosyncratic noise, artificially inflating the estimated variance of bank effects. This concern is likely less severe in our setting because, at a given point in time, we observe the same firm borrowing simultaneously from multiple banks. In labor market applications, by contrast, one would need to observe the same worker employed by multiple firms at the same time for an analogous source of within-period variation, which is infeasible; as a result, the standard AKM estimator instead relies on workers switching jobs over time.

Nonetheless, to address any residual concern related to limited-mobility bias, we report results for a high-mobility subsample. In practice, we restrict attention to bank–province pairs for which we observe at least 15 firms that also borrow from at least one other bank, that is, at least 15 movers.[17]

Results. Table 2 reports the variance decomposition when estimating (2) using interest rates on liquidity credit (i.e., overdraft accounts). Column (1) reports the benchmark decomposition. Firm fixed effects account for 62 percent of residual interest rate dispersion, while bank fixed effects account for 17 percent. The remaining variation is captured by the residual component, and the covariance between firm and bank fixed effects is small and negative.

Column (2) asks to what extent the dispersion in bank–province fixed effects simply reflects geographical differences in lending rates across provinces. To address this question, we absorb province fixed effects and recompute the decomposition. Since each province is a connected set, this is equivalent to imposing that the bank-province fixed effects are on average zero in every province.[18] The variance of bank fixed effects remains sizable and equal to 10 percent of total residual interest rate dispersion. Thus, while geography is an important component of the interest rate dispersion that we document, a large share of the bank component does not reflect pure

[17] Bonhomme et al. (2019) show that the standard AKM estimator may create problems when the true model is not log-additive in the two fixed effects. Our model implies a log-additive decomposition of interest rates.

[18] In the baseline specification we normalized the firm fixed effect to have mean zero within each province, while the mean bank-province fixed effect for each province was equal to the average interest rate in that province.

Table 2: AKM Variance Decomposition for liquidity credit

	Benchmark	Province FE	Loan Controls	Relationship + Loan
Standard deviation	2.87	2.87	2.87	2.87
Variance share: firm FE	0.62	0.62	0.52	0.52
Variance share: bank FE	**0.17**	**0.10**	**0.16**	**0.16**
Variance share: residual	0.24	0.24	0.22	0.22
Variance share: COV firm–bank FE	−0.02	−0.02	−0.03	−0.03
N	48,228	48,228	47,088	47,088

Notes: The table reports the variance decomposition of residual interest rates $\hat{r}_{lib,t}$ estimated from equation (2). All variance components are expressed as fractions of total residual variance $\mathrm{Var}(\hat{r}_{lib,t})$. Column (1) reports the benchmark specification. Column (2) absorbs province fixed effects to isolate the component of bank–province fixed effects that is orthogonal to pure geographical variation in lending rates. Column (3) adds loan-level controls for the type of rate, whether the rate is fixed or variable, the reference rate, the purpose of the loan, the committed amount, and the amount utilized, using fixed effects for bins of continuous variables where appropriate. Column (4) further adds controls for the length of the bank–firm relationship and for the number of credit instruments within the same firm–bank match.

geographical variation but rather differences across banks within the same province or differences across provinces within the same banks.

In Column (3) we further control for observable loan characteristics. In particular, we add a categorical variable to control for whether the overdraft rate is fixed or variable, the reference rate of the contract, and binned fixed effects constructed using the distribution of the committed amount and the amount utilized. These controls have strong predictive power for loan pricing: the variance explained by the firm fixed effect drops from 62% to 52%. However, their introduction does not significantly alter the share of variance explained by the fixed effect of the bank. This finding suggests that the bank component does not primarily reflect specialization in particular loan contracts within the broad loan category we study.

Finally, Column (4) addresses two additional competing explanations for interest rate dispersion. First, the price charged on a given loan may depend on the length of the underlying bank–firm relationship, as emphasized by Gobbi and Sette (2014), Crawford et al. (2018), and Beraldi (2025). Second, loan pricing may reflect cross-subsidization across different contracts held by the same firm with the same bank, for example if overdrafts, working capital loans, and long-term loans are jointly priced. To account for these possibilities, we add controls for the length of the relationship and for the number of credit instruments within the same firm–bank match. These additional controls leave the results essentially unchanged.[19]

While we focus on liquidity credit for the baseline decomposition, we extend the analysis to the other loan categories. Table 3 reports the composition of the loan sample across the main credit categories and the corresponding share of variance explained by bank fixed effects estimated using the fully saturated AKM specification (i.e., Column (4) in Table 2) . Receivable-based credit

[19] Beraldi (2025) documents robust evidence that accounting for relationship lending is essential to explain how interest rates vary over time within a given firm-bank pair. Our evidence suggests that relationship lending has little explanatory power in the cross-section.

Table 3: Loan Categories and Bank Fixed Effects

	Outstanding amount (% of total)	Committed amount (% of total)	Variance share: bank FE
Liquidity credit	11.24	14.48	0.16
Working-capital credit	5.90	7.69	0.18
Receivable-based	35.96	32.72	0.17
Term loans	23.88	22.38	0.12
Trade financing	10.35	9.34	0.11

Notes: The table reports, for each loan category, its share of total outstanding amounts, its share of total committed amounts, and the estimated bank fixed effect from the fully saturated AKM specification.

accounts for the largest share of both outstanding and committed amounts, followed by term loans and liquidity credit. Working-capital credit and trade financing represent smaller shares of the sample. The last column shows that the estimated share of variance explained by bank fixed effects varies across loan categories, with the highest values for working-capital credit and receivable-based credit, and lower values for term loans and trade financing. Overall, we find that results are relatively stable across different loan categories.

Time variation. We now turn to the dataset built by merging the Italian credit register with TAXIA to study how the dispersion in bank fixed effects has evolved over time. To make the analysis comparable with the previous findings, we estimate equation (2) on the effective[20] interest rates for "rischi a revoca", a loan category that comprises mostly overdraft accounts.

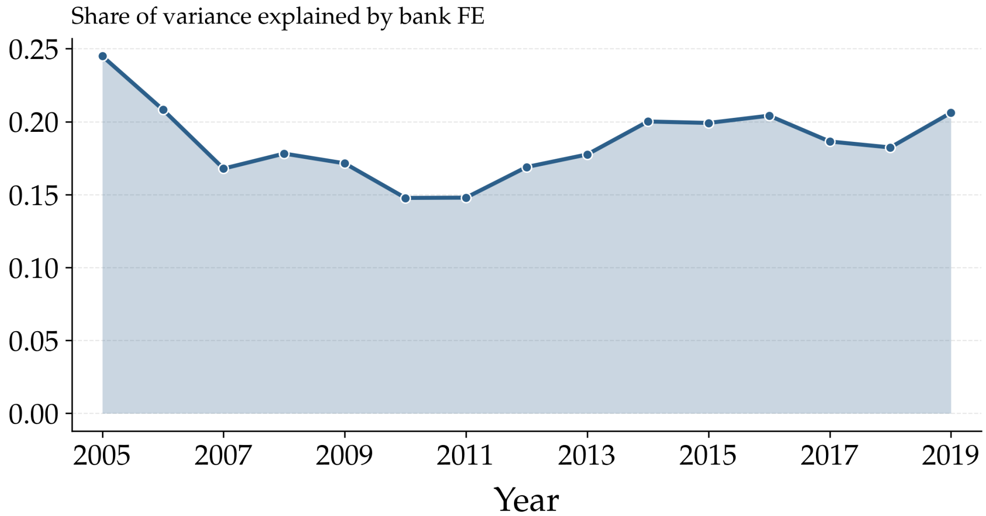
Figure 1 plots $\mathrm{Var}(\alpha_{b,t})/\mathrm{Var}(\hat{r}_{fb,t})$. The first observation is that the 2019 estimate we obtain using the Italian credit register is quite comparable to the one we obtained using AnaCredit (20% vs 17%). This is important for cross-validation, since in the Italian credit register we only observe interest payments at the bank-firm level, not at the loan level as in AnaCredit.

When looking at the time series, we can see interesting patterns. Bank fixed effects explained approximately 25% of the dispersion in interest rates in the first year of our sample, 2005. This share declined meaningfully during the 2000s and at the onset of the Great Recession, before increasing again thereafter.[21] This pattern is consistent with a period of stronger banking competition prior to the financial crisis—stimulated by the financial liberalizations following the adoption of the euro—followed by a consolidation of the banking sector after the financial crises of 2008 and

[20] As discussed in the data section, effective interest rates are computed as the ratio of interest expense to the time-weighted measure of the borrowed amount, which accounts for both the amount borrowed and the portion of the year during which it was outstanding.

[21] While the regressions are year by year, one might be concerned about confounding effects of within year variation in market interest rates. This concern is greatly mitigated because of the characteristics of overdraft loans discussed above (short-term, high percent of floating rates or frequently reset fixed rates). Moreover, because firms do not normally initiate more than one new instrument a year within a certain firm-bank relationship, most of the within year aggregate variation in market rates should be captured by firm fixed effects.

Figure 1: AKM Decomposition Over Time



Notes: The figure plots the ratio $\mathrm{Var}(\alpha_b)/\mathrm{Var}(\hat{r})$ estimated annually from the Italian credit register (TAXIA) over time.

2012, as weaker financial institutions exited and surviving banks gained market share.[22]

2.3 Bank market power

A natural question is whether the bank fixed effects reflect market power or instead capture other bank-specific characteristics, such as differences in monitoring technology or funding costs.[23] The evidence documented in Section 2.2 shows that there are substantial deviations from the law of one price: loan services offered by different banks might not be perfect substitutes and local loan markets might not be perfectly competitive. To shed light on this interpretation and to further document deviations from a hypothetical perfectly competitive benchmark, we study the relationship between the bank fixed effects estimated in Section 2.2 and loan market concentration. For this analysis we use data on "rischi a revoca" obtained by merging the Italian credit register with TAXIA, as this dataset has more information on small banks.[24]

Bank FE and market shares. Table 4 reports the average market shares of banks in each province in different years. The table reports cumulative market shares based on the average

[22]Concentration in the Italian banking industry increased in the early 2000s and again after the Great Recession. De Bonis et al. (2014) document that bank competition in Italy weakened in the late 1990s and early 2000s, a pattern consistent with the merger wave of that period, while Weber (2017) points to a further increase in concentration after the global financial crisis.

[23]It is important to note that firm fixed effects might also capture some of the bank market power related to specific bank-firm relationships (see, for instance, Sharpe, 1990; Petersen and Rajan, 1995).

[24]AnaCredit does not include some small banks with a national market share below 2%, as these credit institutions do not have to report their loans to AnaCredit. However, some of these small credit institutions such as local banks may have a larger market share at the local level.

provincial market share of banks by rank. For each year, the entry in row Top N is the sum of the average market shares of banks ranked 1 through N in each province. These market shares have been calculated using data on overdrafts (i.e., "rischi a revoca"). The average local loan market is fairly concentrated, with the three largest banks serving 60% of the market, and the five largest banks serving 78% of the market in 2019. Overall there has been a marked increase in concentration between 2005 and 2019, with the market share of the five largest banks increasing from approximately 56% to 78%.

