RESEARCH

Swiss Finance Institute
Research Paper Series
N°23-01

"If You Don't Know Me by Now …" Information Collection by Banks in Lending to Private Firms

Stijn Claessens
Yale School of Management and CEPR

Steven Ongena
University of Zurich, KU Leuven, NTNU Business School, Swiss Finance Institute, and CEPR

Teng Wang
Board of Governors of the Federal Reserve System

“If You Don't Know Me by Now ...”

Information Collection by Banks in Lending to Private Firms

Stijn Claessens, Steven Ongena, Teng Wang*

January 20, 2025

What constitutes banks’ information collection of private firms? How does it deepen and change in nature over time? Exploiting a comprehensive Federal Reserve’s supervisory dataset, we first extract three dimensions of private information from banks’ internal credit ratings — depth and better/worse assessments — with all three dimensions related to various loan terms. We then document how private information evolves as firm-bank relationships lengthen, with effects non-linear and peaking after five years. Learning varies by bank and firm characteristics too, with effects particularly salient at longer bank-firm distances, during non-COVID times, for smaller, leveraged, illiquid banks, or for smaller, leveraged firms. (99 words)

* Corresponding author. Claessens is with Yale School of Management and the CEPR and was with the Bank for International Settlements when the paper was started; email: claessens.stijn@outlook.com. Ongena is with the University of Zurich, Swiss Finance Institute, KU Leuven, NTNU Business School, and CEPR; Plattenstrasse 14, CH-8032 Zürich, Switzerland; email: steven.ongena@df.uzh.ch; phone: +41 44 634 39 54. Wang is with the Board of Governors of the Federal Reserve System; 20th St and Constitution Ave NW, Washington, DC 20551, USA; email: teng.wang@frb.gov; phone: +1 202 452 6436. For comments, we are grateful to Paul Calem, Sheng Li, Hirofumi Uchida, Greg Weitzner, seminar participants at the Bank for International Settlements (Basel), Federal Reserve Board, Federal Reserve Bank of Kansas City, and conference participants at the American Economic Association Annual Meetings, Federal Reserve System Committee Meeting on Financial Institutions, Regulation and Markets, Banking and Finance Workshop (Tokyo), and the European Economic Association Annual Meetings. The paper circulated previously under the titles “Out of Sight, Out of Mind?” Banks’ Private Information, Distance, and Relationship Length” and “"If You Don't Know Me by Now ...″ Banks’ Private Information and Relationship Length.” The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Bank for International Settlements, the Board of Governors, or any staff of these institutions.

I. Introduction

Privately held firms are a cornerstone of the U.S. economy. For example, as of 2022, approximately 90% of the 200,000 middle-market businesses in the U.S. — representing one-third of private sector GDP, employing around 48 million people, and generating more than $10 trillion in annual revenue — are privately held.[1] Unlike public companies, which can raise funds through equity and bond markets, private firms depend heavily on bank financing to support their operations and growth. This reliance underscores the pivotal role of banks in overcoming information asymmetries in lending markets, a function that is central to their ability to assess and manage credit risk. Classic models by Leland and Pyle (1977) and Fama (1985) highlight the significance of private information collection in mitigating adverse selection and moral hazard. Banks accumulate valuable private information through screening and monitoring the creditworthiness of their borrowers, which in turn enhances their lending decisions (e.g., Diamond (1984) and Bester (1985)).

Yet despite the central role played by private information in bank lending, it is not well-understood how this private information is acquired, how it varies over time, and how it is shaped by firm and/or bank characteristics, including the physical distance between banks and firms (e.g., Hauswald and Marquez (2003)) and circumstances that affect the modus communicandi and economic conditions (like COVID-19). This paper seeks to address these gaps. In particular, it addresses the following questions. What is in banks' private information about private borrowers? How is private information affected by physical distance? How does

[1] Veronis (2023) and the *National Center for the Middle Market: Q2 2022 Middle Market Indicator*.

it evolve over time, and what bank and firm elements affect its evolution? By examining the depth and dynamics of banks' private information, this study contributes to understanding the mechanisms underpinning the financing of private firms, of which there well over 25 million in the U.S. (in contrast, there are less than 4,000 public firms).

To quantify the nature, formation, and implications of private information embedded in banks' evaluations of private borrowers, we use the Federal Reserve's comprehensive administrative loan-level internal rating data on corporate commercial and industrial (C&I) portfolios held by major banks in the U.S. We develop three key proxies for banks' private information sets and assess their importance for banks' loan terms. We then analyze the factors that drive their evolutions, and similarities and differences across banks and firms. Collectively, this research helps in understanding the intricate process of how banks assess the creditworthiness of corporate borrowers through their private information collection.

While it is widely acknowledged that banks' private information plays a crucial role in determining lending decisions and outcomes, there has been a lack of empirical research into its properties, determining factors, and potential implications. The main reason for this research gap is that banks' private or soft information is only accessible to the bank making the lending decision and often obscure (as noted by Liberti and Petersen (2019)).

For our analyses, we use internal bank credit ratings reported, along with other detailed loan and borrower information, in the Federal Reserve's supervisory Y-14Q quarterly loan-level dataset.[2] This dataset has a much broader representation than other datasets in the US (e.g.,

[2] Note that the Y-14Q is different from a traditional credit bureau in that it does not share any of the data individual banks report with other banks but contains information of internal credit risk assessment of each loan assigned by each individual banks' loan officers.

Berger, Bouwman, Norden, Roman, Udell, and Wang (2021); Faria-e-Castro, Paul, and Sánchez (2021); Beyhaghi, Howes, and Weitzner (2024)) as it covers all commercial loans of $1 million or more extended by the large bank holding companies (BHCs) operating in the U.S. that are subject to the DFAST/CCAR stress tests. Critically, for our study, this dataset uniquely includes the banks' internal ratings of the borrowers, which reflect not only the financial status of the firms and the banks' scoring models and techniques, but also the loan officers' internal credit assessments. Over 90% of the loans in this dataset are granted to non-listed (that is, private) firms, which are not required to disclose their financial data to the SEC or other financial regulators in the U.S.. As the dataset covers a rich cross-section of firms and major banks offering credit to them, it allows us to study how private information varies by firm and bank characteristics.[3]

Using banks' internal ratings for only the non-listed firms in the sample,[4] we first develop three novel proxies of banks' private assessment of corporate borrowers' riskiness: ***depth***, i.e., banks' assessments of firm risk relative to one based on hard information; and its ***direction***, i.e., ***better*** or ***worse*** assessments.

To identify the private information embedded in banks' rating, we follow the seminal work by Morgan (2002) and Agarwal and Hauswald (2010).[5] We deploy a two-stage model, which

[3] When it was put in place in 2011:Q3 all banks with over $50 billion in assets were required to report, but the Economic Growth, Regulatory Relief, and Consumer Protection Act (EGRRCPA) increased the reporting size threshold to $100 billion starting in 2019:Q4 (see further https://www.federalreserve.gov/supervisionreg/stress-tests-capital-planning.htm). Our results are robust, however, to only using banks with over $100 billion in assets in all our analyses.

[4] We drop the public-listed firms as information production for them is very different and largely not done by banks. In earlier versions of our paper, however, we also included the small share of public-listed firms and found no (meaningful) differences in results.

[5] Using the dispersion in the public credit ratings of banks, Morgan (2002) establishes that banks are more opaque than non-financial institutions.

is in a spirit similar to a heteroskedastic regression model (Harvey (1976)). Its first stage extracts from the internal bank credit ratings, the part that is not explained by an encompassing set of observable bank and firm characteristics. Since we run the first stage regression using both bank and Y-14Q firm data, the residual combines part of its process for generating ratings that cannot be explained by the bank characteristics and the information on the firms we include.

The part of the rating that observable characteristics cannot explain is private,[6] as it arises from the bank's specific business model after correcting for those bank features that could reasonably be expected to affect the rating process (e.g., larger size and higher capitalization provide greater scope for diversification benefits and loss absorption, respectively, allowing for relatively more risky firms to be rated higher). Additionally, using all the firm information in the Y-14Q may even give an implicit advantage to the heteroskedastic regression model in explaining the internal ratings as not all banks may have access to this information.[7]

This means that the residual of the first stage regression provides us with an overall measure of private information. We call (the logarithm of) the squared of this residual the *depth* in the banks' private information about firms (the larger the dispersion, the more depth there is to the private information). And we call the positive and negative residuals *better* and *worse* assessments, respectively, as indicators that also capture the *direction* of the banks' private information.

[6] This could also include (but is not limited to) the various (non-linear) ways in which the firm information is combined in the banks' proprietary credit risk modeling.

[7] As noted, only the lending bank demonstrably obtains this information from the borrowing firm and reports it then to the FR Y-14Q, but this firm information is not shared with or necessarily directly available to all other banks.

We show how these measures vary at the initiation of loans with bank and firm characteristics, including distance, i.e., how they relate to the screening process, confirming the roles of bank and firm characteristics typically identified in the literature when loans get made.

Consistent with the existing literature, we confirm that our three novel measures of banks' private information are the profound drivers of loan terms (and also with priors because one would expect the internal rating by construction to have a relationship with loan terms).[8] Specifically, the loan interest rate spread is higher with greater depth and lower (higher) when private information is positive (negative). Also, as expected, maturity, amount, and collateralization are longer, higher, and lower, respectively, with positive private information and vary in the opposite ways with negative private information. To demonstrate robustness, we also show that our three private information proxies contained in the ratings of loans by distinct banks to the same borrower, which only differ in how the banks privately assess the borrower, vary as well with loan terms.

For our main analysis, we investigate how the *depth* and *direction* of the lending banks' private information changes as their relationship with the firm lengthens and how it is affected by the distance between banks and firms, and the characteristics of the lending banks and the borrowing firms.

