Bank Loan Markups and Adverse Selection*

Mehdi Beyhaghi† Cesare Fracassi‡ Gregory Weitzner§

June 2022

Abstract

We analyze market power in local US corporate loan markets. To identify market power, we create a measure of markup that incorporates banks' internal loan-level risk assessments and is orthogonal to the subsequent performance of loans. In contrast to typical theories of competition, we find that markups are higher in regions in which more banks operate. We provide evidence that this result is driven by asymmetric information across banks, which becomes exacerbated as the number of banks increase. We also provide causal support for the adverse selection channel by showing that markups drop following a shock to large banks' lending capacities that reduces asymmetric information in local lending markets. Our findings suggest that adverse selection is an important driver of market power in local bank markets and have implications for antitrust policy.

*We thank Andres Almazan, Aydoğan Altı, Adolfo De Motta, Jeff Gerlach, Christoph Herpfer, Cooper Howes, Brittany Lewis, Elena Loutskina, David Martinez-Miera, Atanas Mihov, Farzard Saidi, Philipp Schnabl, David Schumacher, Denis Sosyura, Sheridan Titman, Zhanbing Xiao, Anthony Lee Zhang and seminar participants at the AFBC Conference, Arizona State, FDIC Bank Research Conference, McGill, MFA Annual Meeting, SFS Cavalcade, UT Austin for the helpful comments and discussions. The views expressed in this this article are solely those of the authors. They do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

†Federal Reserve Bank of Richmond. Email: mehdi.beyhaghi@rich.frb.org.

‡University of Texas at Austin. Email: cesare.fracassi@mccombs.utexas.edu.

§McGill University. Email: gregory.weitzner@mcgill.ca.

1 Introduction

Market concentration has risen significantly across most industries in the United States over the last decades (Grullon, Larkin, and Michaely (2019)). In the banking sector, the top 10 largest bank holding companies have doubled their market share of total assets since 1990 (Fernholz and Koch (2020)). Traditional models of product market competition predict that higher levels of market concentration lead to higher prices and reduced supply. In fact, the antitrust division of the US Department of Justice and other regulatory agencies commonly use various measures of market concentration as important criterion to approve or block mergers, arguing that higher market concentration increases market power, thereby harming consumers with excessively high prices.

However, unlike typical product markets, credit markets are plagued by two levels of asymmetric information: i) borrowers are often better informed than lenders about their own creditworthiness, and ii) some lenders know more about certain borrowers' quality than other lenders. Perhaps surprisingly, this latter form of asymmetric information can create a positive relationship between the number of banks and interest rates in local banking markets.[1] In other words, in regions in which *more* banks operate, banks' market power can *increase*. Intuitively, when there are more banks in a market, individual banks have a more difficult time distinguishing which borrowers have been previously rejected by other banks forcing them to charge higher rates because of adverse selection. In turn, banks that are better informed about certain borrowers are able to charge those borrowers higher interest rates because their outside options are worse.

Despite the importance of understanding the sources of market power in credit markets, disentangling these forces is challenging because the adverse selection channel is driven by banks' private information which is unobservable to the econometrician. For instance, if we were to compare two loans with similar characteristics but different interest rates, we cannot determine whether the loans' interest rates are different due to differences in market power via higher markups or simply differences in risk. In this

[1] E.g., Broecker (1990), Riordan (1995), Shaffer (1998), Dell'Ariccia (2001), Marquez (2002) and Dell'Ariccia and Marquez (2006).

paper, we address this challenge by proposing a novel measure of loan markup that incorporates banks' private information regarding the riskiness of their loans. Consistent with the adverse selection channel of market power, we find a positive relationship between the number of banks in local markets and loan markups. We also find i) banks with an informational advantage due to prior relationships with borrowers charge higher markups and ii) markups drop following a shock that reduces asymmetric information differentially across local markets and borrowers. Together out results provide support for adverse selection being an important driver of market power in local loan markets.

Our analysis uses Federal Reserve's Y-14Q Schedule H.1 data that includes all corporate loans over one million dollars extended by large bank holding companies (BHCs) in the United States. A key advantage of the data is that BHCs are required to report their internal measures of probability of default (PD) and loss given default (LGD) for each loan on their balance sheets. We first show that these risk assessment measures (PD and LGD) are strong predictors of future loan delinquency and default, even after controlling for other determinants of firm performance. We estimate markup as the unexplained residual of a regression of loan interest rate on banks' private risk assessment measures and loan-level controls. We also include bank by quarter fixed effects as controls to account for differences across banks in their internal risk assessments and costs of capital at each point in time. The validity of our estimate of markup relies on the following assumption: after controlling for loan characteristics and banks' loan risk assessments, the variation in interest rates on loans given by the same bank in the same quarter is not driven by the ex-ante riskiness of that loan. We provide support for this assumption by showing that our risk-adjusted estimate of markup does not predict ex-post non-performance or default. Interestingly, when we exclude these risk measures from the model, markup strongly predicts future loan performance, making inferences on market power confounded by the underlying risk of the loan. This result shows that it is critical to control for banks' private information when estimating loan markups.

After having established the validity of our risk-adjusted markup, we test the relationship between local market structure and market power. Consistent with the adverse selection channel, we find that markups are higher in counties with more banks. Specifically, a one standard deviation increase in the number of banks operating in the county (5.5 banks) is associated with a 5.5bp increase in markup, which compares to an average credit spread of about 200bps. Also consistent with adverse selection driving market power, we find that markups are larger for firms likely facing a high degree of asymmetric information (smaller, low profitability, highly-levered and low tangibility firms). While the relationship between the number of banks and markups is consistent with the adverse selection channel, this effect should not be present if there is a single bank operating in the county.[2] Moreoever, a single bank should be able to charge a monopoly price if it indeed has no competition. Consistent with this hypothesis, we find that markups actually decrease from one to two banks, but increase thereafter.

We next develop additional tests based on the predictions of the adverse selection channel. A common theme in this class of models is that a bank's superior information about a particular borrower increases their market power because of adverse selection.[3] In practice, banks that have existing relationships with firms are likely to have better information than other banks (e.g., Greenbaum, Kanatas, and Venezia (1989), Sharpe (1990) and Rajan (1992)). Hence, we test whether banks charge their existing clients higher markups than firms that switch banks. Consistent with this channel, we find that within county and quarter-of-origination, firms that stay with their banks face 13bps higher markups on their loans.

An alternative story put forth by Petersen and Rajan (1995) is that in more concentrated regions, banks are better able to extract future rents from borrowers which incentivizes them to provide credit to lower quality borrowers. We actually find that conditional on observable characteristics, banks estimated PDs are *higher* in counties with *more* banks. This result is consistent with the adverse selection models of Broecker (1990) and Marquez (2002) who show that the average quality of borrowers goes down when there are more banks in a region.[4]

[2]There cannot be asymmetric information across banks if there is only one bank.

[3]E.g., Marquez (2002) and Dell'Ariccia and Marquez (2006).

[4]These results are also consistent with the empirical findings of Shaffer (1998) and Bofondi and Gobbi (2006).

While the above evidence is consistent with the adverse selection channel of market power, ideally we would like to have a shock that differentially affects the level of asymmetric information across counties and borrowers. To address this issue, we use capital surcharges that were imposed on global systemically important banks (GSIBs) in 2016. Favara, Ivanov, and Rezende (2021) show that following the imposition of these capital surcharges, GSIBs reduced their lending relative to other banks. We first confirm the results of Favara, Ivanov, and Rezende (2021) at the local level by showing that GSIBs' county-level market shares drop following the imposition of the surcharges. We argue that this forced reduction in lending reduces the adverse selection problem in local markets. Intuitively, it becomes less of a bad signal if a particular firm does not receive a loan from a GSIB because the bank may have denied the firm credit (or not even considered the firm at all) because it was forced to cut back it's lending, not because it deemed them a lower quality borrower. This effect reduces the adverse selection problem in the local market, thereby reducing the market power of banks. Hence we would expect that markups should drop in regions with a higher pre-surcharge GSIB presence and particularly among borrowers with existing loans from GSIBs. Consistent with this hypothesis, we show that a higher initial aggregate market share of GSIBs in a county leads to a larger drop in markups following the imposition of the capital charges. Moreoever, the drop in markups are concentrated among loans to borrowers with existing loans from GSIBs. These results are consistent with our hypothesis that the adverse selection problem is mitigated as GSIBs are forced to cut back their lending after the imposition of capital surcharges.

Overall, our results speak to the importance of adverse selection driving market power in local loan markets. Furthermore, they suggest a more subtle approach to antitrust policy. The FDIC, DOJ and Federal Reserve predominantly argue that higher concentration make markets less competitive when deciding whether to approve bank mergers. While this approach is certainly reasonable when considering the effect of mergers on deposit market competition, we argue that regulators should be more careful when applying this approach to loan market competitiveness given the importance of adverse selection.[5]

This paper contributes in several ways to the literature on market power and asymmetric information in banking markets. First, to our knowledge this is the first paper documenting a positive relationship between the number of banks and markups in local corporate loan markets. This result is in stark contrast to the large existing literature analyzing the effect of market concentration on loan prices, which find either a positive or no relationship between loan interest rates and market concentration (e.g., Hannan (1991), Petersen and Rajan (1995), Cyrnak and Hannan (1999), Sapienza (2002), Cavalluzzo, Cavalluzzo, and Wolken (2002), Berger, Rosen, and Udell (2007) and Rice and Strahan (2010)).[6] There are a few key differences between our analysis and the existing literature. First, the literature predominantly focuses on much small business loans (Survey of Small Business Finance), while our sample only includes corporate loans of at least one million dollars in size. Second, the vast majority of papers use measures of deposit market concentration rather than measures of loan market concentration due to a lack of data.[7] A problem with this approach is that deposit concentration may not line up with loan market concentration. Indeed we do not find a statistically or economically significant relationship between deposit HHIs and loan markups. Third, our data includes extensive loan and firm level characteristics not generally available in other datasets. Fourth, and most importantly, to our knowledge this is the first paper to systematically measure markups using banks' internal risk measures, thereby controlling for banks' private information regarding the loan. We show that this information is critical to properly measure the effect of bank concentration on markups because without it, our

[5]Deposits are typically insured and because of their low information sensitivity are likely much less subject to asymmetric information.

[6]Berger et al. (2004) and Degryse and Ongena (2008) survey the literature. A related literature analyzing deposits, generally finds a positive relationship between bank concentration and deposit rates (e.g., Hannan (1991), Neumark and Sharpe (1992), Prager and Hannan (1998) Driscoll and Judson (2013), Drechsler, Savov, and Schnabl (2017), Wang et al. (2018)).