Table 4: Cumulative market share of largest banks across provinces

Average market shares in local loan markets				
	2005	**2010**	**2015**	**2019**
Top 1	19.35	23.82	24.07	28.41
Top 3	42.45	51.85	51.69	60.34
Top 5	56.69	67.08	68.34	78.00
Top 10	77.75	85.21	87.65	94.66
Top 15	88.16	92.56	94.67	98.94

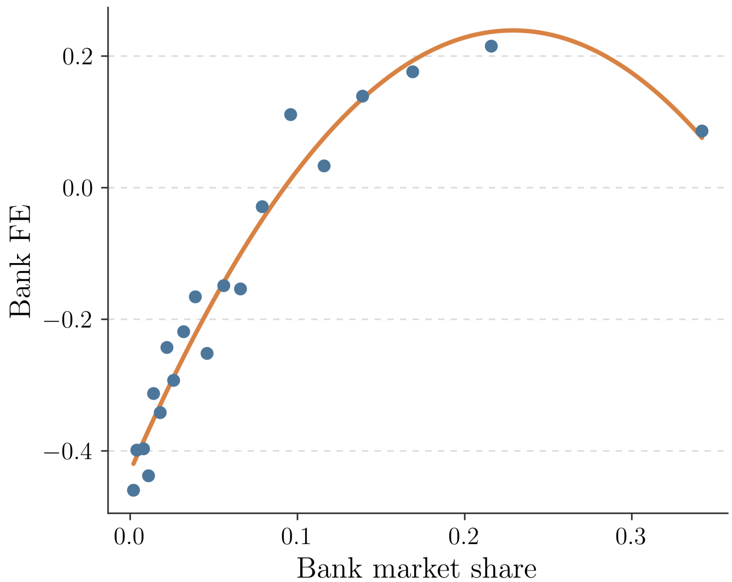
Notes: The table reports cumulative market shares based on the average provincial market share of banks by rank. For each year, the entry in row Top N is the sum of the average market shares of banks ranked 1 through N in each province.

Figure 2 provides a first piece of evidence on how bank fixed effects are related to local market shares. Bank fixed effects have been estimated for each year using cross-sectional data, and then pooled together, so that we have one fixed effect for each bank-province-year combination. Figure 2 plots these bank fixed effects against provincial market shares in a binscatter, without conditioning on any bank-level controls. The relationship is strongly positive and nonlinear: banks with a larger share of the loan market in a given province charge systematically higher rates, conditional on firm characteristics.[25]

The positive relationship in Figure 2 could, however, reflect bank-specific heterogeneity in observable characteristics (e.g., marginal funding costs) and unobservable characteristics (e.g., quality of the loan services provided by each bank) rather than market power. A bank with a high cost of funds will charge high rates, and if such banks also tend to have large market shares, the observed correlation between fixed effects and market shares could simply capture cost heterogeneity rather than differences in markups. To address this concern, we regress the estimated bank fixed effect on market share, allowing for a quadratic relationship and adding controls that proxy for variation in funding costs and other balance-sheet indicators. In particular, we include year fixed effects and a vector of time-varying bank balance-sheet characteristics X_{bt} — interest expense, operating costs, provisions, impairments, and capital, all measured as a share of total assets. To control for unob-

[25]The conditioning on firm characteristics is implicit, as the bank fixed effects have been estimated from the AKM decomposition that isolates firm fixed effects from bank fixed effects.

Figure 2: Bank Fixed Effects and Local Market Share



Notes: This figure presents a binscatter plot of the estimated bank fixed effect coefficients from regression (2) against the bank's market share, and the fitted quadratic line from Table 5.

served heterogeneity across banks we also include bank fixed effects. Importantly, the bank fixed effects included in this regression are time-invariant and should not be confused with the AKM bank effects that are the dependent variable in the regression and vary at the bank-province-year level. Formally, we estimate

$$\alpha_{bt} = c + \beta_1 s_{bt} + \beta_2 s_{bt}^2 + \mu_{\mathcal{N}(b)} + X'_{\mathcal{N}(b)t}\gamma + \delta_t + \varepsilon_{bt}, \tag{4}$$

where α_{bt} denotes the AKM bank effect estimated at the bank-province-year level, s_{bt} is bank b's provincial market share in year t, $\mu_{\mathcal{N}(b)}$ are national bank fixed effects, and δ_t are year fixed effects. The quadratic term s_{bt}^2 captures the concavity visible in Figure 2. The national bank fixed effects $\mu_{\mathcal{N}(b)}$ absorb time-invariant differences across banks while the time-varying controls $X_{\mathcal{N}(b)t}$ capture changes in balance-sheet conditions that may affect the shadow cost of funds.

Table 5 reports the results. Column (1) shows the unconditional relationship. Column (2) adds year fixed effects and time-varying bank-level controls X_{bt}. Column (3) further adds the bank fixed effects. Both $\hat{\beta}_1 > 0$ and $\hat{\beta}_2 < 0$ are stable across specifications, confirming a positive and concave relationship between market share and bank fixed effects. The coefficients decline only modestly as controls are added and remain large and precisely estimated throughout. This suggests that the positive association between bank fixed effects and market shares is not driven by heterogeneity in marginal funding costs, nor by time-invariant differences across banks in unobservable characteristics.

Table 5: Regression of Bank Fixed Effects and Local Market Share

	(1)	(2)	(3)
$\hat{\beta}_1$ (market share)	4.38***	4.02***	1.65***
	(0.22)	(0.22)	(0.22)
$\hat{\beta}_2$ (market share^2)	−7.05***	−6.32***	−2.63***
	(0.57)	(0.56)	(0.51)
Bank controls		✓	✓
Time FE		✓	✓
National Bank FE			✓
Observations	19,571	19,289	19,288
R^2	0.03	0.07	0.33

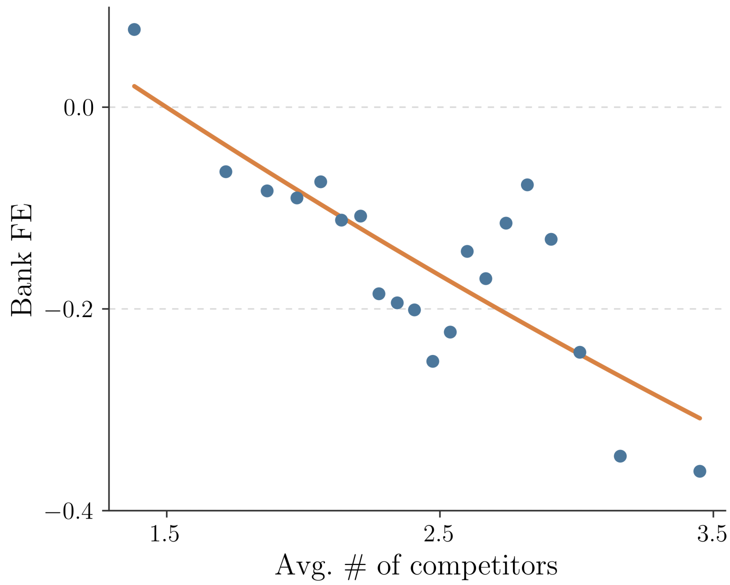
Notes: Dependent variable is the bank fixed effect α_{bt} obtained from the AKM decomposition in equation (2), estimated at the bank-province-year level. Market share is the fraction of total provincial loan volume. Bank controls include interest expense, operating costs, provisions, impairments, and capital (all as a percentage of total assets). Standard errors clustered at the bank level in parentheses. $^{***}p < 0.01$.

Bank FE and number of competitors. We next examine whether the estimated bank fixed effects vary with the average number of competing lenders faced by each bank. Specifically, we ask whether banks that lend to firms with a larger number of concurrent banking relationships tend to have lower estimated bank fixed effects. To do so, we construct, for each bank-province-year, the average number of lenders of the firms borrowing from that bank. This measure captures the extent of competition a bank faces from other lenders within its customer base. Intuitively, when a bank lends to firms that also borrow from many other banks, those firms have access to a broader set of outside financing options, reducing the bank's market power. Consistent with this intuition, we find that banks facing more competing lenders have lower estimated bank fixed effects.

Figure 3 illustrates this relationship using a binscatter. The figure shows a clear negative correlation between the estimated bank fixed effects and the average number of competing lenders faced by each bank. However, this unconditional relationship may partly reflect differences in observable and unobservable characteristics across banks. For example, banks with persistently lower funding costs or different business models may simultaneously attract different types of borrowers and lend to firms with systematically different borrowing patterns.

To address these concerns, we estimate panel regressions relating the bank fixed effects to the average number of competing lenders faced by each bank over time. We include a rich set of time-varying controls that proxy for changes in banks' funding costs and balance-sheet conditions. In particular, we include year fixed effects and a vector of bank-level characteristics, X_{bt}, comprising interest expense, operating costs, provisions, impairments, and capital, all measured as shares of total assets. We further include bank fixed effects to absorb time-invariant heterogeneity across banks. Importantly, these bank fixed effects should not be confused with the estimated AKM bank

Figure 3: Bank Fixed Effects and Number of Competitors



Notes: This figure presents a binscatter plot of the estimated bank fixed effect coefficients from equation (2) against the average number of competing lenders faced by each bank. The number of competing lenders is the average number of other banks from which firms borrowing from bank b also obtain credit.

effects, which constitute the dependent variable in the regression and vary at the bank-province-year level.

Formally, we estimate

$$\alpha_{bt} = c + \beta(N.\text{ of competing lenders})_{bt} + \mu_{\mathcal{N}(b)} + X'_{\mathcal{N}(b)t}\gamma + \delta_t + \varepsilon_{bt}, \tag{5}$$

where $\mu_{\mathcal{N}(b)}$ denotes national bank fixed effects that absorb time-invariant differences across banks, while the time-varying controls $X_{\mathcal{N}(b)t}$ capture changes in balance-sheet conditions that may affect banks' shadow cost of funds.

Table 6 reports the estimation results. Column (1) presents the unconditional relationship. Column (2) adds time-varying bank controls and year fixed effects, while Column (3) further includes bank fixed effects. Across all specifications, we find a stable and economically meaningful negative relationship between the average number of competing lenders faced by each bank and its estimated bank fixed effect. These results indicate that banks charge lower markups when competing for borrowers with more outside lending relationships. Moreover, the stability of the estimates after controlling for funding costs and other balance-sheet characteristics suggests that the negative association between bank fixed effects and measures of market power is not driven solely by heterogeneity in banks' marginal funding costs.

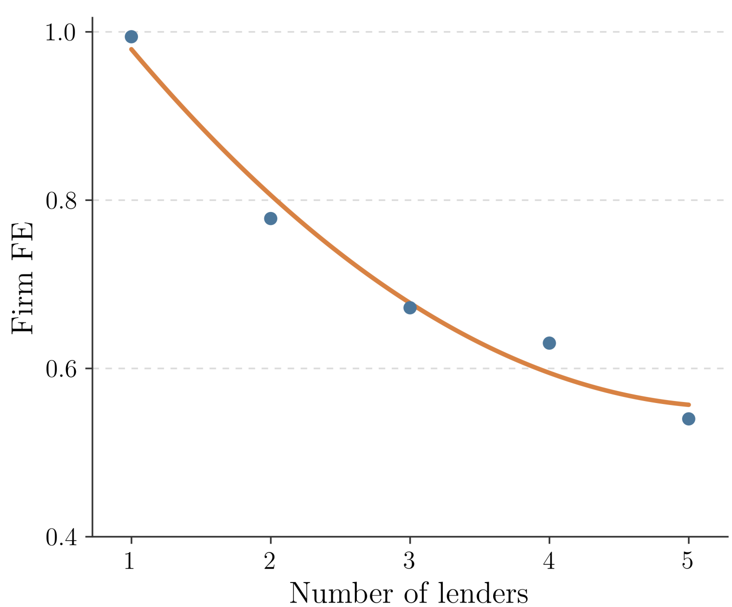
Table 6: Regression of Bank Fixed Effects and Average Number of Competitors

	(1)	(2)	(3)
$\hat{\beta}$ (N. of competing lenders)	-0.17^{***}	-0.13^{***}	-0.12^{***}
	(0.02)	(0.01)	(0.01)
Bank controls		✓	✓
Time FE		✓	✓
National Bank FE			✓
Observations	19,571	19,289	19,288
R^2	0.01	0.06	0.33

Notes: Dependent variable is the bank fixed effect α_{bt} obtained from the AKM decomposition in equation (2), estimated at the bank-province-year level. The number of competing lenders is the average number of other banks from which firms borrowing from bank b also obtain credit. Bank controls include interest expense, operating costs, provisions, impairments, and capital (all as a percentage of total assets). Standard errors clustered at the bank level in parentheses. $^{***}p < 0.01$.