We first focus on relationship length since it has been identified as a salient driver of banks' private information. Much research has argued that relationship length is a proxy for the

[8] We include this assessment since the literature has focused on how outcome variables such as the amounts of lending and the terms thereof relate to banks' private information and internal ratings. But these relationships arise by construction and the focus of our paper is studying the banks' information production processes.

information asymmetry between the bank and the borrower, with the signal becoming more informative in the length of the lending relationship (e.g., Petersen and Rajan (1994)). We accordingly expect that the longer the bank has a relationship with the specific firm, the more private information will be produced about the firm, meaning the banks' internal rating will deviate more from the one based on observable characteristics only, i.e., its depth will be higher. As to the direction of private information, the literature conjectures that the valuable private information acquired during a relationship will make the bank improve its evaluation from the one based solely on observable characteristics and relative to other banks, resulting in higher *better* and lower *worse* private information.

Consistent with our conjectures, we show that the length of the bank-firm relationship contributes to how the depth of private information evolves. Specifically, as relationships between banks and firms progress, the banks' assessments of the firms deviate more from those based on the observable characteristics available to the banks (which themselves are continuously updated), suggesting that banks learn about the borrowers during their relationships. This increase in depth goes hand in hand with an increasingly better assessment by banks of the borrowing firms, providing support for the notion that relationship lending contributes to private information (e.g., Petersen and Rajan (1994)). This effect appears to build up over time with a peak impact at about five years.

In the second part of our main analysis, we study how the information impact of relationships varies with bank and firm characteristics, including bank-firm distance. In other words, does bank learning/private information formation vary alongside these dimensions? Consistent with theoretical literature, yet not empirically documented, we expect the various effects for the length of the relationship to differ by those characteristics reflecting varying abilities and incentives for private information acquisition as well as degrees of information asymmetries. For example, as a relationship lasts longer, we expect the adverse impact of distance on depth (e.g., Hauswald and Marquez (2003) and Hauswald and Marquez (2006)) to decrease as the bank produces valuable private information from its engagement with the firm, and the likelihood of the internal rating being better (relative to observable factors) to increase and it being worse to decrease. We also expect that the COVID period may have interrupted this process of learning, as onsite visits and face-to-face meetings of bankers and customers were less possible. And we expect that private information acquisition will be faster for banks that are smaller and less capitalized because loan officers` and institutional incentives are greater (e.g., Stein (2002)) and for firms that are smaller and riskier because it is easier and more necessary to do so.

To best depict the conjectures and understand how bank and firm characteristics affect the banks' private information production process, we provide, using the regression analyses with interaction terms, graphic illustrations of the banks' learning process in 3-D settings. Importantly, this shows that banks' learning process is highly non-linear and varies significantly across firms and banks of different characteristics. For example, although banks have limited and often negative private information about borrowers that are located further away at the inception of the relationship, this is offset as more private information is collected for distant borrowers as the relationship lengthens. More importantly, the private information collected is largely positive in nature. Similarly, we find that relationship length particularly increases the depth measure and favorably affects both direction measures during non-COVID times for smaller and less capitalized banks and for larger and leveraged firms.

The rest of the paper proceeds as follows. Section II discusses the literature and the paper's contributions. Section III introduces the methodology. Section IV introduces the data. Section V reports the main findings. Section VI concludes.

II. Literature and Contributions

With its findings on what drives the variations in the depth and direction of banks' private information, our paper contributes to three strands of the literature: on internal bank credit ratings, on bank-firm relationships, and on bank-firm distance.

A. Internal bank credit ratings

For publicly-listed firms, a vast literature has analyzed the information content of credit ratings (e.g., Hirtle (2006), Iannotta (2006), Livingston, Naranjo, and Zhou (2007), Bannier, Behr, and Güttler (2010), Iannotta (2011), Jones, Lee, and Yeager (2012), Flannery, Kwan, and Nimalendran (2013), or King, Ongena, and Tarashev (2020). In general, such ratings are shown to be somewhat informative. Hand, Holthausen, and Leftwich (1992), Ederington and Goh (1998) and Kliger and Sarig (2000) for example show that rating changes can explain stock and bond returns of non-financial borrowers, with Sironi (2003), Cavallo, Powell, and Rigobon (2013) and Correa, Lee, Sapriza, and Suarez (2014) finding similar effects for banks.

Regarding internal ratings of private firms, the literature is much smaller. Nakamura and Roszbach (2018) for example use credit rating data from two large Swedish banks to elicit evidence on banks' loan monitoring ability (see also Carling, Jacobson, Lindé, and Roszbach (2007)). Their analyses reveal that banks' internal credit ratings indeed include valuable private information from monitoring, which in their setting increases with the size of loans. Surprisingly, they also show that information from a credit bureau, which is in principle publicly available, is not efficiently impounded in the bank ratings and that this inefficiency is greater for smaller loans, consistent with bank loan officers placing too much weight on their private information, which they deem a form of overconfidence.

B. Bank-firm relationships

Our findings on banks` private information improving in quality during a bank-firm relationship are entirely consistent with seminal theories on how to help overcome informational challenges by Fischer (2001), Sharpe (1990), Rajan (1992), von Thadden (2004), and Hauswald and Marquez (2006), among others.[9] As hypothesized by these theories, private repayment and other information on firms collected by incumbent banks during a relationship generates informational advantages. Loan repayments, for example, are thought to allow the "inside banks" to distinguish high- from low-quality firms, with low-quality firms more likely to switch to "outside banks." As a test of the presence of banks' private information, Beyhaghi et al. (2024) study how changes in losses privately expected by banks predict firms' future stock returns, bond returns, and earnings surprises. Yet they do not study the process of learning, which is at the center of our analysis, and they focus on the smaller sample of publicly-listed firms, whereas we exclusively study the very large sample of private firms.

[9] Boot (2000), Ongena and Smith (2000), Berger and Udell (2002), Elyasiani and Goldberg (2004), Degryse and Ongena (2008), Degryse, Kim, and Ongena (2009), Degryse, Ioannidou, and Ongena (2015), Duqi, Tomaselli, and Torluccio (2018), Degryse, Morales-Acevedo, and Ongena (2019), Bonfim, Nogueira, and Ongena (2021), among others, review (parts) of this literature.

While other papers explore the impact of relationship duration on the level of loan rates (and other loan contract terms),[10] few papers focus on these factors` direct impact on the quality of the information (Cerqueiro, Degryse, and Ongena (2011) analyses effects on loan rate depth). What is new in our paper is that we focus on specific measures of the quality of banks' private information, i.e., its depth as well as its direction, based on internal bank credit ratings, also showing that these measures map into the terms of loans, and how they are affected by relationship length and its interactions with distance, and bank and firm characteristics.

Our findings on how the depth and direction of private information change with length of relationships and physical distance suggest that banks can overcome through longer relationships to some extent the informational challenges posed by distance and modus communicandi (e.g., during COVID), and can tailor their learning processes to bank and firm characteristics.[11]

C. Distance and bank, firm and other characteristics

We also contribute to a growing empirical literature that has documented that the intensity of distance-related credit rationing affecting firms may vary by bank, country, period, governmental lending programs, transportation infrastructure, and/or the characteristics of

[10] See, e.g., Ioannidou and Ongena (2010), Barone, Felici, and Pagnini (2011), Stein (2015), Xu, Saunders, Xiao, and Li (2020), Bonfim et al. (2021), Cao, Garcia-Appendini, and Huylebroek (2024) and Di, Ongena, Qi, and Yu (2024). See Kysucky and Norden (2016) for a meta-analysis of earlier reduced-form findings.

[11] Plosser and Santos (2018) for example show that bank capital affects the probability of default reported by each bank (among a sample of at most 15 banks) for about 75,000 syndicated term loans or revolver credits with at least two banks between 2010Q1 and 2013Q3 (as reported in the Shared National Credit Program, with an average commitment of around $20 million). We confirm this specific finding but extend it in several ways by: studying the impact of relationships on both the depth and direction of private information (as present in ratings that are standardized across banks); analyzing the role played by firm, bank, and bank-firm characteristics in the learning process; and using a larger sample of around 3,400,000 loan-firm-bank-quarter observations over the period 2012M9 to 2021M3.

local (bank) competitors.[12] And Degryse, Laeven, and Ongena (2009) show that the lending bank's geographical reach is determined not only by its own organizational structure but also by organizational choices made by its rivals. They find that the geographical footprint of the lending bank is smaller when rival banks are relatively larger and more hierarchically organized (and may rely relatively more on hard information).

We contribute to this literature by highlighting the importance of distance, and bank and firm characteristics on the nature of private information acquisition and learning as a potential explanation for the observed phenomena.

Our findings also have implications for the way studies could be conducted. Specifically, they indicate that bank internal ratings are less favorable for distanced firms especially at the beginning of the bank-firm relationship. This implies that in reduced-form regressions of the loan rate on a set of variables that include both distance and rating (e.g., Agarwal and Hauswald (2010)),[13] the latter may bias the coefficient estimate of the former leading to a possible underestimation of the importance of distance (and related transportation and communication costs) for loan pricing.

[12] Degryse and Ongena (2005) and Degryse and Ongena (2007) document how the intensity of credit rationing in Belgium relates to distance. In contrast, Carling and Lundberg (2005) and Uchida, Udell, and Watanabe (2008) document the absence of distance-related credit rationing in Sweden and Japan. Petersen and Rajan (2002) and Agarwal and Hauswald (2010) indicate that the distance effect may be economically rather small in the United States (and distances correspondingly large). Interestingly, the distance between banks and borrowing firms varies substantially over the financial cycle (in the US in Granja, Leuz, and Rajan (2022)) and may be affected by governmental lending programs (in the US, the Small Business Administration Preferred Lenders Program in Gupta and Ongena (2022)) and road infrastructure improvements (in Norway Herpfer, Schmidt, and Mjøs (2022)).

[13] Loan rates are regressed on distance in, e.g., Petersen and Rajan (2002), Degryse and Ongena (2005), and Herpfer et al. (2022).