[7]An exception is Sapienza (2002) in the context of Italian small-business loans.

estimate of markup strongly predicts loan performance.[8]

Second, we contribute to the empirical literature testing the relationship between bank market structure and asymmetric information. Zarutskie et al. (2003) find that firms are more likely to get loans in more concentrated regions. Cetorelli and Gambera (2001) and di Patti and Dell'Ariccia (2004) find that higher concentration is associated with higher growth in sectors that are highly dependent on external finance. Consistent with the predictions of Dell'Ariccia and Marquez (2006), Dell'Ariccia, Igan, and Laeven (2012) find that that denial rates on mortgages are lower in areas that experience faster credit demand growth and that lenders attached less weight to applicants' loan-to-income ratios. Crawford, Pavanini, and Schivardi (2018) use a structural model to analyze the effect of asymmetric information between individual banks and borrowers and find evidence that market power moderates the effect of adverse selection. Yannelis and Zhang (2021) find a positive relationship between market concentration and interest rates in the auto loan industry, but do not analyze markups. In contrast to the aforementioned papers, we directly use banks' private information through their risk assessments. Moreover, we arguably show how this private information can be a source of market power in local banking markets.[9]

Finally, to our knowledge, this is the first paper to estimate loan markups by controlling for underlying the risk of the loan.[10] Using banks' private risk assessments to

[8]Other papers using different arguments to cast doubt on concentration as a proxy for competitiveness in banking markets are (Berger (1995), Rhoades (1995), Hannan (1997), Claessens and Laeven (2005), Carbo-Valverde, Rodriguez-Fernandez, and Udell (2009)). Although we focus on the effect of market power on interest rates, other papers highlight benefits of increased competition unrelated to the interest rates of loans (e.g., Jayaratne and Strahan (1996), Cetorelli (2002), Bertrand, Schoar, and Thesmar (2007), Liebersohn (2017) and Saidi and Streitz (2018)). Increased competition can also reduce efficiency if it causes banks charter values to decrease, thereby inducing an increase in risk taking (Keeley (1990)).

[9]Our paper also contributes to the empirical literature testing information hold up problems between borrowers and banks (e.g., Santos and Winton (2008), Hale and Santos (2009), Schenone (2010) and Ioannidou and Ongena (2010)). These papers generally argue that around event studies (e.g., IPOs and new bond offerings) interest rate changes are due to changes in market power and not changes in risk. In contrast, we directly measure the magnitude of the information rents banks extract from borrowers by controlling for the risk of the loan.

[10]Our paper also relates to the literature analyzing bank internal risk-measures (e.g., Jiménez and Saurina (2004) Agarwal and Hauswald (2010), Qian, Strahan, and Yang (2015), Behn, Haselmann, and Vig (2016), Dell'Ariccia, Laeven, and Suarez (2017), Plosser and Santos (2018), Becker, Bos, and Roszbach (2018)), Adelino, Ivanov, and Smolyansky (2019)). Adelino, Ivanov, and Smolyansky (2019) also use Y-14Q data and show that interest rates have minimal predictive power on loan performance after controlling for PDs, which is consistent with our analysis. However, they do not explore the variation in interest rates that is unexplained by risk, which is the main focus of our paper.

estimate markup is critical because, as we show, observable characteristics do not fully account for the underlying risk of loans. Furthermore, we show that loan riskiness varies across regions, even after controlling for observable characteristics. This makes it extremely challenging to estimate the relationship between market structure and market power without access to banks' private information. Hence, our methodology could useful to both regulators and researchers who have access to the Y-14Q data to better understand market power in corporate loan markets.

2 Theoretical Background

In homogeneous product markets, theory typically predicts that fewer competitors (or higher concentration) leads to higher prices. For example, in static Cournot models, where firms compete via quantities, firms better internalize the impact that their production has on prices when there are fewer firms, leading to higher markups and firm profits.[11] In dynamic settings, higher concentration facilitates collusion, which can also lead to higher markups (e.g., Stigler (1964), Friedman (1971) and Abreu (1986)).

However, in credit markets plagued by asymmetric information, the relationship between market concentration and markups can flip due to adverse selection. There are two main forms of asymmetric information in credit markets. First, borrowers are often better informed about their own creditworthiness than the banks that lend to them. Second, some banks might know more about certain borrowers than other banks, either because they are better at screening, or because they have access to private information through ongoing relationships with their existing clients. The latter form of asymmetric information can allow informed banks to charge high quality borrowers prices higher than their marginal cost because those borrowers would be pooled with lower quality borrowers if they tried to find a more competitive price from another bank. Higher concentration can actually limit this form of adverse selection. Intuitively, as the number of banks decreases, the winner's curse problem becomes less severe (e.g., Broecker (1990), Riordan

[11] In contrast, if firms compete via prices à la Bertrand in a static setting, prices are competitive as soon as there are multiple firms.

(1995) and Shaffer (1998)) or information becomes less dispersed (e.g., Marquez (2002) and Dell'Ariccia and Marquez (2006)), as banks have an easier time determining whether borrowers have been been rejected by other banks.[12] A natural consequence of adverse selection, as stated by Marquez (2002), is:

> "..focusing exclusively on the number of banks may not provide a very good indicator of the competitiveness of a market... markets composed of many small banks may actually have higher expected interest rates in equilibrium than markets composed of a few large banks."

Furthermore, this form of adverse selection anecdotally appears to an issue in practice for banks. For instance, Shaffer (1998) states:

> "The chief financial officer of a new bank once told the author that "as soon as you open your doors, every deadbeat in town lines up to try to borrow from you" and that the only solution to this problem was to hire "superior" loan officers. Bankers and bank examiners alike are very familiar with this phenomenon."

In summary, two opposing forces can affect the relationship between market concentration and markups in credit markets: on the one hand, higher market concentration can increase banks' market power and loan markups, while on the other hand, higher market concentration can reduce banks' market power and loan markups by alleviating adverse selection concerns.

As mentioned above, the adverse selection channel of market power implies that banks who are better informed about certain borrowers can charge higher markups due to market power. Hence, we would expect that borrowers that stay with their existing banks face higher markups than those that switch to different banks as existing banks are likely to know more about their borrowers than other banks (e.g., Greenbaum, Kanatas, and Venezia (1989), Sharpe (1990) and Rajan (1992)).

Finally, adverse selection can also affect the composition of borrowers granted credit. In Broecker (1990) and Marquez (2002) the average quality of borrowers goes down as

[12] Bolton, Santos, and Scheinkman (2016) and Fishman and Parker (2015) also show how informed investors profits can increase as more informed investors enter.

the number of banks increases.[13] Hence, the adverse selection channel would imply a positive relationship between the number of lenders and the riskiness of borrowers who receive credit. Note this is contrast to Petersen and Rajan (1995) in which banks provide more credit to lower quality borrowers in highly concentrated markets because they can earn future rents from those borrowers.

3 Empirical Design

3.1 Estimating Markup

In this section we describe our approach to identifying market power by estimating markups. Markups, which measure the extent to which banks are able to price their loans above marginal cost, are the most common proxy for market power in the banking literature a well as finance and economics more broadly (De Loecker and Warzynski (2012)).

The interest rate (IR) charged by a bank to a firm can be decomposed into three parts: (i) the marginal cost due to the credit risk that the banks faces in case the firm defaults (MC_{risk}); (ii) the marginal cost of originating, administering, and monitoring the loan ($MC_{non-risk}$), and (iii) the markup (MU), which by construction is the difference between the price and the marginal cost.

$$IR = MC_{risk} + MC_{non-risk} + MU, \tag{1}$$

In the banking literature, markups are commonly estimated by absorbing both risk and non-risk marginal costs using a host of controls and fixed effects based on bank and loan characteristics. However, the premise of the adverse selection channel, as described in Section 2, is that banks have private information that is not observed by others, including the econometrician. Any measure of markup that does not directly control for the ex-ante

[13]The mechanism in these models is slightly different. In Broecker (1990) each firm gets more chances to get accepted for a loan because it can approach more banks, while in Marquez (2002) information becomes more dispersed.

risk assessment of banks could be contaminated by unobserved risk factors, biasing the estimation of the determinants of loan markups. In fact, we show below that markup predicts ex-post loan performance and realized default if it is estimated without explicitly controlling for banks' ex-ante risk assessments.

We estimate bank loan markup using a two-stage procedure. First we estimate the following linear regression:

$$IR_l = \beta_0 PD_l + \beta_1 LGD_l + \beta_2 \left(PD_l \times LGD_l\right) + \gamma X_l + \delta_{b,t} + \alpha_{i,t} + u_l, \tag{2}$$

where the unit of observation is loan l in industry i originated by bank b in quarter t. The outcome variable IR_l is the loan's interest rate and X_l are loan-level controls, $\delta_{b,t}$ is bank by quarter fixed effects, and $\alpha_{i,t}$ is industry by quarter fixed effects. Our first-stage includes all variables that we believe may influence a loan's marginal cost. Most importantly, we include the bank's estimate of the loan's probability of default PD_l and loss given default LGD_l.

Next, we estimate markup by decomposing the interest rate into two components using the coefficient estimates from Equation (2): the predicted interest rate and the residual. The predicted interest rate represents the marginal cost (both the risk and non-risk component), while the residual is our estimate of loan markup. If the interest rate decomposition is valid, the predicted interest rate component should reflect the marginal cost of the loan, and thus predict future loan performance while the markup should not as it is orthogonal to the risk of the loan. As shown below, we will directly perform this test.[14]

It is important to note that throughout our analysis we will use bank by quarter fixed effects to control for any bank specific factors that may affect the marginal cost of a loan,

[14]One limitation of our analysis is that our measure of markup is the residual of a regression, which by construction has mean of zero. Thus our measure of markup is only relative to other loans in the sample. Hence, we cannot identify the absolute level of bank markups, only how markups vary across loans. An alternative approach to estimating markups is to use a structural model (e.g., De Loecker and Warzynski (2012)). However, this approach requires estimating a production function, which may not be well-suited for a bank loans where the main input is the information regarding the borrower. Moreover, as far as we are aware, it cannot deal with differences in marginal cost across loans due to banks' private information. Hence we view our approach as flexible enough to capture what we view as the main driver of markups in bank loans.

e.g., difference in cost of capital, regulatory costs, monitoring skills, etc. Moreoever, if banks use difference risk models, this approach should absorb such heterogeneity. Relatedly, may not report their risk measures truthfully (Begley, Purnanandam, and Zheng (2017), Plosser and Santos (2018), Behn, Haselmann, and Vig (2016)). While we cannot rule out that banks underreport their average risk measures, aggregate misreporting should not bias our results because our analysis includes bank by quarter fixed effects. Our results could also be biased if banks differentially misreport across loans; however, if this were the case we would expect our estimate of markup to predict the future performance of loans, which as we show below, it does not.

We do not include firm characteristics when we estimate markup for two reasons: first, as long as we control for the riskiness and characteristic of the loan, firm characteristics should not affect the risk-based component of the interest rate. Second, we do not want to control for variables, such as the size of the firm, that may be related to asymmetric information and thereby drive markups. In the Appendix, we show that markup predicts loan performance when we include firm characteristics but not banks' risk assessments in our estimate of markup.

It is also possible to simply use a single regression in which we control for PD and LGD and include variables that we expect to affect the interest rate but not the risk of the loan (e.g., local market concentration). However, it is critical for our analysis that we can separate out the component of the interest rate that is related to the risk of the loan and the component which is orthogonal to the risk of the loan (i.e., the markup). Without doing so, we are unable to tell which variables are affecting the underlying risk of the loan and which are affecting the markup. Moreoever, we can directly test whether the markup is orthogonal to the risk of the loan by regressing loan performance on the estimated markup.

One concern with our approach is that bank and loan characteristics could proxy not only for the marginal cost of a loan, but also for the degree of asymmetric information of the loan. For example, the adverse selection problem may be less severe for a revolving credit line than a term loan because the quality of the borrower may be less important if the loan remains undrawn for a period. By absorbing these characteristics using control variables and fixed effects, our approach will attribute all variation in the interest rate driven by that control variable to marginal cost. However, this potential mis-attribution would only attenuate our results by making our estimate of markup noisier.

A second concern is that the independent variables we include to estimate markup do not absorb all the differences in non-risk components of marginal costs, such as originating, administering, and monitoring costs. For example, monitoring costs could be higher in rural areas, where the distance between banks and borrowers is usually larger. However, as we show later we find that regions with *higher* concentration exhibit *lower* markups. Hence, if present, this effect should only attenuate our results.