Firm FE and number of bank relationships. We next examine whether firms with more banking relationships display sizable differences in the estimated firm fixed effects. If the loan market is not perfectly competitive, firms borrowing from multiple banks may be in a stronger bargaining position: because they have access to several lenders at the same time, they can more easily substitute across banks and obtain credit on more favorable terms. We therefore compute, for each firm-year, the number of distinct banks from which the firm borrows.

Figure 4: Firm Fixed Effects and Number of Bank Relationships



Notes: This figure presents a binscatter plot of the estimated firm fixed effect coefficients from equation (2) against the number of bank relationships of the firm. The number of bank relationships is the number of distinct banks from which the firm borrows in a given year.

Figure 4 illustrates this relationship using a binscatter. The figure shows that firms with more bank relationships tend to have lower estimated firm fixed effects. This pattern is consistent with the idea that firms with broader access to credit markets face lower borrowing costs because they can rely on multiple lending relationships. Of course, this unconditional relationship may also reflect differences in firm size, balance-sheet strength, or other persistent firm characteristics. For example, larger or less levered firms may both borrow from more banks and obtain better loan terms.

To account for these differences, we estimate panel regressions relating the estimated firm fixed effects to the number of bank relationships at the firm-year level. We include year fixed effects and a vector of time-varying firm controls, X_{it}, comprising the number of employees, total assets, and financial leverage. We also include firm fixed effects to absorb time-invariant heterogeneity across firms. As above, these fixed effects are distinct from the estimated AKM firm effects, which are the dependent variable in the regression and vary at the firm-year level.

Formally, we estimate

$$\alpha_{it} = c + \beta(N.\text{ of bank relationships})_{it} + \mu_i + X_{it}'\gamma + \delta_t + \varepsilon_{it}, \tag{6}$$

where μ_i denotes firm fixed effects and the time-varying controls X_{it} capture changes in firm size and balance-sheet conditions that may affect borrowing terms.

Table 7 reports the estimation results. Column (1) presents the unconditional relationship. Column (2) adds firm controls and year fixed effects, while Column (3) further includes firm fixed effects. Across specifications, the coefficient on the number of bank relationships is negative, consistent with firms obtaining more favorable borrowing terms when they have access to a broader set of lenders. The inclusion of firm controls and firm fixed effects helps distinguish this pattern from differences in observable firm characteristics and time-invariant heterogeneity across firms.

Alternative interpretations of bank fixed effects. There may be factors other than market power driving heterogeneity in bank fixed effects. For instance, heterogeneity in bank fixed effects may partly reflect heterogeneity in funding costs, since banks with higher funding costs may charge higher rates. As we show in Section 3, this force is also present in our model: heterogeneity in bank fixed effects reflects both differences in marginal funding costs and differences in markups across banks. Importantly, both in the data and in the model, bank fixed effects are increasing in market shares after controlling for funding costs.

Bank fixed effects might also capture heterogeneity in risk premia across banks. Indeed, even if we residualize interest rates with respect to default risk in the first step of our analysis, bank fixed effects may still partially capture default risk if different banks price the same default risk differently. To show that the substantial dispersion in bank fixed effects we documented above is

Table 7: Regression of Firm Fixed Effects and Number of Bank Relationships

	(1)	(2)	(3)
$\hat{\beta}$ (N. of bank relationships)	-0.12^{***}	-0.13^{***}	-0.12^{***}
	(0.01)	(0.01)	(0.01)
Firm controls		✓	✓
Time FE		✓	✓
Firm FE			✓
Observations	724,791	701,377	669,295
R^2	0.01	0.02	0.55

Notes: Dependent variable is the firm fixed effect α_{it} obtained from the AKM decomposition in equation (2), estimated at the firm-year level. The number of bank relationships is the number of distinct banks from which firm i borrows in a given year. Firm controls include the number of employees, total assets, and financial leverage. Standard errors clustered at the firm level in parentheses. $^{***}p < 0.01$.

not driven by bank-specific risk premia, we estimate two alternative AKM decompositions using TAXIA.

The results are reported in Table 8. Column (1) reports the variance of firm and bank fixed effects after including interactions between national bank fixed effects and credit risk in the first-stage residualization. This specification absorbs not only differences in default risk across firms, but also heterogeneity across banks in the pricing of default risk. It is therefore a particularly demanding specification, since it also absorbs national bank fixed effects. Nevertheless, we find that more than 10 percent of the dispersion in residual interest rates is still explained by bank fixed effects, even after absorbing potential heterogeneity in risk premia across banks. For small local banks that operate in only one province, this specification absorbs all bank-level variation already in the first stage. Thus, the remaining bank fixed effects capture variation in the rates charged by the same bank across different local loan markets. Column (1) therefore implies that at least 10 percent of the dispersion in interest rates is explained by how the same banks charge different rates across local markets. As a second exercise, we estimate the AKM decomposition separately for two subsamples of high-risk and low-risk firms. While the share of variance explained by bank fixed effects is mechanically larger than in our baseline specification because each subsample contains only half of the firms, it remains broadly similar across the two subsamples, as shown in Columns (2) and (3) of Table 8.

Table 8: Robustness: risk premia

	(1)	(2)	(3)
Variance share: firm FE	0.68	0.61	0.61
Variance share: bank FE	0.13	0.30	0.25

Notes: The table reports the share of residual interest-rate variance explained by firm and bank fixed effects in alternative AKM decompositions. Column (1) includes interactions between national bank fixed effects and credit risk in the first-stage residualization. Columns (2) and (3) report results from separate decompositions for high-risk and low-risk firms.

3 Model

Our model builds on the macro-banking framework of Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), augmented with oligopolistic competition in the loan market through a random matching process between firms and banks. We consider an economy populated by a representative household, a continuum of firms, and a finite number of banks. Time is discrete and the horizon is infinite. We start by describing the environment, then define the equilibrium, and end by characterizing the optimal interest rates set by banks and showing how the structural elasticities map to the reduced-form elasticities we later estimate in the data.

3.1 Environment

Households. The representative household maximizes expected discounted utility:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[C_t - \frac{\chi N_t^{1+\psi}}{1+\psi} \right], \tag{7}$$

where C_t is consumption, N_t is labor supply, $\beta \in (0,1)$ is the discount factor, $\chi > 0$ scales the disutility of labor, and $\psi > 0$ is the inverse Frisch elasticity of labor supply. The household supplies labor to firms, holds deposits a_t in banks, receives dividends from banks when they exit, and collects profits from firms. We normalize the price of consumption to 1. The budget constraint is given by

$$C_t = w_t N_t + \Pi_t + a_t i_{at}, \tag{8}$$

where w_t is the real wage and Π_t denotes the aggregate level of dividends distributed in period t. The latter includes all firm profits together with the dividends distributed from banks. The interest rate on deposits is denoted by i_{at}.

Firms. There is a unit measure of firms. Each firm $f \in (0,1)$ produces a perfectly substitutable good using labor as the sole input of production. We assume each firm's production function features decreasing returns to scale, $\sigma \leq 1$:[26]

$$y_f = n_f^{\sigma}, \tag{9}$$

where n_f denotes the amount of labor firm f uses, and y_f its production.

Production requires working capital. A fraction γ of the wage bill must be paid upfront using credit from banks. We denote by ℓ_f the overall loan amount of firm f. The working capital requirement is given by

$$\ell_f \geq \gamma w n_f. \tag{10}$$

We assume each firm f has access to a random subset $B(f) \subseteq \mathcal{B}$ of banks, where $\mathcal{B}$ denotes the set of all banks in the economy. We assume that a firm is matched with each individual bank with probability ξ. This matching process is iid across relationships and across firms. The expected number of available banks per firm therefore equals ξN_B, where N_B is the number of banks in the economy. Loan services from different banks are imperfectly substitutable. In particular, we assume that the overall loan amount of firm f is a CES aggregator across all individual loans. That is,

$$\ell_f = \left(N_{Bf}^{-\frac{1}{\eta}} \sum_{b \in B(f)} \ell_{fb}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}, \tag{11}$$

where $\eta > 1$ is the elasticity of substitution across banking relationships, and N_{Bf} denotes the number of banks firm f is connected with.

Imperfect substitutability across banking relationships captures, in reduced form, frictions that make loans from different banks only partially interchangeable from the firm's perspective. These frictions may reflect relationship-specific information, differences in contract terms or lending technologies, and adjustment costs in reallocating borrowing across lenders. As a result, substitution across banks is feasible but limited, so each bank faces a downward-sloping residual demand and retains some degree of market power over the firm.

We assume that each bank the firm is matched with offers an interest rate specific to that firm which we denote by i_{fb}. The firm takes as given the interest rates offered by its entire banking network, and decides the amount of borrowing from each bank. In summary, the firm's problem

[26] For ease of notation, we omit the time subscript in the firm block.

is given by

$$
\begin{aligned}
\pi_f = \quad & \max_{\{n_f, \ell_f, \{\ell_{fb}\}_{b\in B(f)}\}} \quad && n_f^\sigma - wn_f - \sum_{b\in B(f)} \ell_{fb} i_{fb}, && (12)\\
& \text{s.t.} && \gamma w n_f \le \ell_f, \\
& && \ell_f = \left(N_{Bf}^{-\frac{1}{\eta}} \sum_{b\in B(f)} \ell_{fb}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}.
\end{aligned}
$$

Banks. There is a finite number of banks in the economy. Banks are denoted by $b \in \mathcal{B}$. The net worth of bank b at the start of period t is denoted by k_{bt}. Each bank finances their loan portfolio through a combination of equity (k_{bt}) and household deposits (a_{bt}), so that

$$\int \ell_{fbt} df \le k_{bt} + a_{bt}. \tag{13}$$

Banks are subject to a leverage constraint (capital requirement) that limits the ratio of loans to equity:

$$\frac{k_{bt}}{\int_f \ell_{fbt} df} \ge \theta, \tag{14}$$

where $\theta \in (0,1)$ is the required capital ratio. This constraint can be interpreted in reduced form as capturing regulatory capital requirements, incentive constraints arising from limited commitment, or increasing marginal funding costs associated with higher leverage.

Each bank is matched with a fraction ξ of firms. Because there is a continuum of firms, the matching process does not lead to any uncertainty from the bank's perspective, nor any uncertainty at the aggregate level. The only shock banks are subject to is an "exit" shock. With probability $1-\phi$, the bank survives the period and retains its profits:

$$k_{bt+1} = \int_{f: b\in B_t(f)} (1+i_{fbt})\ell_{fbt} df - (1+i_{at})a_{bt}. \tag{15}$$

With probability ϕ, the bank exits, remits its remaining net worth to the household as a dividend, and is replaced by a new bank with initial net worth k_0. This random exit process gives the model a well-defined stationary distribution of bank sizes.