III. Methodology

To identify the determinants of the depth and better/worse assessments of bank credit ratings, we employ a two-stage regression model which is inspired by a model with multiplicative heteroskedasticity as introduced by Harvey (1976). We introduce this model in Appendix A.1.

In essence we estimate the parameters in a first-stage equation by OLS and use the squared errors as raw estimates of the individual variances. Then, one obtains estimates of the parameters in the second stage equation by regressing the (logarithm of) squared errors, *Ln(Residual squared)*, which we call *Depth*, on a set of covariates.

We also define dependent assessment variables, *Residual if Residual < 0* and *Residual if Residual > 0*, which we regress on the same set of covariates. The former, which we call *Better Private Information* is the (absolute value of the) estimated residual from the first stage equation when it is less than 0 and equals 0 otherwise. *Worse Private Information* is the estimated residual when it is more than 0 and equals 0 otherwise. While *Depth* tells the overall magnitude of private information, *Better Private Information* and *Worse Private Information* capture the degree of the (un)favorable nature of the information.

Figure 1 provides an illustrated example conveying the main intuition of these measures for the variable (physical) distance between the bank and the firm (the intuition applies similarly to other bank and firm characteristics). Merely as a striking example, we consider Firm number 5 to be located quite a distance from Bank 3, maybe a bit closer to Bank 2, but even further from B1. Across all banks and firms, the rating is increasing (i.e., worsening) in observable distance along an estimated line which is depicted in dashed gray. This is captured in the first stage. All banks give the firm F5 a different rating (hence it is dispersed), but all ratings are located above the estimated line. Hence the depth of private information by these three banks about firm F5 is high and the overall assessment across banks is worse than what the first stage estimation suggests in terms of rating based on all included observables. This is captured in the second stage.

In terms of summary statistics, we glean from Appendix Table A.2 that the *Internal bank rating of the firm* ranges between 1 and 10, with a mean (median) equal to 4.98 (5) and a standard deviation equal to 1.03. The mean (median) residual equals 0.00 (-0.06) with a standard deviation equal to 0.96. This translates into the *Depth of Private Information* variable having a mean of -2.13 (-1.62) and a standard deviation of 2.52. The variables *Better and Worse Private Information* both have means (medians) of 0.30 (0.00) and standard deviations of 0.45 and 0.58, respectively.

A. Banks' credit assessment and internal loan rating

Banks' internal loan ratings are central to understanding how private information is collected and utilized in lending decisions. As part of its initial lending and ongoing monitoring processes, a bank assesses the credit quality of its borrowers for which it typically uses an internal credit risk rating scale. These ratings, assigned by loan officers, reflect the bank's internal credit risk assessment of borrowers, incorporating both quantitative metrics such as borrowers' financial information observed and qualitative judgments. As such, they offer a direct and detailed measure of the private information banks gather, making them essential for analyzing how banks mitigate information asymmetries in lending to private firms.

To examine these dynamics, we utilize data from the Y-14Q regulatory filings, a comprehensive dataset collected by the Federal Reserve. This dataset includes banks' internal ratings, standardized by the Federal Reserve to a common scale, allowing for consistent comparisons across institutions and over time. The rating scale ranges from 1 (= best) to 10 (= worst).[14] The internal credit risk rating allows then for cross-bank comparison of private information embedded. Such a straightforward comparison is not possible in other bank or loan data sets such as the Call Reports or DealScan. Leveraging this unique data source enables us to study banks' learning of private information within firm-bank relationships and its impact on loan terms.

B. Data on loans, relationships, and banks

Along with internal rating variables, we gather additional loan-level information from the Federal Reserve's Y-14Q report. This report, which has been collected quarterly since the fall of 2011, provides data from reporting banks regarding their commercial and industrial (C&I) loans.[15] Our data cover all C&I loans in size more than $1 million when originated and held by the largest bank holding companies in the U.S. by assets.[16] Loan sizes range from the $1 million reporting threshold (in commitment) to billions of dollars, thus covering loans to a rich spectrum of middle-market privately held firms of various sizes (as noted, we exclude all publicly listed firms from our base analysis). Each loan-level observation contains the issuing bank's internal rating of the borrower and various loan characteristics (e.g., committed amount, interest rate spread, and maturity). The dataset also includes extensive information

---

[14] In the context of the Y-14Q data used in the supervisory stress tests, the Federal Reserve receives banks' concordance maps that translate their internal ratings to a common S&P-like rating scale (the scale is the one required by the Fed, and the Fed can also unilaterally make certain minor adjustments). The mapping is done on the basis of probability of default (PD). For example, if a class of loans with internal ratings of "3" from bank 1 and a class of loans from bank 2 with internal ratings of "b5" have similar PD and are within the PD range of public-rated A ("single A") loans, then the internal ratings of both classes of loans would map to a converted external rating of A. It is possible that banks use the same concordance map when they communicate with market participants.

[15] More information on the data, including sources and definitions, is provided in Appendix Table A1. Appendix Table A2 provides the summary statistics for all variables.

[16] The number of reporting banks varies over our sample period between 27 and 33.

collected by the lending bank(s) on firm financials and performance, including total assets, ROA, and leverage as well as the identification of the borrower, allowing us to calculate the distance between the bank branch (or HQ) and the firm. As noted, no firm information is shared with other banks.

The Y-14Q data not only includes information on loans that are newly originated, but also tracks the (changes in) characteristics of the loans and of their related borrowers over time. Our sample contains loan-level observations over the period from September 2012 (when data quality converges) to March 2021. For each quarter, we consider loans recorded on banks' balance sheets and apply the following filters to provide a clean sample. We eliminate all loans to other financial institutions and governments (NAICS codes 52 and 92). We also drop loans with a committed exposure below $1 million, the official minimum size requirement to be included in the Y-14Q. Schedule H.1 explicitly excludes "small business loans", that are evaluated based on borrower, not the firm, credit quality or rated on a different scale than other corporate loans. For consistency, we drop all loans reported with "a small business." Observations are deleted as well if the total size of the loan package is larger than the size of the firm, or if the maturity of the loan is negative. Our final regression sample contains over 3.4 million loan-firm-bank-quarter observations.[17] As of 2022:Q4, these loans covered more than 70 percent of the balances of all commercial and industrial (C&I) loans as reported in the Federal Reserve's Y-9C.[18]

[17] Credit ratings frequently change as the bank-firm relationship goes on (and provided the bank-firm exposure remains positive). In our case, there are rating changes for 2.55 million loan-year-quarter observations (i.e., 73.9 percent of all observations in this category).
[18] Federal Reserve Board, Form FR Y-9C, provides the Consolidated Financial Statements for Bank Holding Companies.

In addition, we collect data on bank characteristics from FR Y-9C data. The Y9-C data, which is publicly available, contains a quarterly balance sheet and income statement information — including bank age, size, liquidity, profitability, and capital ratio — for U.S. holding companies and the branches of foreign companies that operate in the U.S.

IV. Findings

A. First stage equation

The dependent variable in the first stage equation is the *firm's Internal bank rating*, which reflects the bank's credit risk assessment of the firm, as sourced from Y-14Q. The first stage equation is mainly there to predict — as good as possible — the bank rating based on information that is observable in principle, including financial information of the borrower firms and the bank that makes the loan.[19] We do this for a combined sample of the initial ratings, i.e., when the bank deals with new customers, defined as those for which the length of the bank-firm relationship is less than a quarter (of a year), and the subsequent ratings. We include bank-firm distance and a comprehensive set of bank and firm controls, disregarding, for example, potential multicollinearity and bad control issues.

As bank controls, the first stage equation includes: *Ln(Bank assets)*, *Bank equity ratio*, *Bank NPL ratio*, *Bank liquid asset ratio*, and *Bank ROA*. As firm controls, it includes: *Ln(Firm assets)*, *Firm ROA*, *Firm leverage*, and two indicator variables for whether the firm is *Green* or *Brown* (recall that Appendix Table A1 contains the precise definitions of all variables).

[19] Banks vary in how they rate specific classes of firms depending on their business model (e.g., some banks specialize in large or small firms or in firms from a specific sector), see, e.g., Plosser and Santos (2018) and Blickle, He, Huang, and Parlatore (2024) for further discussions.

Since the fitted estimates reflect all quantitative information that can, in principle, be collected by any bank, any bank's private assessment of the firm, derived from the specific approach they employ to reach the assessment, is captured in the residual and thus deemed banks' private information (since, as noted, the financial information is not shared among the banks, this favors a more conservative measure of private information).

The estimated coefficients are in Table 1 Model (1). The sign and size of these estimated coefficients are straightforward and reasonable. Better internal ratings (i.e., lower numbers, with 1 = best and 10 = worst) are granted by banks in closer proximity to the firm, by small, leveraged, low performing, illiquid, or less profitable banks and assigned to large, profitable, lowly leveraged, or green firms. Though not the focus of our investigation, these estimates are similar to those already established in the literature and/or intuitive. For example, weaker banks are shown to give better ratings overall (e.g., Plosser and Santos (2018)) and maybe not surprisingly banks grant better ratings to better quality borrowers (e.g., Nakamura and Roszbach (2018)).