3.2 Data

Our main source of data is Schedule H.1 of the Federal Reserve's Y-14Q data. The Federal Reserve began collecting this data in 2011 to support the Dodd-Frank mandated stress tests and the Comprehensive Capital Analysis and Review (CCAR).[15] The sample includes corporate loans from all bank holding companies (BHCs) with $50bn or more in total assets, accounting for 85.9% of all assets in the U.S. banking sector as of 2018:Q4 (Frame, McLemore, and Mihov (2020)). Qualified BHCs are required to report detailed quarterly loan level data on all corporate loans that exceed one million dollars in size. These loans represent 70% of all commercial and industrial loan volume in the U.S. (Bidder, Krainer, and Shapiro (2020)).

The data include detailed loan characteristics (such as interest rate, maturity, amount, collateral, credit guarantee, purpose), quarterly loan performance (past due payments, non-accruals, charge-offs), the ZIP code of the borrowers' headquarters as well as firm financials (balance sheet and income statement). Importantly for our analysis, banks are also required to report their internal estimates of probability of default (PD) and loss given default (LGD) for each loan to the Federal Reserve on their Y-14Q filings.

[15]Other papers that use Y-14Q data include: Bidder, Krainer, and Shapiro (2020), Brown, Gustafson, and Ivanov (2017), Balasubramanyan, Berger, and Koepke (2019), Ivanov, Pettit, and Whited (2020), Abdymomunov, Curti, and Mihov (2020), Beyhaghi (2022) and Greenwald, Krainer, and Paul (2020).

According to the Basel Committee on Banking Supervision, internal estimates of PD and LGD "must incorporate all relevant, material and available data, information and methods. A bank may utilize internal data and data from external sources (including pooled data)."[16]

Following Brown, Gustafson, and Ivanov (2017), we restrict the sample to domestic borrowers and remove financial firms, government entities, individual borrowers, foreign entities, and nonprofit organizations. In addition, we drop loans to special purpose entities, loans with government guarantees, demandable loans, loans with prepayment penalty clauses, loans that are tax-exempted, and loans that are contractually subordinated. We include these additional screens to make the loans in our sample as comparable as possible, thereby allowing us to accurately estimate markups. To keep focus on issues of local market power, we also drop publicly traded firms (firms with a valid ticker information) and syndicated loans because they are usually sourced nationally rather than locally. To ensure that our results are not affected by the sample of public firms with unreported ticker information, we trim the sample on borrower size at the 99th percentile.

To correct reporting errors, we drop loans with interest rates equal to or below 0% or above 100% and loans with PDs missing, zero, or greater than the 99th percentile. Interest rates are reported only in the quarter in which the borrower makes a payment on a loan, otherwise the loan's interest rate is reported as zero. For credit lines this has a material impact because firms may not draw them immediately. Hence, when the interest rate field is zero, we take the interest rate from the next quarter it is populated. Loans that are not utilized within two quarters after initiation are dropped from the sample as no interest rate is reported for these loans. As a loan might remain on the bank's balance sheet for multiple quarters, we only keep the first appearance of a loan in the data (i.e., new loans). Finally, as some firms have an abnormally large number of loans in the sample, we remove loans in which the borrowing firm has more than the 99th percentile in total new loans over the sample. After these filters, we are left with 28,033

[16] The most recent instructions are available at Calculation of RWA for credit risk.

new loans originated from 2014Q4 to 2020Q3 by 23 BHCs.[17]

We define the following firm-level financial variables: profitability (EBITDA/assets), firm size (log assets), tangibility (tangible assets/assets), and leverage (debt/assets), winsorized at the 1% and 99% level. Furthermore, we use two measures of loan performance: (i) non-performance, which is a dummy variable equal to one if the bank reports the loan as 90 days past due or non-accrual, or reports a positive net cumulative charge-off amount, or reports specific reserve for an impaired loan for the loan within the 12 months following the origination of the loan, or if the bank considers the borrower as defaulted as defined below; and (ii) default, which is a dummy variable equal to 1 if the borrower defaults within one year since origination, defined as a borrower rated D (defaulted) or is assigned a PD=100% by the lending bank within one year after the origination of the loan. We use a window of one year because banks' PD estimates are required to reflect one year default rates.

Finally, the data includes ZIP codes corresponding to each borrower's headquarters. As is typical in the banking literature, we define a local market as a county.[18] In order to create concentration measures at the county level, we obtain the ZIP code to county crosswalks from the Housing and Urban Development (HUD). After merging the county data into the Y-14Q dataset, we construct a measure of concentration based on the number of banks that operate in that county in the sample. Specifically, we consider a bank to operate in a county if it gives a loan at any point in the sample. We use the number of banks as our measure of concentration because it has a more direct relationship with theories of adverse selection.[19] However, our main results hold if we use a Herfindahl-Hirschman Indices (HHI). Moreoever, as shown in Figure B1, the correlation between the number of banks and HHI is -.89. We also collect population density data from the Census and county-level wage data from BLS to control for county characteristics.

In Section A of the Appendix we include detailed definitions of all of our variables.

[17] In the early part of the sample, PDs were not reported consistently; hence, our sample begins in 2014Q4.

[18] E.g., Drechsler, Savov, and Schnabl (2017). Our main results are also robust to using MSA as our measure of a local market.

[19] E.g., Broecker (1990), Marquez (2002) and Dell'Ariccia (2001).

3.3 Descriptive Statistics

Table 1 includes summary statistics for the variables used in the paper. The average and median loan size is approximately \$7mm and \$2.6mm, respectively and over 90% of loans are less than \$16mm. The fact that the majority of the loans and firms are relatively small is also important for testing the effect of geographic concentration on markups, as larger firms often source their loans nationally.

The loan sample is approximately evenly split among credit lines and term loans and the median interest rate is 3.66%, which corresponds to about a 200bp credit spread over the average treasury rate. The median firm has \$20mm in assets, 8% profitability, and 30% book leverage. The loans in our sample thus make up a large portion of firms' capital.

Over our sample period, 0.81% of firms default within the first year after loan origination. This compares to an average ex-ante expected PD of 1.38%. This discrepancy is likely due to the fact that the realized aggregate economic conditions in the US over the sample period were positive relative to banks' ex-ante expectations.

3.4 Validity of Bank Risk Assessments and Estimation of Markup

In order to properly estimate markup, banks' reported risk measures must reflect the actual risk of the loans. Therefore, we verify that the ex-ante banks' risk metrics predict ex-post performance (delinquency and defaults) and interest rates.

First, we compare the univariate relationship that realized default has with interest rate and PD. In Figure 1a we place loans into five equal-sized bins sorted on interest rate and plot their average realized default rate. While the overall correlation is positive, the trend is not monotonic: average default rates increase from bins 2-5 but not from bin 1-2. On the other hand, when we place loans into five PD buckets we see a much clearer positive and monotonic relationship between average realized defaults and PD than interest rates (Figure 1b). The preliminary evidence shows that bank risk measures are more strongly correlated to performance than interest rates, suggesting that interest rates may include substantial non-risk components to interest rates, namely non-risk marginal costs and markups.

Second, we formally test whether the bank risk metrics explain loan performance and interest rates after we control for loan characteristics and add a host of fixed effects. We thus formally estimate the following multivariate regression

$$y_l = \beta_0 PD_l + \beta_1 LGD_l + \beta_2 \left(PD_l \times LGD_l \right) + \gamma X_l + \delta_{b,t} + \alpha_{i,t} + u_l, \tag{3}$$

where the unit of observation is each loan l of firm f in industry i originated by bank b in quarter t. The outcome variable y_l is either Non-Performance, Realized Default or Interest Rate and X_l are loan-level controls, which include: log(Maturity), log(Amount), Guarantee and loan type fixed effects, $\delta_{b,t}$ is bank by quarter fixed effects, and $\alpha_{i,t}$ is industry by quarter fixed effects. Bank by quarter fixed effects allow us to control for any differences in internal risk models across banks or within bank over time. Furthermore, by always evaluating two loans given by the same bank in the same quarter, we absorb any differences in banks' cost of capital or financial constraints that may affect interest rates. We add an interaction term between PD and LGD to explicitly take into account the expected loss of each loan. To adjust standard errors for correlations in the residuals, we cluster the standard errors by firm.

The results are displayed in Table 2. In Columns (1) and (3), we estimate the predictive power of loan-level controls and fixed effects in explaining non-performance and realized default. These baseline regressions do not include banks' risk assessment measures. The adjusted R-squared are 8% and 5%, respectively. In Columns (2) and (4) we add PD, LGD, and PDLGD. Consistent with Figure 1b, banks' PDs strongly predict future non-performance and realized default, even after controlling for a host of loan characteristics and fixed effects. For example, a 1pp increase in PD implies an 0.63pp increase in realized default rates. The adjusted R-squared of the regressions also increase significantly to 10% and 7%. Overall, we conclude that the ex-ante bank risk assessments of loan credit risk predict ex-post loan performance, and that the information included in these measures is not fully absorbed by loan and bank characteristics. It is thus crucial to control for such variables when estimating markups.

We then turn to predicting interest rates and estimating markups in Table 3. In Column (1), we establish a baseline model where we predict interest rates using only loan-level characteristics and fixed effects. The adjusted R-squared is 49%. In Column (2), we include PD, LGD, and their interaction term (Expected Loss) to the regression. Consistent with the results in Table 2, bank risk measures also strongly predict interest rates. Even after controlling for loan, bank, and industry characteristics, loans that have greater probability of default, loss given default, and expected loss are charged higher interst rate. The adjusted R-squared increases from 49% to 52%, confirming that these bank assessments can explain a large portion of the heterogeneity in interest rates. The effect of the risk assessment on interest rates is not only statistically significant, but also economically relevant: a 10pp increase in PD leads to a 0.75pp increase in interest rate.

Third, we estimate loan markups by decomposing the interest rate into two components using the coefficient estimates from Equation (3): the predicted interest rate and the residual, which will be our proxy for loan markup. We define two measures of markup: a baseline markup, which uses the residual from the estimates in Column (1) that does not include banks' risk assessment measures, and a risk-adjusted markup, which uses the residual from the estimates in Column (2), that takes into account the banks private risk assessment. It is important to note that our measure of markup is relative, and not absolute. By construction, the markup is a residual with a mean of zero. However, as shown in Table 1, the standard deviation of markup is 0.8pp, which is very similar to the standard deviation of the predicted interest rate. In other words, half of the difference in interest rates charged by banks to firms is driven by observable factors, and half by unobservable factors.

Finally, we test the validity of the baseline and risk-adjusted markups. For the residual to be a plausible measure of markup, it should be unrelated to the future performance of the loan. We directly test the relationship between markup and loan non-performance and default in the following regression:

$$y_l = \beta_0 \widehat{IR}_l + \beta_1 \widehat{MU}_l + \gamma X_l + \delta_{b,t} + \alpha_{i,t} + u_l, \tag{4}$$

where the outcome variable, y_l, is either Non-Performance or Default, $\widehat{IR}_l$ and $\widehat{MU}_l$ are the predicted interest rate and markup estimated from Equation (3). We estimate Equation (4) with the same set of fixed effects from Equation (3) and we cluster our standard errors by firm.[20] The results, which we estimate excluding bank risk measures (baseline) and including bank risk measures (risk-adjusted) are displayed in Table 4 with non-performance as the dependent variable in Panel A and realized default as the dependent variable in Panel B. For reference, we also include regressions with the actual interest rate as an independent variable rather than its decomposition into predicted interest rate and markup.