Banks compete à la Bertrand. Importantly, the matching between firms and banks is predetermined at the start of each period and fixed for the duration of the period, so the oligopolistic competition in interest rates is static conditional on the network. A bank b decides on the interest rate they offer firm f at time t, i_{fbt}, taking as given the interest rates set by all of its competitors. Denote the residual loan demand the bank faces from a firm by $\ell(i_{fb}, i_{f-b})$, where i_{f-b} denotes the interest rates set by all other firms matched with firm f. The bank chooses these interest rates to maximize their end-of-period net worth:

$$\widetilde{k}_{bt+1} = \max_{\{a_{bt},\{i_{fbt}\}\}} \int_{f:b\in B(f)} (1+i_{fbt})\ell(i_{fbt}, i_{f-bt})df - (1+i_{at})a_{bt}, \tag{16}$$
$$\text{s.t.} \quad \int_{f:b\in B(f)} \ell(i_{fbt}, i_{f-bt})df \leq k_{bt} + a_{bt}.$$
$$k_{bt} \geq \theta \int_{f:b\in B(f)} \ell_{fbt}df,$$

3.2 Equilibrium

Given an initial distribution of bank net worth $\{k_{b0}, \widetilde{k}_{b0}\}$, an equilibrium consists of sequences of quantities $\{N_t, C_t, \Pi_t, a_{bt}, \{\ell_{fbt}\}, \{k_{b,t+1}\}, \{\widetilde{k}_{b,t+1}\}, \{\pi_f\}\}$ and prices $\{w_t, i_{at}, \{i_{fbt}\}\}$ for all banks b and firms f such that:

1. Consumption and labor solves the household's problem, given the real wage and interest rate on deposits.

2. For all firms, loans and employment maximize their profits, given the real wage and the vector of interest rates they are offered from banks.

3. For all banks, their interest rates, loans, and deposits maximize their end-of-period net worth, taking as given the real wage, the banking network, and all interest rates charged by their competitors.

4. Aggregate dividends in the economy satisfy

$$\Pi_t = \int \pi_{ft}df + \sum_b \left(k_{bt} - \widetilde{k}_{bt}\right) \tag{17}$$

5. All markets clear. Specifically, aggregate labor demand equals labor supply, aggregate deposit demand equals deposit supply, and consumption demand equals aggregate output produced.

3.3 Loan Demand and Interest Rate Determination

We now characterize firms' demand for loans from each bank and the resulting interest rate setting problem faced by banks.

Loan demand. Fix a firm f and take as given the vector of interest rates i_{fb} it faces. The firm chooses its loans to minimize total borrowing costs for a given level of total loans ℓ_f. Standard cost minimization with a CES aggregator implies that loan demand from bank b is given by

$$\ell_{fb} = \frac{1}{N_{Bf}} \left(\frac{i_{fb}}{i_f} \right)^{-\eta} \ell_f, \tag{18}$$

where

$$i_f = \left(\frac{1}{N_{Bf}} \sum_{b \in B(f)} i_{fb}^{1-\eta} \right)^{\frac{1}{1-\eta}} \tag{19}$$

is the firm-specific CES loan rate index.

The working capital constraint (10) binds in equilibrium, implying $\ell_f = \gamma w n_f$. Substituting into the firm's problem yields the optimal labor demand, $n_f = \left(\frac{\sigma}{w(1+\gamma i_f)} \right)^{\frac{1}{1-\sigma}}$. Combining expressions, the demand faced by bank b from firm f can be written as

$$\ell_{fb} = \frac{1}{N_{Bf}} \left(\frac{i_{fb}}{i_f} \right)^{-\eta} \gamma w \left(\frac{\sigma}{w(1+\gamma i_f)} \right)^{\frac{1}{1-\sigma}}. \tag{20}$$

This expression highlights two key margins: (i) substitution across lenders, governed by η, and (ii) the overall scale of borrowing, which depends on the composite loan rate i_f through firms' labor demand.

Bank pricing. Consider bank b setting the rate i_{fb} for a given firm f, taking as given the rates charged by competing banks and the aggregate conditions. The bank chooses i_{fb} to maximize profits from this relationship:

$$\max_{i_{fb}} \quad (i_{fb} - \lambda_{bt}) \ell(i_{fb}, i_{f-b}), \tag{21}$$

where λ_{bt} denotes the marginal cost of funding for the bank. This marginal cost is given by the Lagrange multiplier on the bank's balance sheet constraint in the bank's problem (16), and reflects the shadow value of internal funds, incorporating both the deposit rate and the leverage constraint.

Solving this profit maximization problem, we obtain the standard markup formula:

$$i_{fb} = \frac{\varepsilon(i_{fb}, i_{f-b})}{\varepsilon(i_{fb}, i_{f-b}) - 1} \lambda_{bt}, \tag{22}$$

where

$$\varepsilon(i_{fb}, i_{f-b}) \equiv -\frac{\partial \ell(i_{fb}, i_{f-b})}{\partial i_{fb}} \frac{i_{fb}}{\ell(i_{fb}, i_{f-b})} \tag{23}$$

is the elasticity of residual loan demand faced by bank b from firm f.

In our setup, the demand elasticity $\varepsilon(i_{fb}, i_{f-b})$ is endogenous and depends on the firm's loan expenditure share with bank b in the following way:

$$\varepsilon(i_{fb}; i_f) = \underbrace{\eta\,(1 - s_{fb})}_{\text{credit substitution}} + \underbrace{\frac{1}{1-\sigma} \cdot \frac{\gamma i_f}{1 + \gamma i_f} \cdot s_{fb}}_{\text{credit demand}}, \tag{24}$$

where $s_{fb} = \frac{\ell_{fb} i_{fb}}{\ell_f i_f}$ denotes the expenditure share of bank b in firm f's total loan expenditure. The coefficient $\frac{1}{1-\sigma}$ in the second term follows directly from the firm's labor demand schedule: the elasticity of labor demand with respect to its interest rate index equals $\frac{1}{1-\sigma}\frac{\gamma i_f}{1+\gamma i_f}$.

The first term, $\eta(1 - s_{fb})$, captures the credit substitution channel: when bank b raises its rate, firm f reallocates borrowing toward other banks in $B(f)$. This effect is proportional to $1 - s_{fb}$: a bank with a larger expenditure share faces less elastic substitution demand, because the firm has less scope to shift borrowing toward alternative lenders. The second term captures the credit demand channel: a higher rate charged by bank b raises the firm's average funding cost i_f, which reduces labor demand and therefore total borrowing through the working capital constraint. This effect is proportional to s_{fb} and becomes stronger when a larger fraction of the wage bill must be financed externally, that is, when γ is higher.

The decomposition in (24) implies that a bank's expenditure share matters for the markup it charges. A bank accounting for a larger share of a firm's borrowing faces, all else equal, a less elastic residual demand curve and therefore optimally sets a higher markup. In particular, the credit substitution term $\eta(1 - s_{fb})$ declines with s_{fb}: as the bank's share increases, the firm can substitute away from it less easily. By contrast, the credit demand term increases in s_{fb}, but its slope is typically smaller when γ is moderate. Hence, whenever the substitution channel dominates, the overall elasticity declines with market share, generating a positive relationship between expenditure share and the interest rate markup.

The demand elasticity faced by the oligopolistic banks resembles the demand elasticity with oligopolistic competition in Atkeson and Burstein (2008). In their setup, aggregate consumption is nested CES across sectors and firms within each sector, and the demand elasticity faced by each firm is a weighted average between the within-nest and across-nests elasticities, with weights equal to the firm's market share in its sector. Similar to their setup, the elasticity in the credit substitution channel is the elasticity of substitution across lenders and the weights equal the bank's market share. However, in our setup, the second term comes from the firm's demand elasticity for loans and not from an outer nest, and depends on the composite interest rate faced by the firm.

3.4 AKM decomposition in the model

We now show how the equilibrium interest rates set by banks map into the firm and bank fixed effects estimated in Section 2. This mapping gives a structural interpretation to the empirical

AKM decomposition and clarifies which forces in the model are needed to rationalize the patterns documented there.

We take logs of the pricing rule in equation (22). Letting $\mu_{fb} \equiv \varepsilon(i_{fb}, i_{f-b})/[\varepsilon(i_{fb}, i_{f-b})-1]$ denote the equilibrium markup, the pricing rule implies $\ln i_{fb} = \ln \mu_{fb} + \ln \lambda_b$. Adding and subtracting cross-sectional means delivers a log-additive decomposition:

$$\ln i_{fb} \;=\; \bar{\iota} \;+\; \underbrace{(\bar{\mu}_f - \bar{\mu})}_{\alpha_f\text{: firm FE}} \;+\; \underbrace{\left(\ln \lambda_b - \bar{\lambda}\right) + (\bar{\mu}_b - \bar{\mu})}_{\alpha_b\text{: bank FE}} \;+\; u_{fb}, \tag{25}$$

where $\bar{\mu} \equiv \mathbb{E}[\ln \mu_{fb}]$ is the average log markup in the economy, $\bar{\mu}_f \equiv \mathbb{E}_f[\ln \mu_{fb}]$ is firm f's average log markup across its lenders, $\bar{\mu}_b \equiv \mathbb{E}_b[\ln \mu_{fb}]$ is bank b's average log markup across its borrowers, $\bar{\lambda} \equiv \mathbb{E}[\ln \lambda_b]$ is the average log marginal cost of funds, and $\bar{\iota} \equiv \bar{\mu} + \bar{\lambda}$. The term $u_{fb} \equiv \ln \mu_{fb} - \bar{\mu}_f - \bar{\mu}_b + \bar{\mu}$ is the interaction component of the markup that is not absorbed by either fixed effect.

Equation (25) reveals the economic content of each fixed effect. The firm fixed effect α_f captures how the average markup faced by firm f deviates from the economy-wide mean. Firms matched with a single large bank have a lower loan demand elasticity and therefore pay a higher markup, $\bar{\mu}_f$. Firms matched with many small banks substitute more effectively across lenders and pay lower markups on average.

The bank fixed effect α_b has two components. The first, $\ln \lambda_b - \bar{\lambda}$, captures dispersion in the shadow cost of funds across banks. We refer to this as the "financial frictions" channel: banks with tighter leverage constraints face a higher λ_b and pass this cost through to their borrowers, even in the absence of any market power. The second, $\bar{\mu}_b - \bar{\mu}$, captures dispersion in the average markup that bank b charges across the firms it lends to. We refer to this as the "market power" channel: a bank that systematically holds a large expenditure share in its borrowers charges, on average, higher markups, and therefore contributes more to the bank component of interest rate dispersion. Both channels are active in the data and both are present in the model. Disentangling them is the central task of the quantitative analysis.

To clarify the role of each ingredient, we consider two special cases. Under perfect competition ($\eta \to \infty$ and $N_B \to \infty$), markups converge to one for all bank-firm pairs and the law of one price holds: neither fixed effect carries any variation. Under monopolistic competition ($\eta < \infty$ and $N_B \to \infty$), each bank faces a downward-sloping demand curve but holds a vanishing expenditure share in every firm. Markups are then symmetric across banks, $\bar{\mu}_b = \bar{\mu}$ for all b, and the bank fixed effect reflects cost heterogeneity alone. The full oligopolistic model, with finite N_B, generates dispersion in both components of α_b.

3.5 Discussion

Before turning to the quantitative analysis, let us discuss some of our modeling assumptions.