B. First stage robustness: changes in the set of controls and fixed effects

So far, we have erred on the side of including as much information as we have available as econometricians. We assumed for example that the firm financial data is available to all banks. But, as noted, in reality, only the lending bank has collected and reported to the Fed the data of the specific firm it lends to. However, when we exclude all financial information and just include firm sector and green/brown type dummies, it turns out our main findings are mostly unaffected by this and other such modulations in the set of controls.[20]

Another way to assess private information — and its relationship with loan terms — is to focus on the smaller sample of firms that borrow from more than one bank. This allows us to comprehensively control with firm-time fixed effects for all time-varying observable and unobservable firm characteristics.[21] However, this level of saturation with firm-time fixed effects is arguably too comprehensive for our objective. For example, assume that for a specific firm, all (or most) banks collect publicly-available information on, say, its CEO that is mainly positive (compared to what the first stage equation based on hard information would tell us). And conversely, for another firm most such information collected is negative. But the firm-time fixed effects negate either of these common findings as "it forces" on average some lending banks to have positive and the other lending banks to have negative private information on each of these firms. To put it differently, this approach ignores the fact that the private information collected by any one individual bank for one firm may be correlated with the private information collected by the other banks during their relationships (also

[20] This should not come as a surprise given that Cerqueiro, Degryse, and Ongena (2013) for example points out that the estimates of the coefficients in the second stage equation are often surprisingly unaffected by changes in the set of variables included in the first stage equation, because the maximum likelihood estimators for the parameters in the first stage and second stage equations are, from a theoretical perspective in expectation, uncorrelated (see Harvey (1976)). This finding also pertains to the bank and firm variables which, in any case, are rather standard when explaining credit ratings (e.g., Altman (1968)). Altering this set also leaves the second stage equation estimates mostly unaffected.

[21] In different contexts, firm*time fixed effects have been deployed extensively in the literature to estimate the impact of monetary, regulatory or other shocks on banks` supply of credit and disentangle it from the demand for credit by assuming homogeneity of loan demand across banks (Khwaja and Mian (2008)). For discussion and further methodological developments on this account see also, e.g., Jiménez, Ongena, Peydró, and Saurina (2012), Jakovljević, Degryse, and Ongena (2015), Degryse, De Jonghe, Jakovljevic, Mulier, and Schepens (2019), De Jonghe, Dewachter, Mulier, Ongena, and Schepens (2020), Greenstone, Mas, and Nguyen (2020) and Berg, Reisinger, and Streitz (2021).

because relationship and calendar time are inevitably correlated). We return to the weaknesses of this approach once more in Table 2.

C. Second stage: private information at inception

Using the estimated residual between the actual ratings and fitted ratings predicted by observables from the first stage in Model (1) in Table 1, we construct our three variables to capture the content of banks' private information. We next relate the three private information content measures at the beginning of the lending relationship to initially known factors, thus abstracting from any influences due to learning. We regress the three measures we constructed on several factors including the geographic distance between bank and firm and their characteristics.

The dependent variable in Table 1 Model (2) is *Depth*, and in Models (3) and (4), the "direction equations", *Better* or *Worse Private Information*, respectively. With the residuals directional, i.e., greater *Better* or *Worse* are both higher values, positive estimated coefficients imply that the specific factor adds to the private information, i.e., betters it, relative to that estimated using observable variables, in Model (3) and conversely in Model (4) worsens it.

Most estimates of the distance, bank and firm coefficients for the *Depth* and *Direction* equations, Models (2) to (4), are intuitive. The estimates on *Distance bank HQ to firm* in Models (2) to (4) equal -0012, -0.007***,[22] and 0.001, respectively, implying that distance lowers *Better Private Information* (i.e., lower ratings), with no statistically significant effects on *Depth* and *Worse* ratings. These estimates imply that an increase in (log) distance from zero to the median (i.e., 6.461 or 638 miles) reduces better ratings by 4.5 percentage points

[22] As in the Tables we indicate statistical significance in the text as follows: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

(pp) (= 0.007 * 6.461), or about 10 percent of the standard deviation of better ratings (= 0.045 / 0.451). Hence, physical distance decreases in economically significant manners initial private information, leading to lower favorable ratings. This finding likely reflects a combination of the greater unfamiliarity of banks with such borrowers and the higher costs of collecting information.

We also find statistically significant roles for bank characteristics, consistent with both differences in business models and banks' financial situation and performance playing roles in the production of private information. Specifically, large, leveraged, low-performing, illiquid, or unprofitable banks display more private information in their ratings.

Firm characteristics are also relevant, not surprisingly. Take for example firm size. In Model (2) the estimated coefficient equals 0.111***, which implies that a one standard deviation increase in *Ln(Firm assets)* increases *Depth* by 0.21 (= 0.111 * 1.932) or around 8 percent of the standard deviation (= 0.21 / 2.522), in Model (3) the estimate of -0.010*** implies that size "lowers" *Better Private Information* by 0.019 (= -0.010 * 1.932) or 4 percent of the standard deviation (= 0.019 / 0.451). And in Model (4) the estimate of 0.038*** implies that it "increases" *Worse Private Information* by 0.073 (= 0.038 * 1.932) or 13 percent of the standard deviation (= 0.073 / 0.577). In other words, banks initially attain more depth in their private information for those private firms that are larger. And the banks' extra private information makes them improve their ratings when dealing with the "better" larger firms (recall that a lower rating is "better", i.e., maps into a lower probability of default), and increase their ratings when dealing with the "worse" larger firms.

These results suggest that the banks studied here (which, recall, are among the larger ones in the US) initially amass more private information on larger firms and then act upon it by adjusting their rating direction, while they assign more “cookie cutter” (and hence based on observables more explainable) ratings for smaller firms (as in, e.g., Cole, Goldberg, and White (2004)). However, as we will see below, relatively there is more subsequent “learning” for the smaller firms than for the larger firms. We also show that banks acquire more private information over time when dealing with profitable, leveraged or other-than-green firms, for which it may be easier to do.

D. Private information and loan terms

To confirm that our measures reflect banks’ credit assessments, we investigate how banks’ loan terms vary with our three measures of bank private information. In Table 2 Panel A, we report the regression results for four key loan terms, i.e., the Loan interest rate spread, the Ln(Loan maturity), the Ln(Loan amount), and d(Collateralized), on our three information measures. All the regressions include the usual set, i.e., bank-firm distance, bank and firm characteristics as controls, as well as bank and industry fixed effects, and a constant. Note that we use here again all observations after the loan was initiated, now to allow for the loan terms to be adjusted as bank and firm characteristics change and the relationship evolves. The overall sample size is somewhat smaller for the spread regression (as Y-14Q does not report the interest rate on undrawn credit) but otherwise it is very similar to the full sample, so results do not reflect the sample choices.[23]

[23] As a robustness, we ran the other three regressions for the same sample for which we have the spread data and the results are very similar.

The estimates show that when lending, greater depth comes with a higher interest rate spread. This finding reflects the fact that when faced with a greater uncertainty about the borrower's quality, the bank charges a higher risk premium.

Directionally, the measures of private information correlate even more strongly to loan terms. A more positive assessment comes with a lower spread, longer maturity and less collateralization, all as expected. A more negative assessment comes with both a higher spread and a greater likelihood of collateralization and a lower loan maturity and amount, all again as would be expected if the bank acts in accordance with the private information measures we constructed. The correlations are the largest for the spread charged, which is lower (higher) by 9.6 (11.3) percent of its standard deviation for one standard deviation higher positive (negative) assessment. Next important are the correlations for collateralization, which imply a 6.1 percent lower and a 2.7 percent higher chance of collateralization upon a similarly defined difference in positive and negative assessment, respectively. The correlations for the other dependent variables, loan maturity and amount, are economically much smaller, 2 percent or less equivalently expressed. Overall, these associations confirm that our private information measures are meaningful representations of the banks' lending procedures.

Next in Table 2 Panel B, we study only multi-bank firms and replace Industry with Firm * Year fixed effects. We now effectively compare at a given point in time the implication of private info sets contained in loans to the same borrower by distinct banks, who differ in their relations. Recall our earlier deliberations — above — on the excessiveness of saturating with firm-time fixed effects. In the estimates reported in Panel B we observe this argumentation to be potentially valid, as our private information measures now explain less of the loan terms, in economic terms only around a quarter or so of the effect explained in Panel A.

E. Private information and lending relationships

To analyze the learning process, we next study the evolution of our three private information measures, *Depth* and *Better* and *Worse* assessments, after the loan has been made, i.e., after the 1st quarter. In Table 3, Models (1) to (6), we provide our full regression results, which include the distance between the bank and the firm as well as the usual bank and firm controls. Given our main interest, we focus on the estimated coefficients on *Length bank-firm relationship*, which in Models (1) to (3) is the number of years (divided by 1,000) the bank has been lending to the specific firm, and in Models (4) to (6), dummies for three buckets for the length of the bank-firm relationship (0.25 << 3 years; 3 << 5 years; and > 5 years). The latter specification allows for the effects of a relationship to vary by period of length.

The estimates on the simple *Length bank-firm relationship* (Models (1) to (3)) equal 3.693***, 6.760***, and -5.492***, respectively. These estimates imply that an increase in the length of the relationship from zero to the median (i.e., 4 years), increases depth by 0.014 (= 3.693 * 0.004), which is around 0.5 percent of its standard deviation (= 0.014 / 2.522), increases the likelihood of a better rating by 0.027 pp (= 6.760 * 0.004), which is around 6 percent of the standard deviation of this residual variable (= 0.027 / 0.451), and reduces the likelihood of a worse rating by 0.022 pp (= -5.492 * 0.004), which is around 4 percent of the standard deviation of this residual variable (= 0.022 / 0.577). Hence, as the relationship lengthens, banks give more weight in their internal ratings to private information and tend to rate firms better.

The estimates in Models (4) to (6) on the dummies for the length of bank-firm relationship show that there are some important non-linearities.[24] For *Depth*, the impact of relationship seems to peak between 3 and 5 years. For *Better Private Information*, there seems to be an increasing value of relationship throughout as the estimated coefficients continuously increase. For *Worse Private Information*, it appears that the impact peaks between 3 and 5 years, after which ratings marginally improve (i.e., get less worse).[25] Overall, these regression results suggest that banks proactively use their relationship to improve their assessments of the firms they lend to.[26]

Note that the estimates on *Distance bank HQ to firm* in Table 2, for both Models (1) to (3) and Models (4) to (6), now equal 0.004, 0.002*, and 0.000 respectively, which vary from those in Table 1 (which were -0012, -0.007***, and 0.001, respectively). This suggests that after the initial lending stage, having a relationship significantly influences how distance affects depth and better and worse ratings, to the point that the adverse effects of distance are mitigated or even overcome. In other words, as relationships lengthen, banks start to deviate for distant borrowers more from the hard information only compared to their initial assessment and rate these borrowers differentially.