Columns (1) and (4) of Panel A and B in Table 4 confirm the univariate results of Figure 1a, showing that a higher interest rate predicts loan non-performance and default with or without the use of fixed effects. In Columns (2) and (5) we decompose the interest rate into the predicted interest rate and markup (residual) for the baseline model without banks' risk assessment measures. In all specifications, the baseline markup still predicts loan performance, suggesting that controlling for loan and bank characteristics does not absorb drivers of loan default risk.[21] In other words, not adjusting for banks' private information leads to a biased measure of markup that correlated with loan performance, and any inference regarding market power using such a measure of markup could be confounded by the underlying risk of the loan.

In Columns (3) and (6), we repeat the interest rate decomposition using the risk-adjusted model estimated using banks' private risk assessments. Unlike the baseline markup, the risk-adjusted markup does not appear to predict ex-post default or non-

[20]Because, the predicted interest rate is estimated from another regression, we can also boostrap our standard errors; however, clustering by firm ends up with more conservative standard errors. Additionally, the predicted interest rate is not the main variable of interest in these regressions.

[21]In the Online Appendix, we show that even if we include firm characteristics in the estimate of the baseline model, the baseline markup still strongly predicts loan performance. We conduct this exercise to give the baseline markup the best chance to capture the variation in interest rates due to the riskiness of the loan without using bank risk assessments.

performance, while the predicted interest rate is positive and statistically significant. It is important to note that we cannot formally tests whether risk-adjusted markup is completely orthogonal to loan performance, as the inability to reject the null hypothesis does not mean that the null is true. However, the point estimate on the risk-adjusted markup is both economically small (less than one order of magnitude of the baseline markup) and statistically insignificant. Overall, the drastic differences we see in the baseline and risk-adjusted markup performance highlights the importance of adjusting for banks' risk assessments to have an unbiased measure of markup that is not contaminated by credit risk.

After having established the validity of the risk-adjusted markup, in the following section we analyze the relationship between market concentration, firm characteristics and markups.

3.5 Markups, Market Concentration and Firm Characteristics

In this section we test how risk-adjusted markup relates to market concentration and firm characteristics. We thus estimate the following loan-level regression:

$$\widehat{MU}_l = \beta_0 NOB_c + \gamma Z_{f,t} + \delta_{b,t} + \alpha_{i,t} + u_l, \tag{5}$$

where risk-adjusted markup ($\widehat{MU}_l$) for loan l originated by bank b to firm f in 2-digit NAICS industry i is the dependent variable measured in percentage points, NOB_c is the the number of banks in the county where the firm is headquartered, $Z_{f,t}$ is a vector of firm characteristics, and $\alpha_{i,t}$ are industry-quarter fixed effects. We also cluster the standard errors by county.

The results are displayed in Table 5. In Column (1) we estimate a univariate regression without firm characteristics. The point estimate for the number of banks is 0.01 and statistically significant. In other words, if the county has one more bank this leads to a 1bp higher markup. This is economically meaningful given that the standard deviation of the number of banks is 5.5 and the average credit spread is about 200bps. This result is consistent with the adverse selection driving market power. However, if there is only one bank operating in a county, this effect should not be present as there should be no possibility of adverse selection and the single bank can charge a monopoly price. To test this hypothesis, in Column (2) we add a dummy variable that equals one if there is one bank operating in that county. Consistent with the hypothesis, the point estimate (0.114) for one bank is positive and statistically significant. Moreoever, it is ten times as large as the coefficient for the number of banks.

We next further test the adverse selection channel by adding firm characteristics to the regression in Column (3). The coefficient on the number of banks does not change. Furthermore, firms that the literature suggests suffer less from asymmetric information problem, i.e., larger, more profitable, with low leverage and with higher tangible assets, receive loans with lower markups.[22] In Columns (4) and (5) we find that results are robust to controlling for population density and average wages in the county. Finally, in Column (6) we show the results are unchanged after controlling for bank by quarter, industry by quarter, and loan fixed effects.

The results in Table 5 suggest a non-monotonic relationship between the number of banks and markup. To further explore this relationship, we plot the coefficients for various numbers of banks operating in the county in Figure 2. The coefficients exhibit a U-shape in which markups first drop from one bank to two banks, but then steadily increase as the number of banks increase, which is in line with what theory predicts.

At first glance, it might seem puzzling that markups are higher for counties with 18-22 banks than a single bank if that single bank is operating as a monopoly. We offer two potential explanations. First, our measure of the local market as a county is somewhat coarse (Stigler and Sherwin (1985)). A firm does not necessarily need to get a loan from within the county so it may not be the case that a single bank in that region implies a pure monopoly. Moreoever, there could be competition from smaller banks that do not

[22]Firm size should not be related to transaction costs from loans because we control for loan size in our estimate of markup. Higher asset tangibility can reduce asymmetric information if the payoff of the assets are easier to observe (Almeida and Campello (2007)). Higher leverage can exacerbate the asymmetric information problem by increasing the sensitivity of a security's payoff to firm quality e.g., (Heider (2003)).

appear in our data and or shadow banks.[23] Nonetheless, we still believe it is reasonable, and our results support, that if there is one large bank operating in a county, that bank likely has a sizable amount of market power from a lack of competition from other large banks.

Relatedly, a concern with our measure of the number of banks in a county is we do not have all loans from all banks in the region because our data only covers U.S. banks with over $50bn in assets. However, as noted earlier, the Y-14Q eligible banks make up the vast majority of corporate volume. We also believe this concern is alleviated by the fact that we see the U-shape in Figure 2. First, if there were many other banks competing, we would not expect markups to be higher when there is one Y-14Q eligible bank. Second, the fact markups are monotonically increasing in then number of banks after one bank suggests our measure is capturing the relative number of banks in the region.

Another possible concern with our results is that our measure of markups might be influenced by unobserved costs of originating, processing, administering, and monitoring loans, which vary within banks, and across loans and regions. We believe by controlling for county characteristics we partially mitigate these concerns.[24] However, in Sections 3.6 and Section 3.7, we provide additional evidence on the adverse selection channel by analyzing i) cross-sectional differences in loans within county quarter and ii) time series variation in markups within counties by exploiting a shock to asymmetric information.

Next, in Table 6 we estimate the similar regressions to Table 5 but include PD as the dependent variable rather than risk-adjusted markup to test how the risk of borrowers varies across counties.[25] In Broecker (1990) and Marquez (2002), the average risk of borrowers increases in the number of banks operating in a market because of adverse

[23] For instance, Gopal and Schnabl (2020) show that FinTech lenders increased their lending to small businesses after the financial crisis.

[24] An additional concern could be that advertising is more expensive in less concentrated areas. In the context of mortgages, a more homogenous market than corporate loans, Gurun, Matvos, and Seru (2016) find that heavily advertised mortgages are more expensive. However, to our knowledge banks do not advertise corporate loans as they do consumer loans. Hatfield and Wallen (2022) provide evidence that the three largest depositary banks (Bank of America, JP Morgan and Wells Fargo) tacitly collude across multiple local markets. Our results are also robust to controlling for the total market share of these three banks and are available upon request.

[25] We also include numerous fixed effects that we include in the first stage of the markup estimate.

selection.[26] Across all specifications, increasing the number of banks in the county is associated with higher PDs. This result not only lends support to the adverse selection channel, but also highlights the importance of having access to banks' private risk assessments as the underlying risk of loans can systematically vary across different market structures *conditional on observables*.

3.6 Markups and Switching Banks

In this section we provide further evidence for the adverse selection channel. In particular, we test whether firms face higher markups when they remain with their existing banks on new loans, a direct prediction of several theories of adverse selection in banking markets.

In order to capture the information effect of repeat borrowers, we restrict the sample to firms with more than one loan and analyze all loans that follow their first loan.[27] After making these restrictions, we estimate the following regression:

$$\widehat{MU}_l = \beta_0 StayBank_l + \gamma X_{f,t} + \delta_{b,t} + \alpha_{i,t} + \lambda_{c,t} + u_l, \tag{6}$$

where $StayBank$ is a dummy that equals one when firms stay with their existing banks on their new loan, $\lambda_{c,t}$ is county by quarter-of-origination fixed effects to control for any unobserved differences in markups across regions and time. Once again, we cluster the standard errors by county. The results, which we estimate with and without our main set of fixed effects and firm characteristics are displayed in Table 7. Consistent with our hypothesis, we find that the estimated coefficient of $StayBank$ is positive and statistically significant for all regressions. For example in Column (2) when we include firm characteristics, firms face 7.4bp higher markups when they remain with their existing bank.

This result provides additional support for the adverse selection channel of markups. Next we analyze a plausibly exogenous shock to asymmetric information.

[26] This is also consistent with Yannelis and Zhang (2021) who show that banks invest more in screening technologies in concentrated markets.

[27] For this filter we also use the data from 2011 up to the beginning of our sample to determine whether a loan follows the firm's first loan.

3.7 GSIB Surcharges as a Shock to Asymmetric Information

In this section, we use the capital surcharges imposed on global systemically important banks (GSIBs) as a shock to asymmetric information in local banking markets. Favara, Ivanov, and Rezende (2021) show that as the capital surcharges were phased in in 2015, affected banks decreased their lending relative to other banks.[28] Intuitively, it becomes less of a bad signal if a particular firm does not receive a loan from a GSIB because the bank may have not granted the firm credit because it was forced to cut back it's lending, not because it deemed them a lower quality borrower. This effect raises borrowers' outside options, particularly in counties with a high initial GSIB presence and among borrowers from GSIBs.

We first confirm the results of Favara, Ivanov, and Rezende (2021) at the county-year level by estimating the following differences in differences regression:

$$MarketShare_{b,t,c} = \beta_0 + \beta_1 GSIB_b \times Post_t + \gamma_{b,c} + \delta_t + u_{b,t,c}, \tag{7}$$

where $MarketShare_{b,t,c}$ is bank b's market share in county c in year t, $GSIB_b$ is a dummy variable that equals one if bank b is affected by the capital surcharges, $Post_t$ is a dummy variable that equals one if the year is 2016 or later and $\gamma_{b,c}$ are bank by county fixed effects. In all regression in this section we cluster our standard errors by county.

The results are displayed in Column (1) of Table 8. Consistent with GSIBs cutting back their lending relative to non-GSIBs, the coefficient on the interaction is negative and statistically significant. GSIBs' market shares drop by 2.2pp which is almost one third of their initial value. We also estimate a year-by-year version of (7) and plot the time series of the coefficients in Figure 3, which shows GSIB market shares drop beginning in 2016 and steadily fall thereafter.

Next, we test whether counties with larger initial GSIB shares experience reductions

[28] The eight US banks that are identified as GSIBs in our sample are: Bank of America Corporation, The Bank of New York Mellon Corporation, Citigroup, Inc., The Goldman Sachs Group, Inc., JPMorgan Chase & Co., Morgan Stanley, State Street Corporation, and Wells Fargo & Company. We thank Ivan Ivanov for sharing the GSIB data.

in markups following the imposition of the surcharges. We create a variable GSIB Market Share which is the sum of GSIBs' market shares in 2015 prior to surcharges being imposed. We then estimate the following regression

$$\widehat{MU}_l = \beta_0 \beta(GSIB\ Market\ Share_c \times Post_{t,c}) + \gamma Z_{f,t} + \gamma_{b,c} + \delta_t + u_l, \tag{8}$$

The coefficient of interest is $GSIB\ Market\ Share_c \times Post_{t,c}$ which can be interpreted as the effect of GSIBs' initial aggregate market shares on change in markups following the imposition of the capital charges. The estimate, which is displayed in Column (2) of Table 8, is negative and statistically significant. Specifically a one standard deviation increase in GSIB Market Share (≈30pp), leads to a 3bp decrease in markups after the imposition of the capital surcharges. Although the coefficient estimate in Column (2) may seem small, In Column (3) we estimate (8) but interact the difference in difference coefficient with a Stay Bank dummy. The triple interaction is negative and statistically significant and three times the magnitude as the baseline interaction. This result implies that the drop in markups is being driven by firms that stay with their existing banks, which is exactly what we would expect. We also plot the time series of annual coefficients in Figure 4, which shows a downward drop in markups beginning in 2016.