We assume that loan services from different banks are imperfect substitutes, with substitutability governed by the CES parameter η. This assumption captures, in reduced form, a variety of frictions that make loans from different banks only partially interchangeable: relationship-specific information, differences in contract terms and lending technologies, and adjustment costs in reallocating borrowing across lenders. An alternative approach would be to microfound the imperfect substitutability through an explicit model of search frictions or switching costs. We view the CES formulation as a parsimonious representation that allows us to focus on the interaction between market structure and financial frictions, which is the main object of interest.

We assume that banks compete in interest rates (Bertrand-Nash) and that each bank sets a single interest rate per firm. The Bertrand assumption reflects the fact that interest rates are the primary dimension along which banks differentiate their loan offers. The restriction to linear pricing rules out quantity discounts, two-part tariffs, and other forms of nonlinear contracts. Allowing for nonlinear pricing would give banks additional instruments to extract surplus and would likely increase banks' rents, but at a considerable cost in tractability given the oligopolistic setting.

Bank-firm matching is assumed to be random and independent of bank characteristics. In practice, matching is likely influenced by geography, relationship lending, and information frictions. We view random matching as a tractable approximation that captures the key feature of the data, namely that most firms borrow from a small number of banks, while remaining agnostic about the specific determinants of match formation. An important consequence of this assumption is that the banking network is exogenous: firms cannot respond to high markups by seeking new banking relationships, and banks cannot compete for new customers. Allowing firms to endogenously form and sever banking relationships would introduce a dynamic competition margin that disciplines markups over time. We abstract from this margin to keep the model tractable and to isolate the static forces that drive interest rate dispersion in the cross section.

Finally, our analysis focuses on market power in lending and assumes that the deposit market is competitive. This is a modeling choice, not a statement about reality: a large literature, including Drechsler et al. (2017), documents significant market power in deposit markets. Allowing for deposit market power will introduce another source of heterogeneity in banks' funding costs. We leave this extension for future work.

4 Quantitative Analysis

In this section, we estimate the model using data from the credit register and show that the estimated model can successfully match key moments of the data, including the AKM decomposition, without targeting them. We then use the calibrated model to study the macroeconomic implications of bank market power and the role that market power plays in shaping the effects of macroprudential policy.

4.1 Loan Demand Elasticities

Loan demand elasticities play a crucial role in the model presented in Section 3, as these elasticities affect markups and interest rates. In this section we will focus on the role played by two distinct structural loan demand elasticities. The total loan demand elasticity $\partial \log(\ell_{fb})/\partial i_{fb}$ captures how firm f's total borrowing from bank b responds to a change in that bank's rate, holding the rates charged by all competing banks fixed. Equation 24 shows that the loan demand elasticity is the sum of two terms, capturing reallocation of credit across banks (i.e., credit substitution) and the change in the overall loan demand from firm f. The second elasticity we focus on is the elasticity of portfolio shares $\partial \log(\ell_{fb}/\ell_f)/\partial i_{fb}$ (i.e., elasticity in shares): it measures how much the firm reallocates borrowing across banks within its network, again holding the rates of competing banks fixed. Formally, these structural elasticities can be expressed as:

$$\left.\frac{\partial \log \ell_{fb}}{\partial \log i_{fb}}\right|_{i_{f,-b}} = -\eta(1-s_{fb}) - \frac{1}{1-\sigma} \cdot \frac{\gamma i_f}{1+\gamma i_f} \cdot s_{fb}, \tag{26}$$

$$\left.\frac{\partial \log(\ell_{fb}/\ell_f)}{\partial \log i_{fb}}\right|_{i_{f,-b}} = -\eta(1-s_{fb}), \tag{27}$$

where $s_{fb} = \ell_{fb} i_{fb}/(\ell_f i_f)$ is the expenditure share of bank b in firm f's total loan expenditure. The first term in equation (26) captures credit substitution across banks within the firm's set; the second captures the reduction in total loan demand as the firm's average funding cost i_f rises. Each expression is evaluated holding $i_{f,-b}$ fixed—these are partial-equilibrium objects, and the relevant demand elasticities for the bank's first-order condition (22).

Structural versus reduced-form semi-elasticities. When banks behave strategically, the structural elasticity cannot be recovered directly from observed loan-level responses to changes in interest rates, not even when changes in rates are driven by well-identified bank-specific shocks. By construction, the structural elasticity is a partial-equilibrium object: it answers the counterfactual question, how much must bank b change the interest rate i_{fb} it offers to firm f in order to change loan demand by one percent, holding the terms offered by competing banks fixed? In equilibrium, however, a shock to one bank induces competing banks to best respond, which in turn affects the original bank's optimal pricing decision. As a result, the response measured in the data after a bank-level shock is a reduced-form elasticity, which incorporates the full equilibrium adjustment of loan quantities and interest rates across all banks in the market. The presence of strategic complementarities in price setting with oligopolistic competition have been studies in the product market, the labor market, and in the credit card market (Amiti et al., 2016; Berger et al., 2022; Morelli and Herkenhoff, 2024).

Although reduced-form elasticities differ from structural elasticities, the former still contain useful information about the latter (Berger et al., 2022). Our model implies that up to a first order approximation the reduced-form elasticity of substitution and the reduced-form elasticity of loan demand are equal to:

$$\frac{d\log(\ell_{fb}/\ell_f)}{d\log i_{fb}} = -\eta(1-s_{fb}) + \sum_{b'\neq b} \eta s_{fb'} \frac{d\log i_{fb'}}{d\log i_{fb}} \tag{28}$$

$$\frac{d\log \ell_{fb}}{d\log i_{fb}} = -\varepsilon(i_{fb}; i_f) + \sum_{b'\neq b} \Big[\eta - \varepsilon(i_{fb'}; i_f)\Big] \frac{d\log i_{fb'}}{d\log i_{fb}} \tag{29}$$

Note that the reduced-form elasticity of substitution is always smaller in absolute value than its structural counterpart. The same applies to the elasticity of loan demand as long as $\eta > \epsilon(i_{fb'}, i_f)$, i.e., as long as $\eta > \frac{1}{1-\sigma}\frac{\gamma i_f}{1+\gamma i_f}$.

Figure 5 illustrates the difference between these elasticities. Suppose that the marginal cost of funds of bank b increases from mc_b to mc'_b. In each panel, the blue solid curve traces the structural demand for bank b's loans as a function of i_{fb}, holding competitors' rates fixed at $i_{f,-b}$; the red solid curve is the demand schedule after competing banks have raised their rates to $i'_{f,-b}$ in response to higher marginal costs. Point A is the initial equilibrium.

Panel A focuses on the elasticity of substitution. When the marginal cost of funds for bank b increases, it is optimal for bank b to increase interest rates. Starting from point A, the structural response from point A to point B moves along the blue demand curve: bank b raises its rate and, if the rates of competing banks are held fixed, the firm substitutes its borrowing toward lower-cost alternatives, so bank b's market share would decline substantially.

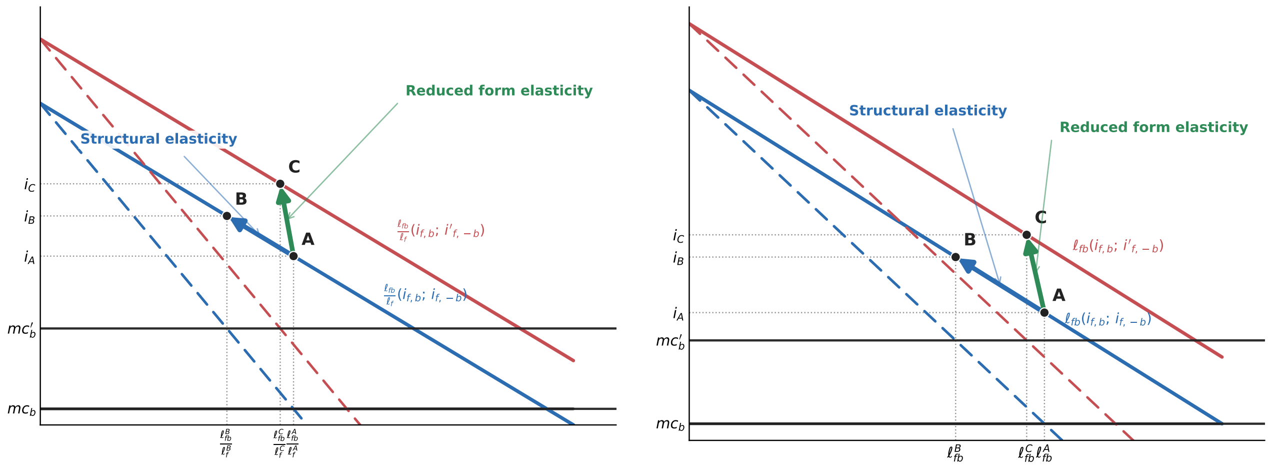
In practice, firm f may have access to other banks and these banks will not keep their rates fixed; rather, they will optimally adjust their rates in response to an increase in i_{fb}. Thus, the loan demand schedule shifts upward, and point C is the new equilibrium resulting from an increase in bank b's marginal cost from mc_b to mc'_b. Since competing banks also raise their rates, the firm's outside option deteriorates, and the demand for bank b's loans shifts upward. Bank b therefore raises its rate further yet retains more of its share—the movement from point A to point C is steeper in $(i_{fb},\ \ell_{fb}/\ell_f)$ space than the movement from point A to point B. The structural elasticity in shares thus would exceed the reduced-form estimate in absolute value; using the reduced-form elasticity directly to gauge the substitution channel would understate the substitutability of loan demand.

Panel B depicts the loan demand elasticity. The structural response from point A to point B features a substantial rate increase and a moderate contraction of total lending: holding competitors' rates fixed limits how much the composite loan rate i_f rises, so the scale effect on total firm borrowing is contained. The reduced-form equilibrium is instead point C, on the red demand curve: competing banks also raise their rates. The loan demand schedule for bank b shifts upwards

Figure 5: Structural and Reduced-Form Loan Demand Semi-Elasticities

A. Loan demand elasticity

B. Total loan demand semi-elasticity



Notes: Panel A (left) shows the difference between the structural loan demand elasticity of substitution, $\partial \log(\ell_{fb}/\ell_f)/\partial i_{fb}$, which isolates credit substitution across banks, and its reduced-form counterpart. Panel B (right) shows the difference between total loan demand elasticities, $\partial \log \ell_{fb}/\partial i_{fb}$. In each panel, the blue solid curve is the demand schedule holding competitors' rates $i_{f,-b}$ fixed, while the red solid curve is the demand schedule after competitors have optimally adjusted their rates to $i'_{f,-b}$; the dashed curves are associated marginal revenues. Point A is the initial equilibrium; B is obtained following an increase in bank b's marginal cost from mc_b to mc'_b along the structural demand; C is the new equilibrium allowing competing banks to best-respond.

provided that $\eta > 1$. The structural loan demand elasticity would thus exceed the reduced-form estimate in absolute value; using the reduced-form elasticity directly to gauge the loan demand elasticity would overstate the rigidity of loan demand.

Estimates of reduced-form elasticities. Since the reduced-form loan demand elasticities are informative about the structural elasticities in the model, we adopt an indirect inference approach by estimating reduced form elasticities in the model and in simulated data.