---

[24] The estimated coefficients on the dummies capture the deviation from the bases, which is the impact on the outcome variable when the length of the bank-firm relationship is below or equal to 0.25 years. Notice that the time period potentially spent in each period bin is somewhat different, i.e., 2.75, 2, and 7 years (the maximum is 12 years), which affects the proportion of observations in each bin, which equals 2, 35, 21, and 42 percent, respectively (see Appendix A.2), and the precision of the estimates.
[25] The high share of relationships longer than 5 years (42 percent) may explain why the negative estimated coefficient on that bin mirrors the overall negative coefficient estimate on relationship length in Model (3).
[26] When we split the length of the bank-firm relationship variable more finely into six dummies capturing lengths of one, two, three, four, five, or six or more years, we confirm that the effects for depth and better private information are somewhat larger for the longer relationship lengths, whereas for worse private information, the peak impact of a relationship occurs in the fifth year.

F. Private information formation across banks and firms

We next explore how the distance between the bank and firm, the specific time period, and bank and firm characteristics, including factors such as size and riskiness, attenuate the formation of banks' private information over time through the various kinds of engagement between the bank and the firm.

In Tables 4 to 6 we report the regression results, but just in terms of the main coefficient estimates, and display corresponding figures that illustrate the main interactions of interest. Specifically, we add to the regression a variable chosen respectively among the distance between the bank and the firm, a dummy for the COVID period (2020:Q1 - 2021:Q1),[27] bank size and capital (and we briefly discuss the bank liquidity ratio), and firm size and leverage. We include every time the variable itself, the relationship length, and the interaction of the variable with relationship length. For each of the results, only three estimated coefficients are reported, that is, the coefficients for the specific characteristic, the relationship length, and their interaction. The regressions do include the usual bank-firm, bank, and firm controls but these are not reported. The sample size is the same for all regressions, i.e., 2,994,729 observations, so any variations in results do not reflect sample choices. Of most interest are the estimated coefficients for the interactions. These terms reflect how various bank or firm characteristics differently affect the formation of private information over time and thus show light on the process of bank learning. We can also assess the impact of the interaction estimates from the various accompanying figures, where, to benchmark the economic

[27] To avoid any impact of the COVID crisis, instead of using a dummy, we also employ a sample that excludes observations from the year 2020 on and find that results do not change materially.

relevance of the impact, the red arrows display 50 percent of the standard deviation of the respective outcome variable.

Several findings emerge from the analysis. In terms of distance in Table 4, we find that a longer relationship leads to more private information being collected for distant borrowers and that the information being collected turns out especially positive as the relationship lengthens. Hence a longer relationship modulates the adverse distance effects to a degree.

The graphic analysis further illustrates the differential process behind banks' private information formation across borrowers located close to versus farther away. First, we show that banks, over time, accumulate private information of greater depth, as reflected in the tilted plane moving from darker blue to bright yellow as the relationship lengthens. Importantly, the learning curve is steeper for distant borrowers who are located further away. Turning to the direction of private information, we observe that although initially being assigned a less favorable and more unfavorable rating, reflecting the lower better private information and higher worse information content, firms located further away quickly gain banks' trust as the relationship lengthens. And banks accumulate much better information and much less worse information about those borrowers located further away over time as the relationship lengthens. This result is consistent with the classic banking literature, which finds that geographic distance can create significant hurdles for banks to tackle information asymmetry problems (e.g., Dell'Ariccia and Marquez (2004)) initially. As a result, firms located further away can be subject to a more severe hold-up problem when seeking new credit from non-local banks (e.g., Hauswald and Marquez (2003)). Our finding contributes to this literature by highlighting the steep learning curve for banks *after* relationship formation

– confirming the importance of relationship lending in fostering bank learning, especially in acquiring positive private information.

Lockdown measures from the COVID-19 crisis may have hindered interactions between bank loan officers and borrowing firms, which are essential for collecting private information. Supportive of this hypothesis, as shown in Table 4 in the lower panel, during COVID times, the length of the relationship is meaningfully less important for determining the *Depth* of private information, as indicated by the large and highly significant negative coefficient for the dummy. As the size of the coefficient for the interaction is comparable in absolute size with that for the length of the relationship, there appears to have been no learning as to *Depth* during the full COVID period. Directionally, it appears that for firms with *Better* assessments, ratings did not suffer, but those with *Worse* assessments got better ratings. This suggests that during the COVID period, the combination of difficulty in meeting with the firm in person with ample general fiscal and monetary support led banks to maintain their rating for firms with a *Better* assessment but upgraded it for the other firms (potentially, also in light of the more ample support, forbearing), even though they had less, or no, interactions and *Depth* declined.[28]

In term of bank characteristics, the results in Table 5 and the graphic analysis show that relationship length adds relatively less to *Depth*, worsens *Better* assessments, and adds more to *Worse* assessments for larger and better capitalized (and in unreported regressions, more liquid) banks. In sum, smaller and lowly capitalized banks are more engaged in private information acquisition, a result that may explain the earlier findings in the literature on small

[28] Because COVID is a dummy variable, plots are not meaningful and not provided.

bank – small firm matching (e.g., Berger, Klapper, and Udell (2001); Cole et al. (2004); Uchida et al. (2008)).

Note that these effects are, as expected, asymmetrical for firms with *Better* versus with *Worse* assessments, but larger in absolute size for *Better* assessments, making overall for less favorable ratings. This suggests that large and well capitalized banks are less willing to learn about the firms they have lent to, and if they do, it is more likely to result in less favorable assessments (effects for more liquid banks are similar as for well capitalized banks, not reported).

Finally, in terms of firm characteristics, in Table 6, those typically thought to relate to private information acquisition and learning are firm size and riskiness. The size result is as expected in that as the relationship goes on, banks reduce their *Depth* and have a smaller *Worse* private assessment for firms that are larger. We use leverage to proxy firms' risks. The leverage effect is that over time banks increase the *Depth* of their assessment and raise their *Better* private view of those riskier firms that are more leveraged. Overall, these results suggest that banks are more incentivized to learn about firms and adjust their private assessment when the firm is smaller and more leveraged (and thus riskier).

Overall, the figures, for Table 6 in particular, confirm that in the relevant ranges, the magnitudes of the effects are economically meaningful and often display strongly nonlinear patterns. Notably, longer relationships much improve the *Better* scores for large and highly leveraged firms, by smaller and less leveraged and less liquid banks, and for firms that are further away. In terms of *Worse* private information, longer relationships meaningfully improve scores for large, more levered firms, to some extent for smaller and less leveraged and less liquid banks, and to some degree for firms that are further away. These associations suggest that the various types of banks make meaningfully different choices as to how to enhance their private information and they do vary this learning process by specific types of borrowers.

Together, these findings suggest that banks' business models, firm size and riskiness, as well as the geographic distance between the pairs affect how strong the influence of learning from relationship is on banks' private assessments of borrowers. A possible common thread to these findings is that the acquisition by smaller banks of private information on smaller firms is more relationship-based, i.e., making the setting more conducive to learning, whereas for larger banks and firms and banks, more transactional lending is involved, making it less subject to learning.

V. Conclusion

We document how bank-firm relationships affect the depth and degree of positiveness or negativeness in bank-specific private information about a firm's quality. We do this using a unique dataset on bank internal ratings, covering much of corporate sector lending to private firms by banks in the US over the period 2012-2021.

Our contributions are several. First, we develop new measures of the dimensions of banks' private information, i.e., the depth of banks' internal credit ratings as well as the direction of those ratings, all relative to assessments solely based on observables, and show that our three private information proxies are related to the terms of the loans granted.

Second, we show how the length of relationship impacts these dimensions and how impacts vary with bank-firm, bank, and firm characteristics. In this way, we gain additional insights as to the process of learning through relationships.

We document that increasing the length of a relationship substantially increases the depth of private information, improves positive private assessments, and reduces negative assessments. The impacts of the length of relationships peak at about five years. Importantly, effects are particularly salient at longer firm-bank distances, and during non-COVID times, and for smaller and leveraged banks, for smaller and leveraged firms. They are often strongly nonlinear and economically meaningful.

Our findings suggest that some banks have specific business models that make them more likely to invest resources to overcome the information asymmetries related to lending to private firms. Specifically, larger or highly capitalized banks accumulate much less positive information on their borrowers over time, whereas smaller or worse capitalized banks seem more willing to update their private information set in a favorable way.

Our findings also suggest that existing analyses featuring both distance and rating jointly as explanatory variables in reduced form loan rate specifications may have biased coefficient estimates.

Overall, we contribute to the literature on relationship length and distance and their effects on information asymmetries by analyzing how relationship length affects the depth and direction of bank internal ratings and how the effects of the relationship differ by bank and firm characteristics. Such analysis has not been conducted before, especially not for private firms.

References

Agarwal, S., and R. Hauswald. 2010. Distance and Private Information in Lending. *Review of Financial Studies* 23: 2757-88.

Altman, E. I. 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance* 23: 589-610.

Bannier, C. E., P. Behr, and A. Güttler. 2010. Rating Opaque Borrowers: Why Are Unsolicited Ratings Lower? *Review of Finance* 14: 263-94.

Barone, G., R. Felici, and M. Pagnini. 2011. Switching Costs in Local Credit Markets. *International Journal of Industrial Organization* 29: 694-704.

Berg, T., M. Reisinger, and D. Streitz. 2021. Spillover Effects in Empirical Corporate Finance. *Journal of Financial Economics* 142: 1109-27.