Next, we test whether GSIBs in particular reduce the markups on their loans following the imposition of the surcharges. To do so we estimate the same regression as (7) but we include risk-adjusted markup as our dependent variable. The results are displayed in Column (4) of Table 8. The estimated difference in difference coefficient is negative and statistically significant. Specifically, markups drop by 10bps for GSIBs as compared to non-GSIBs after the imposition of the capital charges. Once again, we plot the time series of coefficients in Figure 5, which exhibits a steady drop following the implementation of the surcharges. Finally, in Column (5) we interact the coefficient of interest with *StayBank* to see if the drop in markups is concentrated among firms that stay with their GSIBs. Consistent with this hypothesis the coefficient is negative and statistically significant with a point estimate of -.18.

4 Conclusion

We find that banks' private information creates market power through adverse selection in corporate bank loan markets. To identify market power, we provide a novel estimate of markup by controlling for banks' private information about borrower quality. While the existing literature finds evidence that more concentrated banking markets have higher deposit rates, we find the opposite in loan markets. Hence, a potential unintended consequence of antitrust policies is that by making banking markets less concentrated these policies may also raise interest rates on local bank loans.

Consistent with the adverse selection channel, we show that i) markups are higher for firms that remain with their existing banks and ii) the imposition of capital surcharges on the largest US banks led to a reduction in markups by mitigating the adverse selection problem. Overall, our results suggest that adverse selection is an important driver of market power in local corporate loan markets.

References

Abdymomunov, Azamat, Filippo Curti, and Atanas Mihov, 2020, Us banking sector operational losses and the macroeconomic environment, *Journal of Money, Credit and Banking* 52, 115–144.

Abreu, Dilip, 1986, Extremal equilibria of oligopolistic supergames, *Journal of Economic Theory* 39, 191–225.

Adelino, Manuel, Ivan Ivanov, and Michael Smolyansky, 2019, Humans vs machines: Soft and hard information in corporate loan pricing, *Available at SSRN 3596010* .

Agarwal, Sumit and Robert Hauswald, 2010, Distance and private information in lending, *The Review of Financial Studies* 23, 2757–2788.

Almeida, Heitor and Murillo Campello, 2007, Financial constraints, asset tangibility, and corporate investment, *The Review of Financial Studies* 20, 1429–1460.

Balasubramanyan, Lakshmi, Allen N Berger, and Matthew M Koepke, 2019, How do lead banks use their private information about loan quality in the syndicated loan market?, *Journal of Financial Stability* 43, 53–78.

Becker, Bo, Marieke Bos, and Kasper Roszbach, 2018, Bad times, good credit, *Swedish House of Finance Research Paper* .

Begley, Taylor A, Amiyatosh Purnanandam, and Kuncheng Zheng, 2017, The strategic underreporting of bank risk, *The Review of Financial Studies* 30, 3376–3415.

Behn, Markus, Rainer FH Haselmann, and Vikrant Vig, 2016, The limits of model-based regulation .

Berger, Allen N, 1995, The profit-structure relationship in banking–tests of market-power and efficient-structure hypotheses, *Journal of Money, Credit and Banking* 27, 404–431.

Berger, Allen N, Asli Demirgüç-Kunt, Ross Levine, and Joseph G Haubrich, 2004, Bank concentration and competition: An evolution in the making, *Journal of Money, credit and Banking* 433–451.

Berger, Allen N, Richard J Rosen, and Gregory F Udell, 2007, Does market size structure affect competition? the case of small business lending, *Journal of Banking & Finance* 31, 11–33.

Bertrand, Marianne, Antoinette Schoar, and David Thesmar, 2007, Banking deregulation and industry structure: Evidence from the french banking reforms of 1985, *The Journal of Finance* 62, 597–628.

Beyhaghi, Mehdi, 2022, Third-party credit guarantees and the cost of debt: Evidence from corporate loans, *Review of Finance* 26, 287–317.

Bidder, Rhys M, John R Krainer, and Adam Hale Shapiro, 2020, De-leveraging or de-risking? how banks cope with loss, *Review of Economic Dynamics* .

Bofondi, Marcello and Giorgio Gobbi, 2006, Informational barriers to entry into credit markets, *Review of Finance* 10, 39–67.

Bolton, Patrick, Tano Santos, and Jose A Scheinkman, 2016, Cream-skimming in financial markets, *The Journal of Finance* 71, 709–736.

Broecker, Thorsten, 1990, Credit-worthiness tests and interbank competition, *Econometrica: Journal of the Econometric Society* 429–452.

Brown, James R, Matthew Gustafson, and Ivan Ivanov, 2017, Weathering cash flow shocks, *Available at SSRN 2963444* .

Carbo-Valverde, Santiago, Francisco Rodriguez-Fernandez, and Gregory F Udell, 2009, Bank market power and sme financing constraints, *Review of Finance* 13, 309–340.

Cavalluzzo, Ken S, Linda C Cavalluzzo, and John D Wolken, 2002, Competition, small business financing, and discrimination: Evidence from a new survey, *The Journal of Business* 75, 641–679.

Cetorelli, Nicola, 2002, Life-cycle dynamics in industrial sectors. the role of banking market structure .

Cetorelli, Nicola and Michele Gambera, 2001, Banking market structure, financial dependence and growth: International evidence from industry data, *The Journal of Finance* 56, 617–648.

Claessens, Stijn and Luc Laeven, 2005, Financial dependence, banking sector competition, and economic growth, *Journal of the European Economic Association* 3, 179–207.

Crawford, Gregory S, Nicola Pavanini, and Fabiano Schivardi, 2018, Asymmetric information and imperfect competition in lending markets, *American Economic Review* 108, 1659–1701.

Cyrnak, Anthony W and Timothy H Hannan, 1999, Is the cluster still valid in defining banking markets? evidence from a new data source, *The Antitrust Bulletin* 44, 313–331.

De Loecker, Jan and Frederic Warzynski, 2012, Markups and firm-level export status, *American economic review* 102, 2437–71.

Degryse, Hans and Steven Ongena, 2008, Competition and regulation in the banking sector: A review of the empirical evidence on the sources of bank rents, *Handbook of financial intermediation and banking* 2008, 483–554.

Dell'Ariccia, Giovanni, 2001, Asymmetric information and the structure of the banking industry, *European Economic Review* 45, 1957–1980.

Dell'Ariccia, Giovanni, Luc Laeven, and Gustavo A Suarez, 2017, Bank leverage and monetary policy's risk-taking channel: evidence from the united states, *the Journal of Finance* 72, 613–654.

Dell'Ariccia, Giovanni and Robert Marquez, 2006, Lending booms and lending standards, *The Journal of Finance* 61, 2511–2546.

Dell'Ariccia, Giovanni, Deniz Igan, and Luc UC Laeven, 2012, Credit booms and lending standards: Evidence from the subprime mortgage market, *Journal of Money, Credit and Banking* 44, 367–384.

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

Driscoll, John C and Ruth Judson, 2013, Sticky deposit rates, *Available at SSRN 2241531* .

Favara, Giovanni, Ivan Ivanov, and Marcelo Rezende, 2021, Gsib surcharges and bank lending: Evidence from us corporate loan data, *Journal of Financial Economics* 142, 1426–1443.

Fernholz, Ricardo T. and Christoffer Koch, 2020, Why are big banks getting bigger?, *Unpublished working paper.* .

Fishman, Michael J and Jonathan A Parker, 2015, Valuation, adverse selection, and market collapses, *The Review of Financial Studies* 28, 2575–2607.

Frame, W Scott, Ping McLemore, and Atanas Mihov, 2020, Haste makes waste: Banking organization growth and operational risk .

Friedman, James W, 1971, A non-cooperative equilibrium for supergames, *The Review of Economic Studies* 38, 1–12.

Gopal, Manasa and Philipp Schnabl, 2020, The rise of finance companies and fintech lenders in small business lending, *NYU Stern School of Business* .

Greenbaum, Stuart I, George Kanatas, and Itzhak Venezia, 1989, Equilibrium loan pricing under the bank-client relationship, *Journal of Banking & Finance* 13, 221–235.

Greenwald, Daniel L, John Krainer, and Pascal Paul, The credit line channel (Federal Reserve Bank of San Francisco 2020).

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, 2019, Are us industries becoming more concentrated?, *The Review of Finance* 23, 697–743.

Gurun, Umit G, Gregor Matvos, and Amit Seru, 2016, Advertising expensive mortgages, *The Journal of Finance* 71, 2371–2416.

Hale, Galina and Joao AC Santos, 2009, Do banks price their informational monopoly?, *Journal of Financial Economics* 93, 185–206.

Hannan, Timothy H, 1991, Bank commercial loan markets and the role of market structure: Evidence from surveys of commercial lending, *Journal of Banking & Finance* 15, 133–149.

Hannan, Timothy H, 1997, Market share inequality, the number of competitors, and the hhi: An examination of bank pricing, *Review of Industrial Organization* 12, 23–35.

Hatfield, John and Jonathan Wallen, 2022, Many markets make good neighbors: Multimarket contact and deposit banking .

Heider, Florian, 2003, Leverage and asymmetric information about risk and value, *Unpublished working paper. New York University* .

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

Ivanov, Ivan, Luke Pettit, and Toni M Whited, 2020, Taxes depress corporate borrowing: Evidence from private firms, *Available at SSRN 3694869* .

Jayaratne, Jith and Philip E Strahan, 1996, The finance-growth nexus: Evidence from bank branch deregulation, *The Quarterly Journal of Economics* 111, 639–670.

Jiménez, Gabriel and Jesús Saurina, 2004, Collateral, type of lender and relationship banking as determinants of credit risk, *Journal of banking & Finance* 28, 2191–2212.

Keeley, Michael C, 1990, Deposit insurance, risk, and market power in banking, *The American economic review* 1183–1200.

Liebersohn, Jack, 2017, How does competition affect bank lending? quasi-experimental evidence from bank mergers, Working paper, Technical report.

Marquez, Robert, 2002, Competition, adverse selection, and information dispersion in the banking industry, *The Review of Financial Studies* 15, 901–926.

Neumark, David and Steven A Sharpe, 1992, Market structure and the nature of price rigidity: evidence from the market for consumer deposits, *The Quarterly Journal of Economics* 107, 657–680.

di Patti, Emilia Bonaccorsi and Giovanni Dell'Ariccia, 2004, Bank competition and firm creation, *Journal of Money, Credit and Banking* 225–251.

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

Plosser, Matthew C and Joao AC Santos, 2018, Banks' incentives and inconsistent risk models, *The Review of Financial Studies* 31, 2080–2112.

Prager, Robin A and Timothy H Hannan, 1998, Do substantial horizontal mergers generate significant price effects? evidence from the banking industry, *The Journal of Industrial Economics* 46, 433–452.