We estimate the two reduced-form semi-elasticities using the Italian credit register. Our focus is on semi-elasticities so that results are independent on the level of interest rates in the data that is associated to default risk. The estimating equations use Δi_{fb}—the change in the interest rate level—so the estimated coefficients are semi-elasticities, measuring the change in log loan demand per unit change in the interest rate:

$$\Delta \log \ell_{fb} = \epsilon_{\text{tot}}\, \Delta i_{fb} + e_{fb}, \tag{30}$$

$$\Delta \log(\ell_{fb}/\ell_f) = \epsilon_{\text{subst}}\, \Delta i_{fb} + e_{fb}. \tag{31}$$

Identification relies on a shift-share instrument following Amiti and Weinstein (2018). For each bank-firm pair (b, f), we construct the instrument as the change in the average rate that bank b charges to all *other* firms in the same province p, excluding firm f itself. Formally, let $\bar{i}_{bt}^{-f}$ denote

Table 9: Estimated Loan Demand Elasticities

	Total (ϵ_{tot})	Substitution (ϵ_{subst})
Estimate	-19.044^{***}	-12.153^{***}
	(2.056)	(2.003)
Observations	88,521	88,428

Notes: The table reports IV estimates of the total and substitution loan demand elasticities from equations (30) and (31). *** denotes significance at the 1% level.

this leave-one-out average. The instrument isolates bank-level variation in lending rates that is common across the bank's other borrowers and therefore plausibly orthogonal to the idiosyncratic demand conditions facing firm f. To remove common macroeconomic and regional variation, we partial out aggregate year fixed effects and province-year fixed effects before constructing the instrument, so that the identifying variation is purely cross-bank within a province in a given year.

Estimates of reduced-form loan demand elasticities are reported in Table 9. The point estimate for loan demand elasticity implies that when the interest rate charged by bank b to firm f increases by one percentage point, loan demand of firm f from bank b falls by 19%. Similarly, the point estimate for the elasticity of substitution implies that when the interest rate charged by bank b to firm f increases by one percentage point, then the market share of bank b in the portfolio of loans of firm f falls by 12%. The estimates are broadly consistent with those reported by Dempsey and Faria-e-Castro (2025) and Altavilla et al. (2022). We use these estimates to discipline the demand parameters σ and η in the model using an indirect inference approach: for each candidate parameter vector (η, σ), we simulate a panel of bank-firm relationships from the model and apply the Amiti-Weinstein estimation procedure to the simulated data, recovering model-implied reduced-form semi-elasticities. We then choose η and σ to minimize the distance between these model-implied moments and the data estimates in Table 9.

4.2 Calibration and Validation

Calibration. The model has eleven parameters: $\Theta = [\psi, \chi, \beta, \gamma, \theta, \xi, B, \phi, k_0, \sigma, \eta]$. Our quantitative strategy consists of choosing Θ to match a set of moments from the Italian credit register.

A first set of common parameters is set equal to standard values in the macroeconomic literature. We set the inverse Frisch elasticity $\psi = 1$ and $\beta = 0.99$. We set $\gamma = 0.40$ following Bocola and Bornstein (2023), who document this ratio of short-term debt to revenues for Italian firms. The capital requirement $\theta = 0.08$ corresponds to the Basel I/II minimum Tier 1 capital ratio applied to corporate loans, which carry a risk weight of one so that the effective requirement equals 8 percent of loan volume. The scale parameter χ in the utility function is normalized to one.

The remaining parameters $[\xi, B, \phi, k_0, \sigma, \eta]$ are jointly estimated by method of simulated moments. While these parameters are jointly estimated, we provide some intuition on how they are

Table 10: Calibrated Parameter Values

Parameter	Description	Value
η	Loan CES index	4.5
σ	Returns to scale	0.98
ϕ	Exit probability	0.12
B	Number of banks	10
k_0	Initial bank net worth	10^{-6}
ξ	Probability of matching with each bank	0.18
θ	Minimum net-worth-to-asset ratio	0.08
β	Discount factor	0.99
ψ	Frisch elasticity	1
χ	Labor disutility	1
γ	Working capital ratio	0.40

Table 11: Calibration Targets: Data and Model

Calibration target	Data	Model
Semi-elasticity (ϵ_{tot})	−19	−20
Semi-elasticity in shares (ϵ_{subst})	−12	−12
Three largest banks' market share (%)	65	61
Five largest banks' market share (%)	84	84
Ten largest banks' market share (%)	99	100
Average number of banks per firm	2.1	2.1

identified by the targeted moments. We set $B = 10$ to match the number of banks that account for 99 percent of the market in the sample. The matching probability ξ is disciplined by the average number of banks per firm, which equals 2.1 in the data. The exit probability ϕ and the initial net worth k_0 are jointly identified from the distribution of bank market shares—specifically, the combined loan share of the three largest banks and the five largest banks. To discipline the demand parameters $[\sigma, \eta]$, we target the estimated loan demand semi-elasticities reported in Table 9, using the indirect inference approach described in Section 4.1.

Tables 10 and 11 report the calibrated parameter values and the fit of the model to the targeted moments.

Validation. A remarkable feature of our model is that, despite its parsimonious parameterization, it matches several untargeted moments of interest. In particular, the model replicates key features of the interest rate distribution that were not directly targeted in the calibration, such as the cross-sectional standard deviation of interest rates after controlling for bank-specific default probabilities. Table 12 reports the results. In the data, the standard deviation of residualized loan rates is 2.49, compared with 2.51 in the model.

We apply the AKM decomposition (2) to the simulated data using the same estimation proce-

Table 12: AKM Decomposition: Data vs. Model

	Data	Model
Standard deviation of i_{fb}	2.49	2.51
Variance share: firm FE	0.52	0.57
Variance share: bank FE	0.17	0.24
Covariance($\hat{\alpha}_f$, $\hat{\alpha}_b$)	-0.02	0.09

Notes: The table compares the AKM decomposition of loan interest rates i_{fb} from the data (after controlling for default risk and collateral) and from simulated model data. Variances of fixed effects are reported as percent of total variance.

dure as in Section 2. We find that dispersion in bank fixed effects accounts for 24% of the overall dispersion in interest rates, very close to what we document in the data.

In the model, banks are heterogeneous in their marginal cost of funds, and the relevant state variable is the distribution of banks' net worth. While the overall number of banks is fixed, bank dynamics originate from entry and exit, which induce changes in market structure. For example, when a large bank exits, other banks gain market share and are therefore able to charge higher markups.

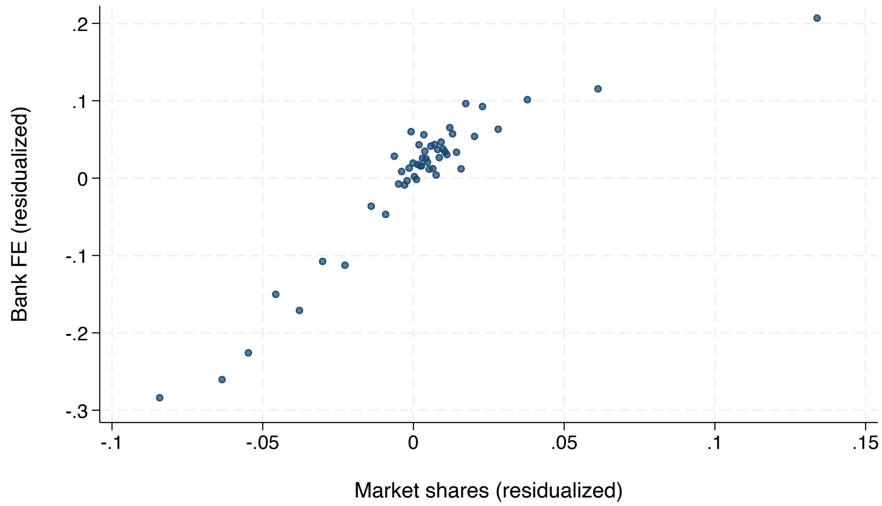
We estimate in the model how bank fixed effects change over time with changes in bank market shares, controlling for bank marginal cost of funds, bank fixed effects, and time fixed effects, as we do in Section 2. Since in the model banks are heterogeneous only in their net worth, as in Gertler and Karadi (2011), the unconditional correlation between bank fixed effects and market shares is negative in the cross-section; large banks have high market shares because their low marginal cost of funds.

Figure 6 plots this relationship between residualized bank fixed effects and market shares in the simulated model data. The positive relationship between residualized bank fixed effects and market shares is due to oligopolistic competition, since markups are increasing in market shares. Moreover, the model generates the same gradient between bank fixed effects and market shares that we document in the data (see Figure 2 in Section 2): an increase in bank market share is associated with higher markups in equilibrium.

While it is not surprising that bank fixed effects play an important role in the model given the rich heterogeneity we allow across banks, we also find that firm fixed effects account for a large share of interest rate dispersion. In particular, firm fixed effects explain 57% of the variance in interest rates in the model, which is close to the contribution of firm fixed effects documented in the data. In our model firm heterogeneity arises entirely from the set of banks with which a firm is matched. Thus, firm fixed effects are driven by the number of banks each firm is matched with and the identity of these banks.

Notably, the number of banking relationships is not the only determinant of the firm fixed effect, as there is substantial dispersion within each group. This dispersion reflects differences in

Figure 6: Bank Fixed Effect vs. Market Share in the Model



Notes: Binscatter from simulated model data. The bank fixed effect (bank-year) and market share are both residualized on the marginal cost of funds, bank fixed effects and time fixed effects.

the funding costs of the banks the firm is matched with. A firm whose banking network consists of two banks with a high marginal cost of funds may face higher borrowing costs than a firm with access to only one bank with a very low marginal cost of funds.

4.3 Macroeconomic Effects of Bank Market Power

The first counterfactual exercise aims to quantify the macroeconomic effects of bank market power. In the calibrated model, banks face downward-sloping loan demand curves and therefore charge interest rates in excess of their marginal cost of funds. These markups are heterogeneous across bank-firm pairs: a firm matched with a single large bank, or with fewer competing lenders, faces a more inelastic residual demand and thus pays a higher markup than a firm with access to several competing lenders. Heterogeneous markups across firms translate into heterogeneous effective borrowing costs i_f across firms. Because firms operate under decreasing returns to scale in production, heterogeneity in borrowing costs across firms distort the allocation of productive inputs. Specifically, a firm facing a higher average loan rate incurs a higher marginal cost of production and consequently employs fewer workers and produces less output than an equally productive firm facing a lower rate. This misallocation reduces aggregate output relative to a frictionless benchmark in which all firms face the same effective cost of credit.

To quantify these effects, we compare the baseline economy to a counterfactual in which each bank sets its loan rate equal to its marginal cost of funds, so that $i_{fb} = mc_b$ for all bank-firm pairs (b, f). An important subtlety in this exercise is that the elimination of markups does not necessarily imply lower equilibrium interest rates.[27] The main reason why lower markups might not lead to lower rates in equilibrium is that the marginal cost of funds mc_b is endogenously determined by the distribution of bank net worth in equilibrium. When banks are constrained to price at marginal cost, their profits fall. Lower profits imply a slower accumulation of net worth over time, which increases the likelihood that banks face binding leverage constraints. A tighter constraint raises the shadow cost of funds, pushing mc_b upward and partially offsetting the direct benefit of eliminating the markup. The net effect on equilibrium interest rates therefore reflects the balance between the static reduction arising from the elimination of markups and the dynamic increase arising from the endogenous deterioration of bank capitalization.