Berger, A. N., C. H. S. Bouwman, L. Norden, R. A. Roman, G. F. Udell, and T. Wang. 2021. Is a Friend in Need a Friend Indeed? How Relationship Borrowers Fare during the COVID-19 Crisis, Indiana University Kelley School of Business, Bloomington IN, Research Paper, 1.

Berger, A. N., L. F. Klapper, and G. F. Udell. 2001. The Ability of Banks to Lend to Informationally Opaque Small Businesses. *Journal of Banking and Finance* 25: 2127-67.

Berger, A. N., and G. F. Udell. 2002. Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure. *Economic Journal* 112: 32-53.

Bester, H. 1985. Screening vs. Rationing in Credit Markets with Imperfect Information. *American Economic Review* 75: 850-55.

Beyhaghi, M., C. Howes, and G. Weitzner. 2024. The Information Advantage of Banks: Evidence From Their Private Credit Assessments, Board of Governors of the Federal Reserve System, Washington DC, Mimeo.

Blickle, K., Z. He, J. Huang, and C. Parlatore. 2024. Information-based Pricing in Specialized Lending, National Bureau of Economic Research, Cambridge MA, Working Paper, 32155.

Bonfim, D., G. Nogueira, and S. Ongena. 2021. “Sorry, We're Closed" Bank Branch Closures, Loan Pricing and Information Asymmetries. *Review of Finance* 25: 1211-59.

Boot, A. W. A. 2000. Relationship Banking: What Do We Know? *Journal of Financial Intermediation* 9: 3-25.

Cao, J., E. Garcia-Appendini, and C. Huylebroek. 2024. Banking on Deposit Relationships: Implications for Hold-Up Problems in the Loan Market, Norges Bank, Oslo NO, Working Paper, 4.

Carling, K., T. Jacobson, J. Lindé, and K. Roszbach. 2007. Corporate Credit Risk Modeling and the Macroeconomy. *Journal of Banking and Finance* 31: 845-68.

Carling, K., and S. Lundberg. 2005. Asymmetric Information and Distance: An Empirical Assessment of Geographical Credit Rationing. *Journal of Economics and Business* 57: 39-59.

Cavallo, E., A. Powell, and R. Rigobon. 2013. Do Credit Rating Agencies Add Value? Evidence from the Sovereign Rating Business. *International Journal of Finance and Economics* 18: 240-65.

Cerqueiro, G., H. Degryse, and S. Ongena. 2011. Rules versus Discretion in Loan Rate Setting. *Journal of Financial Intermediation* 20: 503-29.

———. 2013. Using Heteroskedastic Models to Analyze the Use of Rules versus Discretion in Lending Decisions. In *Handbook of Research Methods and Applications in Empirical Finance.* Eds. A. Bell, C. Brooks and M. Prokopczuk. London: Edward Elgar, 216-37.

Cole, R. A., L. G. Goldberg, and L. J. White. 2004. Cookie-Cutter versus Character: The Micro Structure of Small Business Lending by Large and Small Banks. *Journal of Financial and Quantitative Analysis* 39: 227-52.

Correa, R., K.-H. Lee, H. Sapriza, and G. A. Suarez. 2014. Sovereign Credit Risk, Banks' Government Support, and Bank Stock Returns around the World. *Journal of Money, Credit and Banking* 46: 93-121.

De Jonghe, O., H. Dewachter, K. Mulier, S. Ongena, and G. Schepens. 2020. Some Borrowers Are More Equal than Others: Bank Funding Shocks and Credit Reallocation. *Review of Finance* 24: 1-43.

Degryse, H., O. De Jonghe, S. Jakovljevic, K. Mulier, and G. Schepens. 2019. Identifying Credit Supply Shocks with Bank-Firm Data: Methods and Applications. *Journal of Financial Intermediation* 40: 100813.

Degryse, H., V. Ioannidou, and S. Ongena. 2015. Bank-Firm Relationships: A Review of the Implications for Firms and Banks in Normal and Crisis Times. In *The Economics of Interfirm Networks.* Eds. W. Tsutomu, I. Uesugi and A. Ono. Berlin: Springer, 177-89.

Degryse, H., M. Kim, and S. Ongena. 2009. *Microeconometrics of Banking: Methods, Applications and Results*. Oxford University Press.

Degryse, H., L. Laeven, and S. Ongena. 2009. The Impact of Organizational Structure and Lending Technology on Banking Competition. *Review of Finance* 13: 225-59.

Degryse, H., P. Morales-Acevedo, and S. Ongena. 2019. Competition in Banking. In *Oxford Handbook of Banking*. Eds. A. N. Berger, P. Molyneux and J. O. S. Wilson. Oxford: Oxford University Press, 776-813.

Degryse, H., and S. Ongena. 2005. Distance, Lending Relationships, and Competition. *Journal of Finance* 60: 231-66.

———. 2007. The Impact of Competition on Bank Orientation. *Journal of Financial Intermediation* 16: 399-424.

———. 2008. Competition and Regulation in the Banking Sector: A Review of the Empirical Evidence on the Sources of Bank Rents. In *Handbook of Financial Intermediation and Banking.* Eds. A. V. Thakor and A. W. A. Boot. Amsterdam: Elsevier, 483-554.

Dell'Ariccia, G., and R. Marquez. 2004. Information and Bank Credit Allocation. *Journal of Financial Economics* 72: 185-214.

Di, G., S. Ongena, S. Qi, and Y. Yu. 2024. "Time for a Change of Scenery": Loan Conditions When Firms Switch Bank Branches, Centre for Economic Policy Research, London UK, Discussion Paper, 18985.

Diamond, D. W. 1984. Financial Intermediation and Delegated Monitoring. *Review of Economic Studies* 51: 393-414.

Duqi, A., A. Tomaselli, and G. Torluccio. 2018. Is Relationship Lending Still a Mixed Blessing? A Review of Advantages and Disadvantages for Lenders and Borrowers. *Journal of Economic Surveys* 32: 1446-82.

Ederington, L. H., and J. C. Goh. 1998. Bond Rating Agencies and Stock Analysts: Who Knows What When? *Journal of Financial and Quantitative Analysis* 33: 569-85.

Elyasiani, E., and L. G. Goldberg. 2004. Relationship Lending: A Survey of the Literature. *Journal of Economics and Business* 56: 315-30.

Fama, E. F. 1985. What's Different about Banks? *Journal of Monetary Economics* 15: 5-29.

Faria-e-Castro, M., P. Paul, and J. M. Sánchez. 2021. Evergreening, Federal Reserve Bank of St. Louis, St. Louis MO, Working Paper, 2021-012.

Fischer, K. H. 2001. Banken und unvollkommener Wettbewerb. Empirische Beitrage zu einer Industrieokonomik der Finanzmarkte, Goethe University Frankfurt, Frankfurt, PhD Dissertation.

Flannery, M. J., S. H. Kwan, and M. Nimalendran. 2013. The 2007–2009 Financial Crisis and Bank Opaqueness. *Journal of Financial Intermediation* 22: 55-84.

Granja, J., C. Leuz, and R. G. Rajan. 2022. Going the Extra Mile: Distant Lending and Credit Cycles. *Journal of Finance* 77: 1259-324.

Greenstone, M., A. Mas, and H.-L. Nguyen. 2020. Do Credit Market Shocks Affect the Real Economy? Quasi-experimental Evidence from the Great Recession and "Normal" Economic Times. *American Economic Journal: Economic Policy* 12: 200-25.

Gupta, M., and S. Ongena. 2022. The Impact of the SBA Funding Programs on the Distance and Pricing of Loans to Small Businesses, Nottingham University Business School, Nottingham UK, Mimeo.

Hand, J., R. M., R. W. Holthausen, and R. W. Leftwich. 1992. The Effect of Bond Rating Agency Announcements on Bond and Stock Prices. *Journal of Finance* 47: 733-52.

Harvey, A. C. 1976. Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica* 44: 461-65.

Hauswald, R., and R. Marquez. 2003. Information Technology and Financial Services Competition. *Review of Financial Studies* 16: 921-48.

———. 2006. Competition and Strategic Information Acquisition in Credit Markets. *Review of Financial Studies* 19: 967-1000.

Herpfer, C., C. Schmidt, and A. Mjøs. 2022. The Causal Impact of Distance on Bank Lending. *Management Science* 69: 723-40.

Hirtle, B. 2006. Stock Market Reaction to Financial Statement Certification by Bank Holding Company CEOs. *Journal of Money, Credit and Banking* 38: 1263-91.

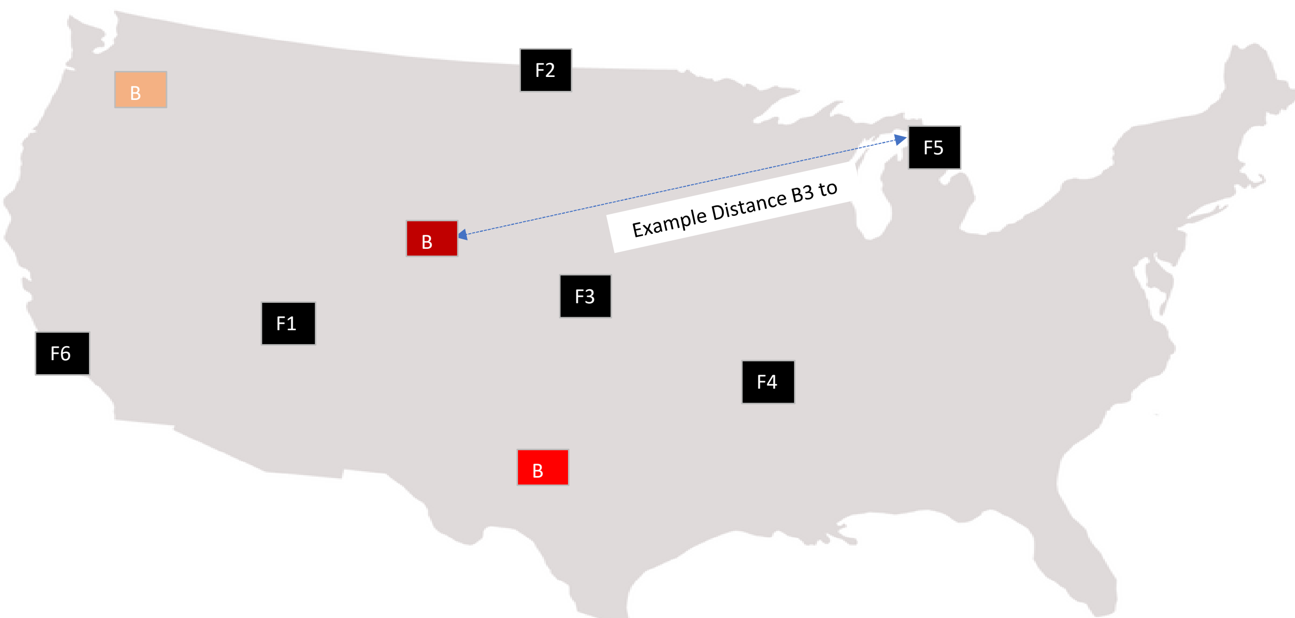
Iannotta, G. 2006. Testing for Opaqueness in the European Banking Industry: Evidence from Bond Credit Ratings. *Journal of Financial Services Research* 30: 287-309.