Qian, Jun, Philip E Strahan, and Zhishu Yang, 2015, The impact of incentives and communication costs on information production and use: Evidence from bank lending, *The Journal of Finance* 70, 1457–1493.

Rajan, Raghuram G, 1992, Insiders and outsiders: The choice between informed and arm's-length debt, *The Journal of finance* 47, 1367–1400.

Rhoades, Stephen A, 1995, Market share inequality, the hhi, and other measures of the firm-composition of a market, *Review of industrial organization* 10, 657–674.

Rice, Tara and Philip E Strahan, 2010, Does credit competition affect small-firm finance?, *The Journal of Finance* 65, 861–889.

Riordan, Michael H, 1995, 11 competition and bank performance: a theoretical perspective, *Capital markets and financial intermediation* 328.

Saidi, Farzad and Daniel Streitz, 2018, Bank concentration and product market competition, *Work. Pap., Stockh. Sch. Econ., Stockholm Google Scholar Article Location* .

Santos, Joao AC and Andrew Winton, 2008, Bank loans, bonds, and information monopolies across the business cycle, *The Journal of Finance* 63, 1315–1359.

Sapienza, Paola, 2002, The effects of banking mergers on loan contracts, *The Journal of finance* 57, 329–367.

Schenone, Carola, 2010, Lending relationships and information rents: Do banks exploit their information advantages?, *The Review of Financial Studies* 23, 1149–1199.

Shaffer, Sherrill, 1998, The winner's curse in banking, *Journal of Financial Intermediation* 7, 359–392.

Sharpe, Steven A, 1990, Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships, *The journal of finance* 45, 1069–1087.

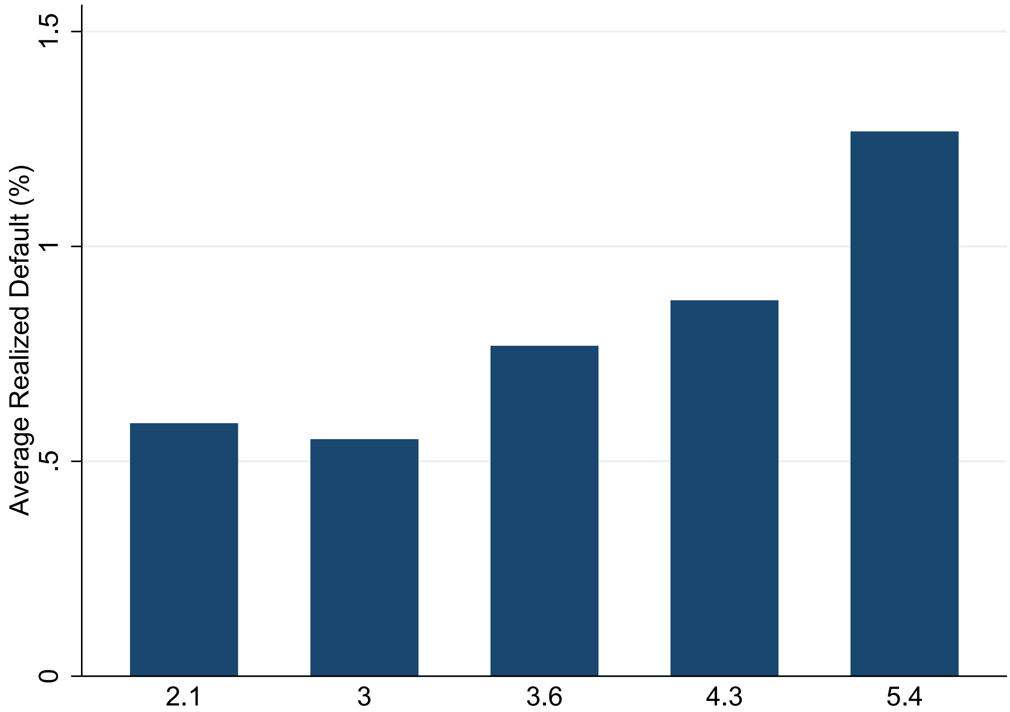
Stigler, George J, 1964, A theory of oligopoly, *Journal of political Economy* 72, 44–61.

Stigler, George J and Robert A Sherwin, 1985, The extent of the market, *The Journal of Law and Economics* 28, 555–585.

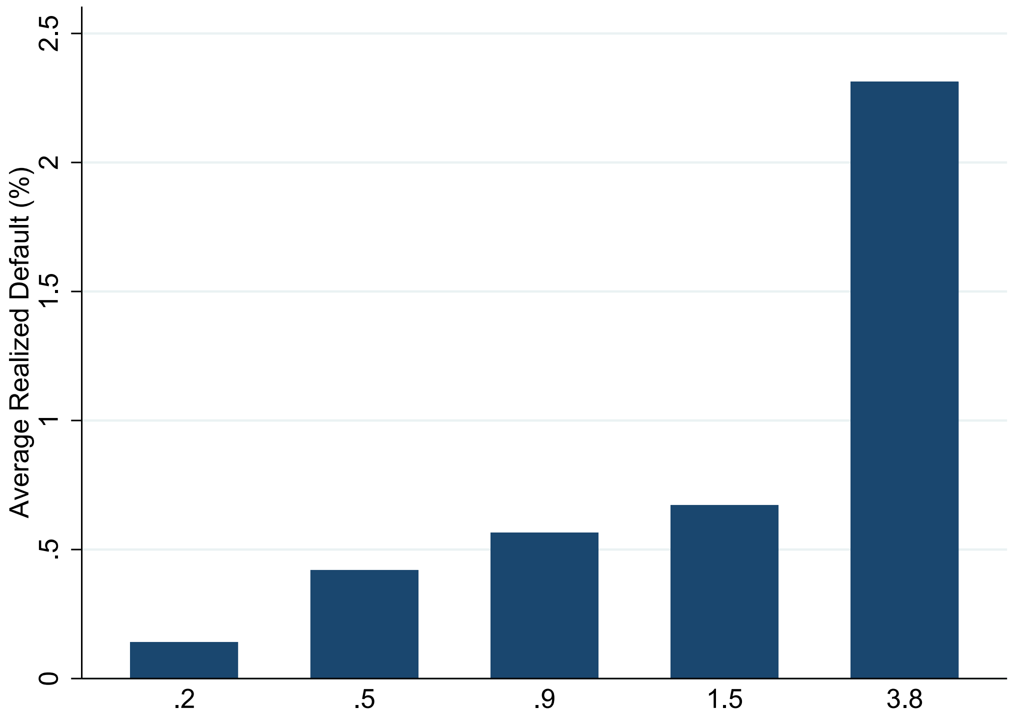
Wang, Yifei, Toni M Whited, Yufeng Wu, and Kairong Xiao, 2018, Bank market power and monetary policy transmission: Evidence from a structural estimation, *Available at SSRN 3049665* .

Yannelis, Constantine and Anthony Lee Zhang, 2021, Competition and selection in credit markets, Working paper, National Bureau of Economic Research.

Zarutskie, Rebecca et al., Does bank competition affect how much firms can borrow? new evidence from the us, *Federal Reserve Bank of Chicago Proceedings*, 856 (2003).



(a) Average Realized Default Rate Across Interest Rate Bins



(b) Average Realized Default Rate Across PD Bins

Figure 1a plots the average realized default rates over the twelve months following origination across five interest rate bins. Figure 1b plots the average realized default rates over the twelve months following origination across five PD bins. The average interest rate or PD in each bin is listed below each bar.

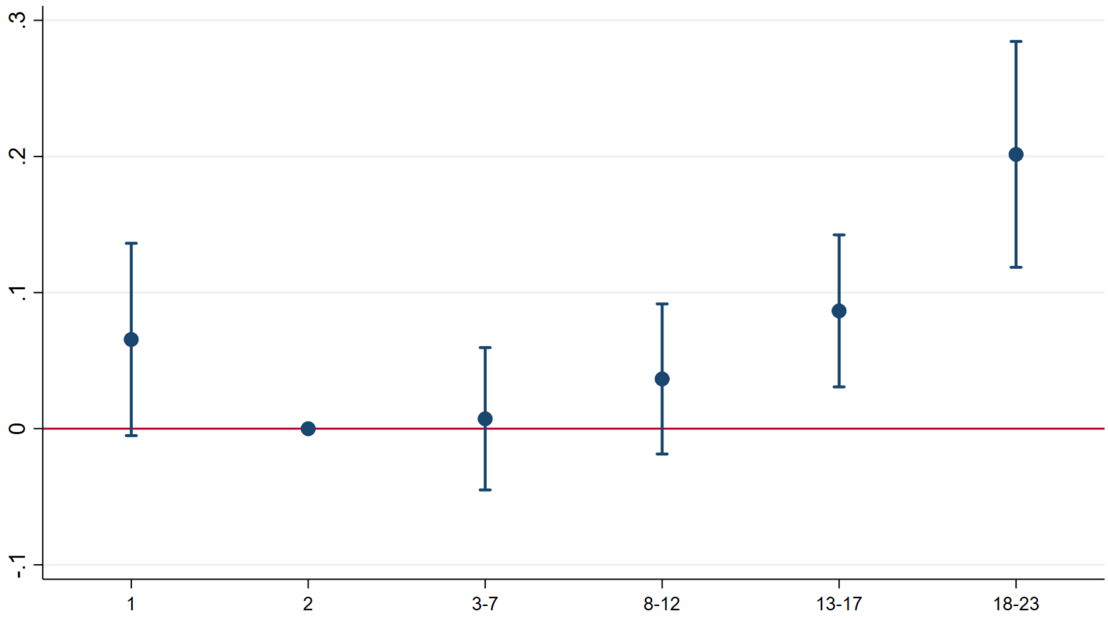


Figure 2: Number of Banks in the County and Markups

Figure 2 plots estimated coefficients, with 90% confidence intervals, for regressions of risk-adjusted markup on different number of bank group dummies. The number of banks in each group is listed below where reference number of banks is 2. Standard errors are clustered by county.

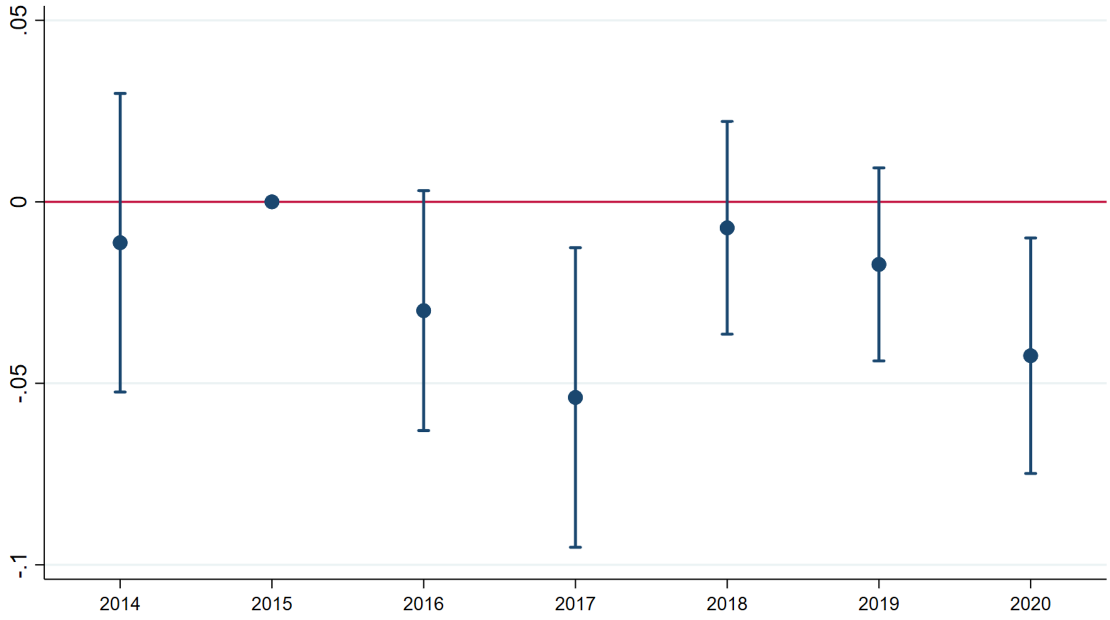


Figure 3: The Effect of GSIB Capital Surcharges on Market Shares

Figure 3 plots estimated regression coefficients with 90% confidence intervals from a version of (7) with annual interaction terms. Standard errors are clustered by county.