A second reason why removing markups need not lower equilibrium interest rates is the presence of leverage constraints. Even holding fixed the distribution of bank net worth, some banks may remain financially constrained, so that the quantity of loans they can extend is determined by available net worth rather than by unconstrained pricing decisions. In that case, interest rates are pinned down by the loan demand schedule at the constrained level of lending. As a result, compressing markups may have limited effects on the total quantity of credit supplied and may instead operate primarily through a reallocation of credit across firms.[28]

To trace the dynamic effects of removing markups, we design the counterfactual exercise as follows. We initialize both the baseline economy and the counterfactual at the stationary distribution of bank net worth from the calibrated model, and we impose competitive pricing—$i_{fb} = mc_b$ for all (b, f)—from period $t = 0$ onward. At $t = 0$, the two economies share an identical distribution of bank net worth, so any difference in aggregate output is driven entirely by the elimination of markups, with the financial structure of the banking sector held constant across the two economies. As time progresses, the distributions of bank net worth diverge: banks in the counterfactual economy accumulate less net worth due to lower profits, gradually tightening their leverage constraints and causing the marginal cost of funds to rise relative to the first period of the experiment.

The removal of markups affects aggregate output through two distinct channels. The first is the *level-of-loans channel*: on impact, the elimination of markups reduces loan rates, stimulating firm borrowing and employment and raising aggregate output. Over time, however, declining bank net worth raises the marginal cost of funds, partially eroding the initial rate reduction and attenuating the output gain. The second channel is *resource misallocation*: since firms pay heterogeneous effective borrowing costs i_f, they face different marginal costs of production and therefore optimally choose different input levels. To formalize this, note that aggregate output can

[27]Similarly, if a share of markups would capture rents due to asymmetric information or cost of information acquisition, removing markups might not necessarily lead to lower rates in equilibrium, as banks might not be willing to supply loans.

[28]Constrained banks may still choose how to allocate their limited lending capacity across firms, and in doing so may tilt credit toward borrowers from whom they can extract higher markups.

be decomposed as:

$$\log Y = \log N - \log \left(\int \frac{1}{N} \left(\frac{\sigma}{w[1+\gamma i_f]} \right)^{\frac{\sigma}{1-\sigma}} df \right), \tag{32}$$

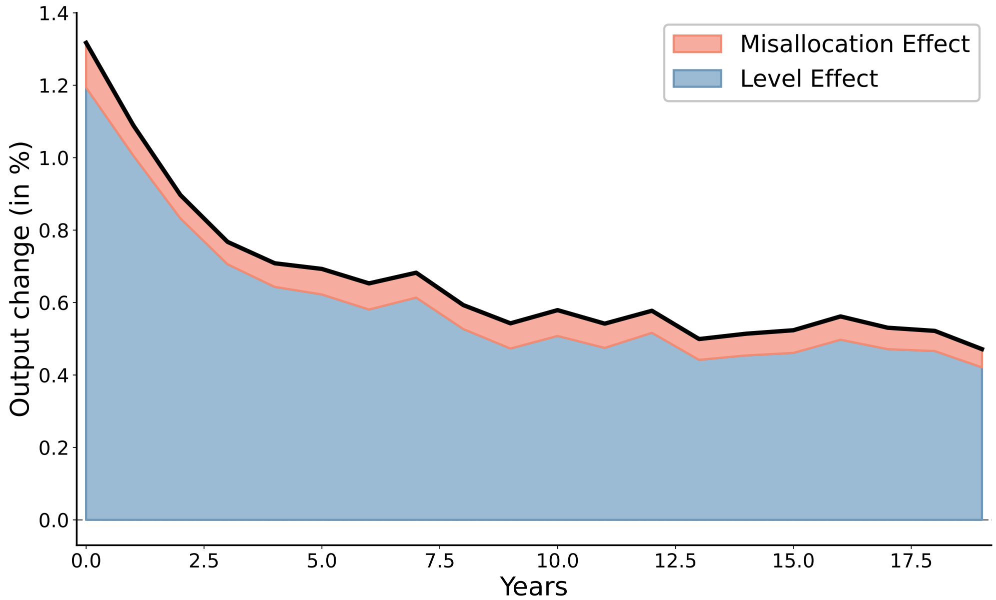
where the first term reflects the contribution of aggregate employment and the second term is a misallocation wedge arising from the cross-firm dispersion in effective borrowing costs i_f. A reduction in this dispersion—which follows from eliminating heterogeneous markups—compresses the misallocation wedge and provides an additional source of output gain beyond the level-of-loans channel. Intuitively, the dispersion in i_f across firms reflects differences in the composition of each firm's lending network: a firm matched with a single large bank that charges a high markup will face a higher average loan rate than a firm with access to multiple competing lenders.

Figure 7 plots the percentage deviation of output in the counterfactual economy relative to the baseline as a function of time, together with a decomposition into the two channels described above. On impact, forcing banks to price their loans at marginal cost raises aggregate output by 1.3 percent relative to the baseline level. This initial gain reflects the direct reduction in loan rates that follows from the elimination of markups, while the distribution of bank net worth remains identical across the two economies at $t = 0$. Over time, however, the reduction in bank profits leads to a gradual erosion of net worth in the counterfactual economy, tightening leverage constraints and causing the marginal cost of funds to rise relative to the first period of the experiment. As a consequence, the output gain diminishes over time: after twenty years, output in the counterfactual remains above the baseline, but the magnitude of the gain is substantially smaller than the effect on impact. This result highlights that an evaluation of the macroeconomic costs of bank market power based solely on static considerations or on observed markups would substantially overstate the long-run benefits of its elimination, and that accounting for the endogenous adjustment of the banking sector's capitalization is essential for a quantitative assessment.

The decomposition in Figure 7 further reveals the relative importance of misallocation channel that, while present, is quantitatively modest relative to the level-of-loans channel. The majority of the output effect—both on impact and over time—operates through the average level of borrowing costs across firms rather than through the cross-firm dispersion in rates. This finding is consistent with Faria-e-Castro et al. (2025b), who document a small role for credit misallocation in normal times and in line with Schivardi et al. (2022).

Finally, the change in output implied by making banks price loans at marginal cost should be interpreted as a counterfactual exercise that holds fixed all forces outside the model. In particular, the exercise does not speak to how lower markups would affect banks' risk-taking incentives, the riskiness of credit, or the probability of bank distress. Instead, it measures the output gains that arise from eliminating the wedge between banks' marginal cost of funds and the interest rates they charge to firms, taking the risk environment as given.

Figure 7: Decomposition of Output Effects: Average Rate vs. Misallocation



Notes: The figure plots the percentage deviation of aggregate output in the counterfactual economy—in which all banks price at marginal cost—relative to the baseline, as a function of time. We compute $2 \times M$ simulations of length T. In the first M simulations we set $i_{fb} = mc_b$. The figure plots the difference in logs between the first and second set of simulations, averaging across M. The total output effect is decomposed into two channels: the *level-of-loans channel* and the *misallocation channel*.

4.4 Interaction between Capital Requirements Policies and Bank Market Structure

Capital requirements have been a key feature of banking regulations since the introduction of the Basel Accord framework in the 1980s. They have the objective of strengthening the resilience of the banking system to aggregate shocks and limiting the procyclicality of credit. Their use has further increased since the Great Financial Crisis and the implementation of the Basel III reforms. The costs and benefits of such requirements have been studied extensively; Jiménez et al. (2017), Corbae and d'Erasmo (2021), Bonaccorsi di Patti et al. (2023), and Plosser and Santos (2023), among others, document that changes in capital requirements have quantitatively significant effects on bank lending and real economic activity.

Our model offers a new perspective on assessing the aggregate *cost* of capital requirements. We argue that policies targeting bank leverage may have substantial effects on the market structure of the banking sector—specifically, the distribution of market shares across banks—and that such changes in market structure have quantitatively important implications for the aggregate cost of borrowing. This channel operates through the link between bank size, market power, and the incidence of the leverage constraint.

Our analysis should be interpreted as a quantitative evaluation of how market structure shapes some of the costs of capital requirements, focusing on how these policies affect the cost of lending.

Quantitatively, the model-implied effects of changes in capital requirements on interest rates are similar to those estimated by Bonaccorsi di Patti et al. (2023). The main contribution of our analysis is to show that a large share of these effects is driven by changes in bank market shares and markups. Importantly, the analysis is not meant to provide a comprehensive assessment of capital requirements, since the model does not capture their potential benefits in reducing systemic bank risk or the additional complexities due to the relationship between borrowers' defaults and bank fragility.

To study these effects quantitatively, we consider a counterfactual economy in which the minimum capital ratio θ takes a lower value. We compare the baseline economy—in which $\theta = 0.08$—to a counterfactual with a more relaxed requirement, $\theta = 0.04$. As in Section 4.3, we implement the change at $t = 0$ in a counterfactual economy initialized at the stationary distribution of bank net worth from the calibrated baseline, and we compare the subsequent evolution of the two economies over time.

Before turning to the quantitative results, it is useful to describe the main mechanisms. A relaxation of the capital requirement reduces the shadow cost of the leverage constraint, allowing banks to expand their balance sheets by issuing more deposits and extending more loans.[29] The aggregate supply of credit therefore expands. However, the response is heterogeneous across banks. The largest banks are unconstrained in equilibrium: their accumulated net worth exceeds the regulatory minimum by a sufficient margin that their leverage constraint does not bind. A relaxation of the capital requirement therefore does not directly expand their lending capacity. By contrast, smaller banks whose leverage constraints are binding respond directly to the relaxation: for these banks, the total volume of loans is determined by the supply constraint rather than by the profit-maximizing choice along the demand curve, so that an expansion of the constraint translates into an expansion of lending.[30] The direct effect of a change in capital requirements is therefore a reallocation of market share from large banks—which are unconstrained—to small banks—which were previously constrained—and a reduction in aggregate market concentration.

The indirect effect of a change in capital requirements arises from the reallocation of market shares described above. As small banks gain market share and concentration falls, the average markup in the economy also falls, because the markups of large banks decline with their market shares. Thus, while a change in capital requirements does not directly affect the interest rates charged by large unconstrained banks, it indirectly affects their rates through changes in market concentration.

Table 13 reports the quantitative effects of this experiment. The aggregate effects are consistent with the mechanism described above: the relaxation of the capital requirement allows small banks

[29]The present analysis focuses on the costs of regulatory constraints and abstracts from their potential benefits in terms of limiting systemic risk and ensuring bank solvency.

[30]The allocation of loans across firms by a given constrained bank is nonetheless shaped by the firm-specific loan demand schedule and, in particular, by the firm-specific loan demand elasticity.

Table 13: Effects of Reducing Capital Requirement from 8% to 4%

	1 year	5 years	20 years
Average rate	−0.43 pp	−0.77 pp	−0.82 pp
of which: cost of funds	−0.21 pp	−0.41 pp	−0.45 pp
of which: markups	−0.22 pp	−0.36 pp	−0.37 pp
Market share of three largest banks	−7 pp	−10 pp	−10 pp
Market share of five largest banks	−5 pp	−8 pp	−8 pp

Notes: The table reports the effects of reducing the minimum capital ratio θ from 8 to 4 percent. All entries are deviations from the baseline equilibrium. "Average rate" refers to the average loan rate across all bank-firm pairs. "Cost of funds" and "Markups" decompose the change in the average rate into contributions from changes in marginal funding costs and changes in markups, respectively. Bank concentration is measured by the combined loan market share of the three and five largest banks, respectively.

to expand their lending, which stimulates firm borrowing, employment, and output, while the increase in credit supply exerts downward pressure on loan rates.