———. 2011. Market Discipline in the Banking Industry: Evidence from Spread Dispersion. *European Journal of Finance* 17: 111-31.
Ioannidou, V. P., and S. Ongena. 2010. "Time for a Change": Loan Conditions and Bank Behavior When Firms Switch Banks. *Journal of Finance* 65: 1847-78.
Jakovljević, S., H. Degryse, and S. Ongena. 2015. A Review of Empirical Research on the Design and Impact of Regulation in the Banking Sector. *Annual Review of Financial Economics* 7: 423-43.
Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina. 2012. Credit Supply and Monetary Policy: Identifying the Bank Balance-Sheet Channel with Loan Applications. *American Economic Review* 102: 2301-26.
Jones, J. S., W. Y. Lee, and T. J. Yeager. 2012. Opaque Banks, Price Discovery, and Financial Instability. *Journal of Financial Intermediation* 21: 383-408.
Khwaja, A. I., and A. Mian. 2008. Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market. *American Economic Review* 98: 1413-42.
King, M. R., S. Ongena, and N. Tarashev. 2020. Bank Standalone Credit Ratings. *International Journal of Central Banking* 16: 101-44.
Kliger, D., and O. Sarig. 2000. The Information Value of Bond Ratings. *Journal of Finance* 55: 2879-902.
Kysucky, V., and L. Norden. 2016. The Benefits of Relationship Lending in a Cross-Country Context: A Meta-Analysis. *Management Science* 62: 90-110.
Leland, H. E., and D. H. Pyle. 1977. Informational Asymmetries, Financial Structure, and Financial Intermediation. *Journal of Finance* 32: 371-87.
Liberti, J. M., and M. A. Petersen. 2019. Information: Hard and Soft. *Review of Corporate Finance Studies* 8: 1-41.
Livingston, M., A. Naranjo, and L. Zhou. 2007. Asset Opaqueness and Split Bond Ratings. *Financial Management* 36: 49-62.
Morgan, D. 2002. Rating Banks: Risk and Uncertainty in an Opaque Industry. *American Economic Review* 92: 874-88.
Nakamura, L. I., and K. Roszbach. 2018. Credit Ratings, Private Information, and Bank Monitoring Ability. *Journal of Financial Intermediation* 36: 58-73.
Ongena, S., and D. C. Smith. 2000. Bank Relationships: a Review. In *The Performance of Financial Institutions.* Eds. P. Harker and S. A. Zenios. London: Cambridge University Press, 221-58.
Petersen, M. A., and R. G. Rajan. 1994. The Benefits of Lending Relationships: Evidence from Small Business Data. *Journal of Finance* 49: 3-37.
———. 2002. Does Distance Still Matter? The Information Revolution in Small Business Lending. *Journal of Finance* 57: 2533-70.
Plosser, M. C., and J. A. C. Santos. 2018. Banks' Incentives and Inconsistent Risk Models. *Review of Financial Studies* 31: 2080-112.
Rajan, R. G. 1992. Insiders and Outsiders: the Choice between Informed and Arm's-Length Debt. *Journal of Finance* 47: 1367-400.
Sharpe, S. A. 1990. Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships. *Journal of Finance* 45: 1069-87.
Sironi, A. 2003. Testing for Market Discipline in the European Banking Industry: Evidence from Subordinated Debt Issues. *Journal of Money, Credit and Banking* 35: 443-72.
Stein, I. 2015. The Price Impact of Lending Relationships. *German Economic Review* 16: 367-89.
Stein, J. C. 2002. Information Production and Capital Allocation: Decentralized versus Hierarchical Firms. *Journal of Finance* 57: 1891-922.
Uchida, H., G. F. Udell, and W. Watanabe. 2008. Bank Size and Lending Relationships in Japan. *Journal of the Japanese and International Economies* 22: 242-67.
Veronis, N. 2023. The Importance of Private Markets, iCapital Institutional Capital Network, Inc., New York NY, Blog, September 26.
von Thadden, E.-L. 2004. Asymmetric Information, Bank Lending, and Implicit Contracts: the Winner's Curse. *Finance Research Letters* 1: 11-23.
Xu, Y., A. Saunders, B. Xiao, and X. Li. 2020. Bank Relationship Loss: The Moderating Effect of Information Opacity. *Journal of Banking and Finance* 118: 105872.

Figure 1. Depth and Direction of Private Information

Panel A depicts example geographical distances between three banks (in various shades of red) and six firms (in black). Panel B plots on the horizontal axis the distance between banks and firms, and on the vertical axis example ratings. For the Distance from Bank 3 to Firm 5 and example arrow is placed in Panel A and Panel B to facilitate the visual mapping. The resultant distance-rating cells are in the red shades of the banks. An example line for the Rating Explained by Observables is added. For two firms, i.e., Firm 2 and Firm 5, the deviations from this Rating line are indicated with green and yellow arrows. The three banks rate Firm 2 (mostly) the same as the rating explained by observables, so lack Depth in their private information, but have (weakly) better private information, while the three banks rate Firm 5 very different from the rating explained by observables, so have Depth in their private information, but each has worse private information than is publicly observable.

Panel A. Geographical Distribution of Banks and Firms



Panel B. Standardized Bank Rating of Firm, and Depth of Private Information and Direction of Private Information

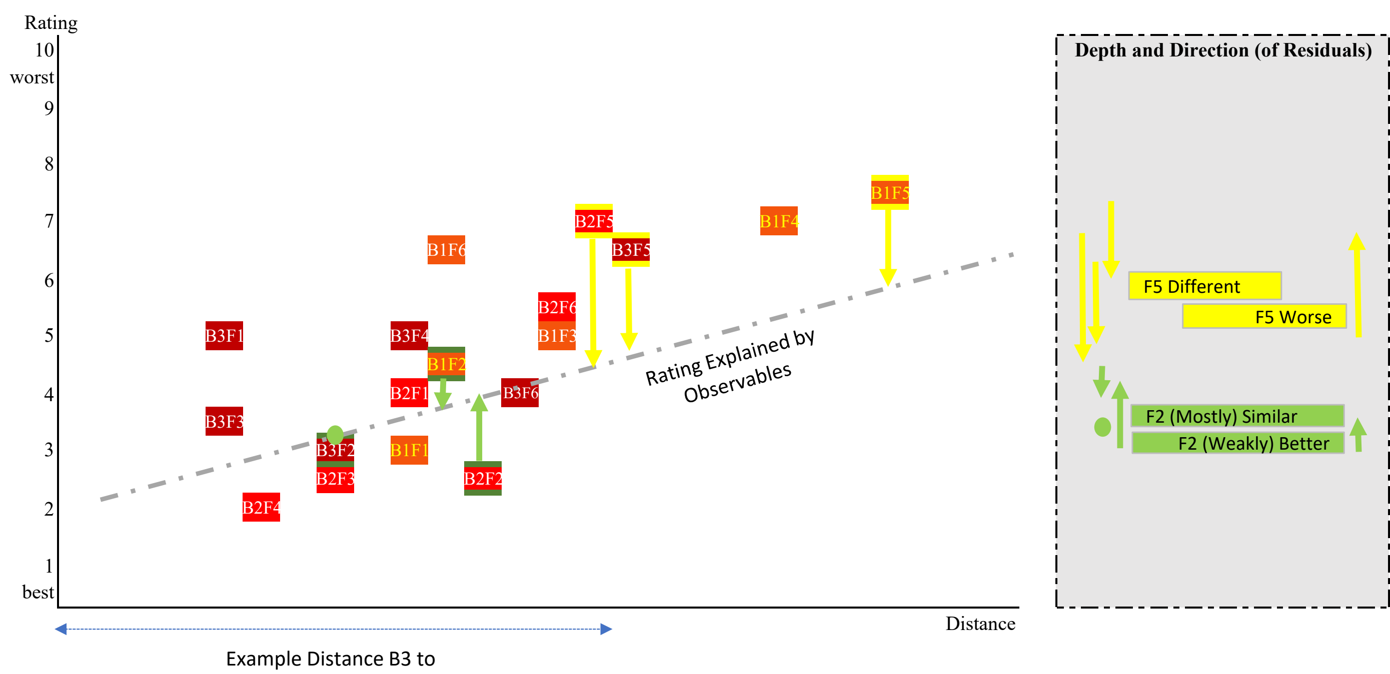


Table 1. Main Results: Bank Rating of Firm, Depth and Direction of Private Information

The table reports estimates from ordinary least squares regressions. The sample in Model (1) includes all bank ratings given to non-listed firms, in Models (2) to (4) only the bank ratings given to firms when the length of the bank-firm relationship is shorter than 0.25 years. The dependent variables are: in Model (1) the Internal bank rating of firm, which is the rating given by the bank to the firm transferred to a common scale; in Model (2) the Ln(Residual squared), which is the natural log of the squared residuals; and, in Models (3) and (4) the Residual itself, and 0 otherwise, if the residual is smaller or larger, respectively, than zero. In all cases "the residual" is the estimated residual from the first stage equation in Model (1). The definition for each independent variable is given in Table A.1. Coefficients are listed in the first row, t-statistics based on robust standard errors are reported in the row below in parentheses, and the corresponding significance levels are adjacent to the coefficient. Standard errors are clustered at the firm level. *** Significant at 1%, ** significant at 5%, * significant at 10%.