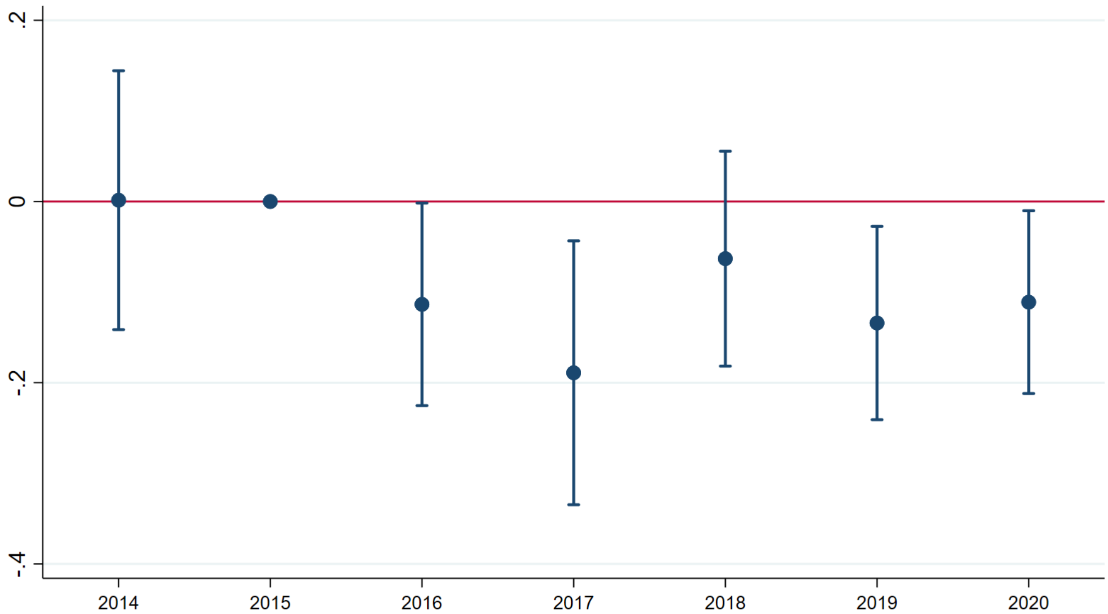


Figure 4: The Effect of GSIB Market Share on Markups

Figure 4 plots estimated regression coefficients with 90% confidence intervals from version of (8) with annual interaction terms. Standard errors are clustered by county.

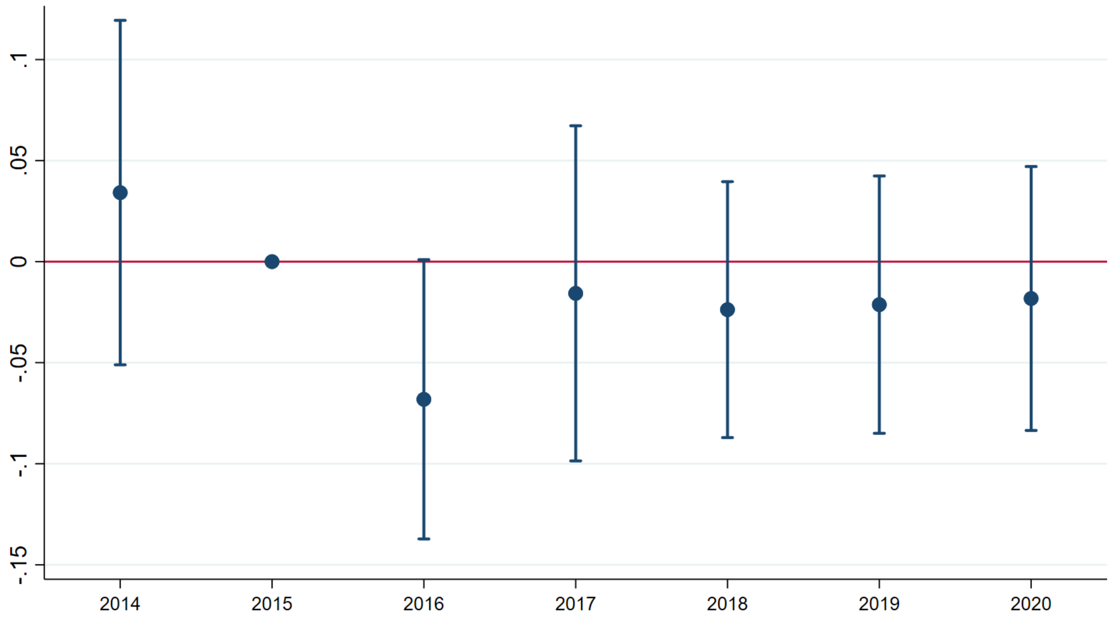


Figure 5: GSIB Banks and Markups

Figure 4 plots estimated regression coefficients with 90% confidence intervals from a version of (7) with annual interaction terms and risk-adjusted markup as the dependent variable. Standard errors are clustered by county.

Table 1: Summary Statistics

This table contains summary statistics for loan-level, firm and geographic characteristics. Section A of the Appendix includes detailed definitions of all of our variables and Section 3.2 explains our filters.

	Mean	SD	10%	Median	90%	N
Loan Characteristics						
Amount (million USD)	6.97	14.15	1.04	2.54	15.75	28,033
Collateral	0.90	0.30	0.00	1.00	1.00	28,033
Non-Performance (%)	2.01	14.03	0.00	0.00	0.00	28,033
Floating Interest Rate	0.79	0.41	0.00	1.00	1.00	28,033
Guaranteed	0.49	0.50	0.00	0.00	1.00	28,033
Interest Rate (%)	3.66	1.17	2.16	3.63	5.25	28,033
Probability of Default (%)	1.38	1.74	0.21	0.90	2.83	28,033
Loss Given Default (%)	35.28	14.88	15.00	36.00	50.50	28,033
Expected Loss (%)	0.46	0.61	0.06	0.29	0.99	28,033
Markup (Baseline) (%)	0.00	0.82	-0.94	-0.08	1.05	28,033
Markup (Risk-Adjusted) (%)	0.00	0.80	-0.91	-0.07	1.02	28,033
Maturity (months)	41.32	31.50	10.00	36.00	84.00	28,033
Predicted IR (Baseline) (%)	3.66	0.84	2.62	3.62	4.81	28,033
Predicted IR (Risk-Adjusted) (%)	3.66	0.86	2.58	3.62	4.83	28,033
Revolver/Line of Credit	0.51	0.50	0.00	1.00	1.00	28,033
Realized Default (%)	0.81	8.96	0.00	0.00	0.00	28,033
GSIB	0.42	0.49	0.00	0.00	1.00	28,033
Firm Characteristics						
Assets (million USD)	109.66	410.70	2.71	20.06	175.80	28,033
Leverage	0.33	0.26	0.01	0.30	0.69	27,507
Profitability	0.14	0.26	-0.01	0.08	0.33	28,033
Tangibility	0.91	0.17	0.67	1.00	1.00	27,963
Geographic Characteristics						
Number of Banks	10.88	5.48	3.00	11.00	18.00	28,033
One Bank	0.03	0.17	0.00	0.00	0.00	28,033
Population Density	6.78	1.54	4.71	6.98	8.27	28,033
Average Wages	9.50	0.27	9.19	9.48	9.82	26,118
Deposit HHI	0.20	0.08	0.11	0.18	0.29	28,033
Loan HHI	0.48	0.24	0.23	0.41	0.89	28,033
GSIB Market Share	0.50	0.30	0.00	0.53	0.91	26,376

Table 2: Predictiveness of Risk Assessments on Loan Performance

This table tests whether banks internal risk assessments predict non-performance and default. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by firm. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Non-Performance (%)		Realized Default (%)	
	(1)	(2)	(3)	(4)
Probability of Default (%)		1.461***		0.629***
		(5.994)		(2.864)
Loss Given Default (%)		0.024**		-0.001
		(2.392)		(0.134)
Expected Loss (%)		-0.837		0.408
		(1.212)		(0.634)
Loan Characteristics Controls	YES	YES	YES	YES
Bank-Quarter FE	YES	YES	YES	YES
Industry-Quarter FE	YES	YES	YES	YES
Loan Type FE	YES	YES	YES	YES
Loan Purpose FE	YES	YES	YES	YES
Observations	28,033	28,033	28,033	28,033
Adj. R-squared	0.08	0.09	0.05	0.07

Table 3: Estimating Markup

This table tests whether banks internal risk assessments predict loan interest rates and is used to calculate both baseline and risk-adjusted markup. Columns (1) and (2) contains the estimation of the baseline markup and risk-adjusted markup, respectively. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by firm. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Interest Rate (%)	
	(1)	(2)
Probability of Default (%)		0.075***
		(7.201)
Loss Given Default (%)		0.002***
		(3.191)
Expected Loss (%)		0.160***
		(5.353)
Loan Characteristics Controls	YES	YES
Bank-Quarter FE	YES	YES
Industry-Quarter FE	YES	YES
Loan Type FE	YES	YES
Loan Purpose FE	YES	YES
Observations	28,033	28,033
Adj. R-squared	0.49	0.52

Table 4: Validity of Markup and Predicted Interest Rate

This table tests the validity of the baseline and risk-adjusted markup. Panel A uses Non-Performance as the dependent variable, while Panel B uses Realized Default. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by firm. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Non-Performance (%)						
	(1)	(2)	(3)	(4)	(5)	(6)
Interest Rate (%)	0.264***			0.607***		
	(3.274)			(5.427)		
Baseline Markup (%)		0.603***			0.603***	
		(4.991)			(5.109)	
Baseline Predicted IR (%)		-0.064			0.691*	
		(0.554)			(1.647)	
Risk-Adjusted Markup (%)			0.083			0.083
			(0.649)			(0.685)
Risk-Adjusted Predicted IR (%)			0.420***			5.065***
			(3.565)			(10.235)
Bank-Quarter FE				YES	YES	YES
Industry-Quarter FE				YES	YES	YES
Loan Type FE				YES	YES	YES
Loan Purpose FE				YES	YES	YES
Observations	28,033	28,033	28,033	28,033	28,033	28,033
Adj. R-squared	0.00	0.00	0.00	0.07	0.07	0.08