The most notable finding concerns the decomposition of the change in the average loan rate. Decomposing the average loan rate as the sum of the average marginal cost of funds and the average markup,

$$\bar{i} = \overline{mc} + \bar{\mu}, \tag{33}$$

the table reports the contribution of each component separately. Roughly half of the total decline in the average loan rate is attributable to a reduction in the average markup—that is, to a reduction in bank market power—rather than to a reduction in the marginal cost of funds. The mechanism operates through the change in market structure described above: the relaxation of capital requirements disproportionately expands the lending capacity of small banks. As small banks gain market share and concentration falls, the average markup in the economy also falls. This reallocation is confirmed by the last two rows of Table 13: the combined market share of the three largest banks falls by 7 percentage points at the one-year horizon and by 10 percentage points in the long run, reflecting a persistent fall in concentration of the banking sector. Since banks are non-atomistic and their market power is directly related to their market share, a less concentrated banking sector charges lower markups and thereby reduces loan rates.

Table 13 also reveals that the decline in interest rates is more pronounced at longer horizons: the average interest rate falls by 43 basis points after one year and by 83 basis points after 20 years. This pattern reflects the gradual adjustment of the distribution of bank net worth in response to the regulatory change. At $t = 0$, all banks begin from the same distribution of net worth as in the baseline, so the initial effect is driven entirely by the direct relaxation of binding constraints for small banks. Over time, however, the distribution of net worth adjusts: small banks that were previously constrained now accumulate net worth at a faster pace and progressively gain market share from large banks. This reallocation of market share from large banks—which charge high markups—toward small banks—which charge lower markups—amplifies the competitive channel over time and explains why both the output gains and the decline in average markups are larger at longer horizons than on impact.

5 Conclusion

This paper studies bank market power in local loan markets and its macroeconomic consequences. We propose a measurement strategy based on an AKM decomposition of loan interest rates from administrative credit registers and use the resulting bank fixed effects as a structural measure of deviations from the law of one price. Applied to Italian credit register data, we find that bank fixed effects explain approximately 15 percent of residual interest-rate variation—roughly half the contribution of default risk and collateral—and are strongly correlated with local market shares even after controlling for bank funding costs.

To assess the aggregate implications, we develop a dynamic general equilibrium model with oligopolistic competition, random matching between firms and banks, and leverage constraints. The model fits the empirical distribution of bank-firm relationships, loan demand elasticities, and the pattern of interest rate dispersion observed in the credit register. The model replicates the qualitative structure of the AKM regression in the data despite being an untargeted moment.

Our counterfactual experiments produce two main findings. The first counterfactual analysis in which lending markups are compressed to zero raises output by 1.3 percent in the short run, with the effect shrinking to 0.4 percent in the long run, reflecting the role of bank profits in sustaining capitalization and the resulting trade-off between short-run gains from lower markups and weaker long-run balance-sheet capacity. Accounting for credit risk and systemic bank risk may further dampen the long-run output gain. Our second counterfactual exercise focuses instead on the interaction between policies impacting bank capital requirements and the market structure: a relaxation of capital constraints allows banks closer to their regulatory constraint to expand, reallocating market shares and compressing markups, which amplifies the decline in rates. While the analysis abstracts from the potential benefits of capital requirements in reducing systemic risk and from borrower-default/bank-fragility feedbacks, it speaks to the lending-cost channel of capital requirements and contributes to the debate by isolating the forces arising from the interaction between capital regulation and bank market structure.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

Aguirregabiria, V., R. Clark, and H. Wang (2025): "The Geographic Flow of Bank Funding and Access to Credit: Branch Networks, Synergies, and Local Competition," *American Economic Review*, 115, 1818–56.

Altavilla, C., M. Boucinha, and P. Bouscasse (2022): "Supply or Demand: What Drives Fluctuations in the Bank Loan Market?" Tech. Rep. 2646, European Central Bank.

Altavilla, C., R. S. Gürkaynak, and R. Quaedvlieg (2024): "Macro and micro of external finance premium and monetary policy transmission," *Journal of Monetary Economics*, 147, 103634.

Amiti, M., O. Itskhoki, and J. Konings (2016): "International Shocks and Domestic Prices: How Large Are Strategic Complementarities?" Working Paper 22119, National Bureau of Economic Research.

Amiti, M., A. K. Kashyap, A. Kovner, and D. E. Weinstein (2026): "Why Do Firms Pay Different Interest Rates on Their Bank Loans?" Working paper.

Amiti, M. and D. E. Weinstein (2018): "How Much Do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data," *Journal of Political Economy*, 126, 525–587.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008): "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society Series A*, 171, 673–697.

Atkeson, A. and A. Burstein (2008): "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 98, 1998–2031.

Baqaee, D. R. and E. Farhi (2020): "Productivity and misallocation in general equilibrium," *The quarterly journal of economics*, 135, 105–163.

Begenau, J. and E. Stafford (2018): "Do Banks Have an Edge?" Harvard Business School Working Paper No. 18-060.

Benetton, M., A. Gavazza, and P. Surico (2025): "Mortgage Pricing and Monetary Policy," *American Economic Review*, 115, 823–863.

Beraldi, F. (2025): "Banking Relationships and Loan Pricing Disconnect," Working paper, Working Paper.

Berger, D., K. Herkenhoff, and S. Mongey (2022): "Labor Market Power," *American Economic Review*, 112, 1147–93.

Bocola, L. and G. Bornstein (2023): "The Macroeconomics of Trade Credit," Working Paper 31026, National Bureau of Economic Research.

Bonaccorsi di Patti, E., M. Moscatelli, and S. Pietrosanti (2023): "The Impact of Bank Regulation on the Cost of Credit: Evidence from a Discontinuity in Capital Requirements," *Journal of Financial Intermediation*, 55, 101040.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2023): "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics*, 41, 291–322.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87, 699–739.

Bordeu, O., G. González, and M. Sorá (2026): "Bank Branches and the Allocation of Capital across Cities," Working paper, Working Paper.

Cavalcanti, T. V. d. V., J. P. Kaboski, B. S. Martins, and C. Santos (2023): "Dispersion in Financing Costs and Development," NBER Working Paper No. 28635.

Corbae, D. and P. d'Erasmo (2021): "Capital Buffers in a Quantitative Model of Banking Industry Dynamics," *Econometrica*, 89, 2975–3023.

Crawford, G. S., N. Pavanini, and F. Schivardi (2018): "Asymmetric Information and Imperfect Competition in Lending Markets," *American Economic Review*, 108, 1659–1701.

De Bonis, R., G. Marinelli, and F. Vercelli (2014): "The Long Journey of Bank Competition: New Evidence on Italy from 1890 to 2013," Banca d'Italia conference paper.

De Loecker, J., J. Eeckhout, and S. Mongey (2021): "Quantifying Market Power and Business Dynamism in the Macroeconomy," Working Paper 28761, National Bureau of Economic Research.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 135, 561–644.

Dempsey, K. (2025): "Capital Requirements with Non-Bank Finance*," *The Review of Economic Studies*, rdaf061.

Dempsey, K. and M. Faria-e-Castro (2025): “A Quantitative Analysis of Bank Lending Relationships,” *Journal of Financial Economics*, 170, 104083.

Drechsler, I., A. Savov, and P. Schnabl (2017): “The Deposits Channel of Monetary Policy,” *Quarterly Journal of Economics*, 132, 1819–1876.

Edmond, C., V. Midrigan, and D. Y. Xu (2023): “How costly are markups?” *Journal of Political Economy*, 131, 1619–1675.

European Central Bank (2021): “Bank Mergers and Acquisitions in the Euro Area: Drivers and Implications for Bank Performance,” Special feature in the ECB Financial Stability Review, November 2021.

Faria-e-Castro, M., S. Jordan-Wood, and J. Kozlowski (2025a): “An Empirical Analysis of the Cost of Borrowing,” Working paper.

Faria-e-Castro, M., J. Kozlowski, and J. Majerovitz (2025b): “The Cost of Capital and Misallocation in the United States,” Federal Reserve Bank of St. Louis Working Paper No. 2025-013.

Ferreira, T. and N. Kozeniauskas (2025): “Beyond Risk: Firm Financing and Interest Rates,” Working paper.

Gertler, M. and P. Karadi (2011): “A Model of Unconventional Monetary Policy,” *Journal of Monetary Economics*, 58, 17–34.

Gertler, M. and N. Kiyotaki (2010): “Financial Intermediation and Credit Policy in Business Cycle Analysis,” in *Handbook of Monetary Economics*, ed. by B. M. Friedman and M. Woodford, Elsevier, vol. 3, chap. 11, 547–599.

Gobbi, G. and E. Sette (2014): “Do Firms Benefit from Concentrating their Borrowing? Evidence from the Great Recession,” *Review of Finance*, 18, 527–560.

Guiso, L., L. Pistaferri, and F. Schivardi (2012): “Credit within the Firm,” *The Review of Economic Studies*, 80, 211–247.

Hsieh, C.-T. and P. J. Klenow (2009): “Misallocation and manufacturing TFP in China and India,” *The Quarterly journal of economics*, 124, 1403–1448.

Ivashina, V., L. Laeven, and E. Moral-Benito (2022): “Loan types and the bank lending channel,” *Journal of Monetary Economics*, 126, 171–187.

Jamilov, R. and T. Monacelli (2025): “Bewley Banks,” *Review of Economic Studies*, advance article online. DOI: 10.1093/restud/rdaf062.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina (2017): "Macroprudential Policy, Countercyclical Bank Capital Buffers, and Credit Supply: Evidence from the Spanish Dynamic Provisioning Experiments," *Journal of Political Economy*, 125, pp. 2126–2177.

Kline, P. (2024): "Firm Wage Effects," NBER Working Paper No. 33084. Prepared for the *Handbook of Labor Economics*, Vol. 5.

Kosekova, K., A. Maddaloni, M. Papoutsi, and F. Schivardi (2025): "Firm-Bank Relationships: A Cross-Country Comparison," *The Review of Corporate Finance Studies*, cfaf004.

Maingi, Q. (2026): "Regional Banks, Aggregate Effects," *Journal of Financial Economics*, 176, 104226.

Morelli, J. M. and K. Herkenhoff (2024): "Welfare Costs of Credit Card Oligopoly," *Available at SSRN 4859945*.

Petersen, M. A. and R. G. Rajan (1995): "The Effect of Credit Market Competition on Lending Relationships," *The Quarterly Journal of Economics*, 110, 407–443.

Plosser, M. C. and J. A. C. Santos (2023): "The Cost of Bank Regulatory Capital," *The Review of Financial Studies*, 37, 685–726.

Schivardi, F., E. Sette, and G. Tabellini (2022): "Credit Misallocation During the European Financial Crisis," *The Economic Journal*, 132, 391–423.

Sharpe, S. A. (1990): "Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships," *The Journal of Finance*, 45, 1069–1087.

Ulate, M. (2021): "Going Negative at the Zero Lower Bound: The Effects of Negative Nominal Interest Rates," *American Economic Review*, 111, 1–40.

Villa, A. T. (2025): "Macro Shocks and Firm Dynamics with Oligopolistic Financial Intermediaries," Federal Reserve Bank of Chicago Working Paper No. 2022-41 (revised 2025).

Wang, Y., T. M. Whited, Y. Wu, and K. Xiao (2022): "Bank Market Power and Monetary Policy Transmission: Evidence from a Structural Estimation," *The Journal of Finance*, 77, 2093–2141.

Weber, A. (2017): "Bank Consolidation, Efficiency, and Profitability in Italy," IMF Working Paper 17/175, International Monetary Fund.

Yeh, C., C. Macaluso, and B. Hershbein (2022): "Monopsony in the US Labor Market," *American Economic Review*, 112, 2099–2138.