Panel B: Realized Default (%)						
	(1)	(2)	(3)	(4)	(5)	(6)
Interest Rate (%)	0.211***			0.366***		
	(3.830)			(4.588)		
Baseline Markup (%)		0.407***			0.407***	
		(4.650)			(4.811)	
Baseline Predicted IR (%)		0.022			-0.469*	
		(0.311)			(1.819)	
Risk-Adjusted Markup (%)			0.067			0.067
			(0.729)			(0.764)
Risk-Adjusted Predicted IR (%)			0.336***			2.910***
			(4.099)			(6.856)
Bank-Quarter FE				YES	YES	YES
Industry-Quarter FE				YES	YES	YES
Loan Type FE				YES	YES	YES
Loan Purpose FE				YES	YES	YES
Observations	28,033	28,033	28,033	28,033	28,033	28,033
Adj. R-squared	0.00	0.00	0.00	0.05	0.05	0.06

Table 5: The Relationship Between County and Firm Characteristics and Markups

This table tests the relationship between the number of banks, firm and county characteristics, and risk-adjusted markups. The dependent variable is risk-adjusted markup. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by county. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Risk-Adjusted Markup (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Banks	0.010***	0.011***	0.011***	0.010***	0.007**	0.007**
	(4.960)	(5.266)	(4.913)	(3.145)	(2.188)	(2.253)
One Bank		0.114***	0.122***	0.127***	0.112***	0.119***
		(2.997)	(3.134)	(3.045)	(2.798)	(2.876)
Log(Assets)			−0.059***	−0.059***	−0.062***	−0.095***
			(14.098)	(14.236)	(14.155)	(19.187)
Leverage			0.087***	0.088***	0.090***	0.105***
			(2.947)	(2.970)	(2.980)	(3.704)
Tangibility			−0.314***	−0.313***	−0.331***	−0.492***
			(8.405)	(8.392)	(8.590)	(12.094)
Profitability			−0.047*	−0.047*	−0.055*	−0.067**
			(1.776)	(1.785)	(1.957)	(2.377)
Population Density				0.005	−0.006	−0.005
				(0.364)	(0.331)	(0.242)
Average Wages					0.170**	0.214**
					(2.345)	(2.554)
Bank-Quarter FE						YES
Industry-Quarter FE						YES
Loan Type FE						YES
Loan Purpose FE						YES
Observations	28,033	28,033	27,487	27,487	25,610	25,608
R-squared	0.00	0.01	0.03	0.03	0.03	0.05

Table 6: The Relationship Between County and Firm Characteristics and PDs

This table tests the relationship between the number of banks, firm and county characteristics, and probability of default. The dependent variable is probability of default. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by county. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Probability of Default (%)				
	(1)	(2)	(3)	(4)	(5)
Number of Banks	0.010***	0.008***	0.011***	0.006*	0.008**
	(3.855)	(3.044)	(4.565)	(1.838)	(2.487)
One Bank		−0.144**	−0.134**	−0.108*	−0.116**
		(2.527)	(2.453)	(1.920)	(2.036)
Log(Assets)			−0.148***	−0.148***	−0.149***
			(15.475)	(15.511)	(15.168)
Leverage			0.823***	0.825***	0.841***
			(12.710)	(12.739)	(12.475)
Tangibility			−0.289***	−0.289***	−0.288***
			(3.269)	(3.273)	(3.050)
Profitability			−1.064***	−1.064***	−1.048***
			(18.912)	(18.942)	(21.966)
Population Density				0.029**	0.037***
				(2.262)	(2.884)
Average Wages					−0.131**
					(1.975)
Loan Characteristics Controls	YES	YES	YES	YES	YES
Bank-Quarter FE	YES	YES	YES	YES	YES
Industry-Quarter FE	YES	YES	YES	YES	YES
Loan Type FE	YES	YES	YES	YES	YES
Loan Purpose FE	YES	YES	YES	YES	YES
Observations	28,033	28,033	27,485	27,485	25,608
R-squared	0.19	0.19	0.24	0.24	0.24

Table 7: Switching Banks and Markups

This table tests whether firms that stay with their existing banks face higher markups. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by county. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Risk-Adjusted Markup (%)		
	(1)	(2)	(3)
Stay Bank	0.059**	0.074***	0.077***
	(2.563)	(3.401)	(3.428)
Log(Assets)		−0.051***	−0.082***
		(7.608)	(10.916)
Leverage		0.134***	0.167***
		(3.069)	(3.905)
Tangibility		−0.369***	−0.556***
		(6.476)	(8.480)
Profitability		−0.126***	−0.121***
		(3.376)	(3.203)
County-Quarter FE	YES	YES	YES
Bank-Quarter FE			YES
Industry-Quarter FE			YES
Loan Type FE			YES
Loan Purpose FE			YES
Observations	15,491	15,118	15,068
R-squared	0.31	0.32	0.39

Table 8: GSIB Capital Surcharges and Adverse Selection

This table contains regression results from the tests described in Section 3.7. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by county. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Market Share	Risk-Adjusted Markup (%)			
	(1)	(2)	(3)	(4)	(5)
GSIB × Post	−0.022*			−0.100***	0.001
	(1.689)			(3.151)	(0.011)
GSIB Market Share × Post		−0.110**	0.092		
		(2.216)	(0.727)		
Stay Bank			−0.067		−0.060
			(1.021)		(1.175)
Stay Bank × Post			0.252***		0.171***
			(3.322)		(2.945)
Stay Bank × GSIB Market Share			0.194*		
			(1.752)		
Stay Bank × GSIB					0.203***
					(2.594)
Stay Bank × GSIB Market Share × Post			−0.317**		
			(2.392)		
Stay Bank × GSIB × Post					−0.180**
					(2.082)
Firm-Level Controls		YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES	YES
Bank-County FE	YES	YES	YES	YES	YES
Observations	25,872	24,365	17,089	25,363	17,789
R-squared	0.70	0.25	0.28	0.26	0.29

Appendix A. Variable Definitions

Baseline Markup: The estimated residual from Equation (3) in percentage points, excluding PD, LGD and PDLGD, from Y-14Q.

Collateral: Dummy variable that equals one if the loan is collateralized, from Y-14Q.

Deposit HHI: The average annual Deposit HHI (sum of squared bank market shares) from each county, from Drechsler, Savov, and Schnabl (2017).

Firm Size: log(assets) trimmed at the 99th percentile, from Y-14Q.

Floating Interest Rate: Dummy variable that equals one if the loan is floating rate, from Y-14Q.

GSIB: Dummy variable that equals one if the loan is from a global systemically important bank (GSIB), from Y-14Q.

GSIB Share: The sum of 2015 market shares of GSIBs in the county, from Y-14Q.

HHI: The average annual county HHI (sum of squared bank market shares) for each county over the entire sample period, from Y-14Q.

Interest Rate: Loan interest rate in percentage points, trimmed at [0,1), from Y-14Q.

Leverage: total debt/assets, winsorized at [1%, 99%], from Y-14Q.

LGD: The bank's estimated loss given default, from Y-14Q.

Number of Banks: Number of unique banks to have given a loan in a county at any point over the entire sample, from Y-14Q.

Maturity: Log of loan maturity in months, from Y-14Q.

Non-Performance: Dummy variable that equals one if the bank reports the loan as 90 days past due or non-accrual, or reports a positive net cumulative charge-off amount, or reports specific reserve for an impaired loan for the loan within the 12 months following the origination of the loan, or if the bank considers the borrower as defaulted as defined above, from Y-14Q.

Probability of Default (PD): The bank's expected annual default rate over the life of the loan, trimmed if $PD = 0$ or above the 99th percentile, from Y-14Q.

PDLGD: Expected Loss. $PD \times LGD$, from Y-14Q.

Population Density: Average county population per square mile, from Census.

Profitability: EBITDA/assets, winsorized at [1%, 99%], from Y-14Q.

Predicted Interest Rate: The predicted interest rate from regression (3) in percentage points, the baseline model excludes PD, LGD and PDLGD, while the risk-adjusted model includes PD, LGD and PDLGD, from Y-14Q.

Realized Default: Dummy variable that equals one if the borrower is rated D (defaulted) or is assigned a PD=100% by the lending bank within one year after the origination of the loan, from Y-14Q.

Revolver/Line of Credit: Dummy variable that equals one if the loan is a revolver or line of credit, from Y-14Q.

Risk-Adjusted Markup: The estimated residual from Equation (3) in percentage points, including PD, LGD and PDLGD, from Y-14Q.

Stay Bank: Dummy variable that equals one if the firm borrows from the bank it received its previous loan from, from Y-14Q.

Tangibility: tangible assets/assets, winsorized at [1%, 99%], from Y-14Q.

Wages: Average county level wages in logs, from BLS.

Appendix B. Additional Tests

Table B1: Correlation Matrix

This table contains a correlation matrix containing different measures of market concentration at the county level.

Variables	Number of Banks	Deposit HHI	Loan HHI
Number of Banks	1.00		
Deposit HHI	-0.25	1.00	
Loan HHI	-0.89	0.24	1.00

Table B2: Alternative Measures of Market Concentration and Markups

This table tests the relationship between alternative measures of market concentration and risk-adjusted markup. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by county or MSA depending on whether county or MSA measures of concentration are used in the regression. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Risk-Adjusted Markup (%)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Loan HHI	−0.145***	−0.175***						
	(3.502)	(4.134)						
Number of Banks MSA			0.007***	0.009***				
			(4.596)	(5.120)				
MSA Loan HHI					−0.131***	−0.154***		
					(2.931)	(3.388)		
Deposit HHI							−0.081	−0.102
							(0.818)	(0.947)
Log(Assets)	−0.059***	−0.091***	−0.060***	−0.091***	−0.061***	−0.094***	−0.059***	−0.091***
	(14.107)	(18.492)	(14.300)	(18.484)	(13.730)	(18.434)	(14.063)	(18.313)
Leverage	0.084***	0.101***	0.084***	0.100***	0.077**	0.091***	0.075**	0.094***
	(2.829)	(3.623)	(2.799)	(3.594)	(2.485)	(3.140)	(2.545)	(3.329)
Tangibility	−0.319***	−0.471***	−0.321***	−0.472***	−0.338***	−0.492***	−0.325***	−0.476***
	(8.474)	(12.016)	(8.536)	(12.061)	(8.693)	(11.957)	(8.554)	(12.095)
Profitability	−0.044*	−0.059**	−0.045*	−0.059**	−0.050*	−0.066**	−0.038	−0.057**
	(1.696)	(2.208)	(1.730)	(2.194)	(1.781)	(2.265)	(1.460)	(2.103)
Bank-Quarter FE		YES		YES		YES		YES
Industry-Quarter FE		YES		YES		YES		YES
Loan Type FE		YES		YES		YES		YES
Loan Purpose FE		YES		YES		YES		YES
Observations	27,487	27,485	27,487	27,485	25,068	25,066	27,487	27,485
R-squared	0.02	0.04	0.03	0.04	0.03	0.04	0.02	0.04

Table B3: Baseline Markup Estimated with Firm Characteristics

This table tests whether an estimate of markup using firm characteristics but not banks' private risk assessments predicts loan performance. T-statistics are shown below the parameter estimates in parenthesis and are calculated using robust standard errors clustered by firm. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Non-Performance (%)		Realized Default (%)	
	(1)	(2)	(3)	(4)
Markup (%)	0.630***	0.630***	0.417***	0.417***
	(4.914)	(5.070)	(4.539)	(4.702)
Predicted IR (%)	−0.015	0.591*	0.054	0.046
	(0.129)	(1.913)	(0.739)	(0.240)
Bank-Quarter FE		YES		YES
Industry-Quarter FE		YES		YES
Loan Type FE		YES		YES
Loan Purpose FE		YES		YES
Observations	27,317	27,317	27,317	27,317
Adj. R-squared	0.00	0.07	0.00	0.05