Model	(1)	(2)	(3)	(4)
Sample	***All bank ratings of firms***	***Length bank-firm relationship < 0.25 Years***		
Definition	**Internal bank rating of firm (1 = best, 10 = worst)**	**Ln(Residual squared)**	**Residual if Residual < 0**	**Residual if Residual > 0**
Dependent Variable Name	*Bank Rating of Firm*	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables				
Static Bank-Firm Variable				
Distance bank HQ to firm	0.008***	-0.012	-0.007***	0.001
	(3.68)	(-1.32)	(-4.29)	(0.37)
Bank Variables				
Ln(Bank assets)	0.099***	0.084***	0.005**	-0.003
	(36.09)	(6.68)	(2.19)	(-0.94)
Bank equity ratio	3.535***	-9.831***	1.164***	-4.271***
	(18.63)	(-10.54)	(6.76)	(-20.69)
Bank NPL ratio	-4.994***	-2.646***	1.557***	-1.057***
	(-31.15)	(-3.08)	(10.73)	(-4.84)
Bank liquid asset ratio	0.475***	-1.513***	0.401***	-0.532***
	(9.18)	(-6.23)	(7.75)	(-8.97)
Bank ROA	12.326***	-25.303***	0.580	-2.004
	(13.13)	(-3.25)	(0.44)	(-0.93)
Firm Variables				
Ln(Firm assets)	-0.110***	0.111***	-0.010***	0.038***
	(-36.30)	(18.33)	(-5.75)	(19.76)
Firm ROA	-1.942***	0.543***	0.156***	-0.012
	(-88.48)	(12.11)	(17.95)	(-1.05)
Firm leverage	0.449***	1.404***	-0.311***	0.519***
	(43.92)	(19.89)	(-28.33)	(21.90)
Green	-0.032*	-0.159**	-0.027**	-0.016
	(-1.85)	(-2.33)	(-2.25)	(-1.27)
Brown	0.057	0.140	0.107***	-0.069***
	(1.60)	(1.05)	(3.43)	(-2.78)
Observations	2,717,102	61,149	61,149	61,149
Adjusted R-squared	0.129	0.019	0.030	0.043

Table 2. Impact of Private Information on Loan Terms

This table reports OLS regression estimates to assess how Depth of Private Information and Better or Worse Private Information affect loan terms. The indicated loan term as dependent variable is regressed on one of the three private information variables, the distance between the bank headquarters, the bank variables, the firm variables, and bank and industry fixed effects in Panel A and bank and fim * year fixed effects in Panel B. The sample includes corporate loans to non-listed firms reported in the Y-14Q by bank holding companies between September 30, 2012, and March 31, 2021. Standard errors are clustered at the bank × industry level. The fourth and eighth column also report the impact of a one standard deviation change in depth or direction of private information on the loan term, in percent scaled by the standard deviation of this loan term. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

Dependent Variable / Independent variables	(1)	(2) *Loan Interest Rate Spread*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*	(1)	(2) *Ln(Loan Maturity)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Depth of Private Information	0.023***			*2.92%*	-0.001			*0.26%*
	(4.31)				(-1.04)			
Better Private Information		-0.423***		*-9.62%*		0.042***		*1.97%*
		(-22.34)				(2.80)		
Worse Private Information			0.388***	*11.28%*			-0.042***	*-2.52%*
			(25.78)				(-6.95)	
Bank-Firm, Bank and Firm Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	2,121,365	2,121,365	2,121,365		2,713,058	2,713,058	2,713,058	
Adjusted R-squared	0.112	0.121	0.128		0.283	0.283	0.283	

Dependent Variable / Independent variables	(1)	(2) *Ln(Loan Amount)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*	(1)	(2) *d(Collateralized)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Depth of Private Information	-0.001			*-0.22%*	-0.003***			*-2.75%*
	(-0.53)				(-8.18)			
Better Private Information		0.022		*-0.88%*		-0.037***		*-6.07%*
		(1.45)				(-5.68)		
Worse Private Information			-0.013*	*-0.67%*			0.013***	*2.73%*
			(-1.82)				(8.17)	
Bank-Firm, Bank and Firm Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	2,715,622	2,715,622	2,715,622		2,715,622	2,715,622	2,715,622	
Adjusted R-squared	0.393	0.393	0.393		0.148	0.151	0.148	

Panel B

Dependent Variable / Independent variables	(1)	(2) *Loan Interest Rate Spread*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*	(1)	(2) *Ln(Loan Maturity)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*
Depth of Private Information	0.002			*0.25%*	0.000			*0.00%*
	(1.60)				(0.18)			
Better Private Information		-0.121***		*-2.75%*		0.004		*0.19%*
		(-9.97)				(0.90)		
Worse Private Information			0.120***	*3.49%*			-0.013***	*-0.78%*
			(9.63)				(-2.75)	
*Bank-Firm, Bank and Firm Controls, and Bank and Firm * Year Fixed Effects*	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	1,979,232	1,979,232	1,979,232		2,553,341	2,553,341	2,553,341	
Adjusted R-squared	0.787	0.787	0.787		0.724	0.724	0.724	

Dependent Variable / Independent variables	(1)	(2) *Ln(Loan Amount)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*	(1)	(2) *d(Collateralized)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Depth of Private Information	0.003***			*-0.67%*	-0.000***			*0.00%*
	(6.43)				(-2.61)			
Better Private Information		0.054***		*-2.17%*		-0.014***		*-2.30%*
		(7.94)				(-6.25)		
Worse Private Information			-0.022***	*-1.13%*			0.008***	*1.68%*
			(-3.84)				(3.66)	
*Bank-Firm, Bank and Firm Controls, and Bank and Firm * Year Fixed Effects*	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	2,555,699	2,555,699	2,555,699		2,555,699	2,555,699	2,555,699	
Adjusted R-squared	0.696	0.696	0.696		0.678	0.678	0.678	

Table 3. Bank-Firm Relationship Length and Banks' Private Information

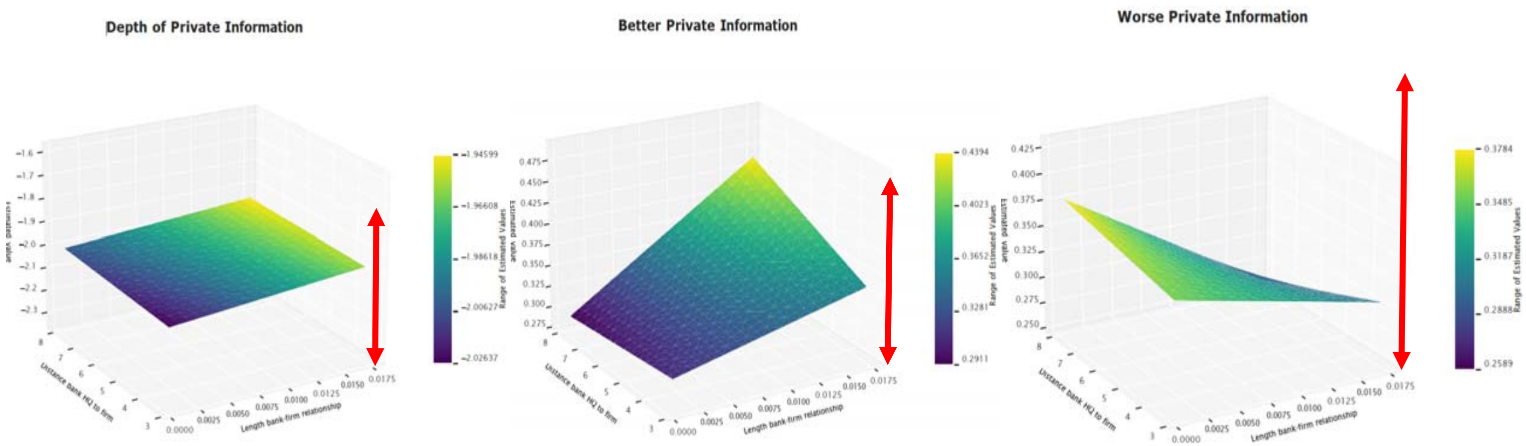
The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to non-listed firms. The number of observations equals 2,994,729. The dependent variables are: in Panel A the Depth of Private Information, which is the natural log of the squared residuals; and, in Panels B and C the Better and Worse Private Information which is equal to the absolute value of the residual, and 0 otherwise, if the residual is larger or smaller, respectively, than zero. In all cases "the residual" is the estimated residual from the first stage equation in Model (1) in Table 1. The definition for each independent variable is given in Table A.1. Coefficients are listed in the first row, t-statistics based on robust standard errors are reported in the row below in parentheses, and the corresponding significance levels are adjacent to the coefficient. Standard errors are clustered at the firm level. *** Significant at 1%, ** significant at 5%, * significant at 10%.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables						
Dynamic Bank-Firm Variable						
Length bank-firm relationship	3.693***	6.760***	-5.492***			
	(3.31)	(19.98)	(-20.95)			
Length bank-firm relationship (0.25 << 3 years)				0.176***	0.006***	0.049***
				(14.53)	(3.08)	(17.15)
Length bank-firm relationship (3 << 5 years)				0.245***	0.036***	0.054***
				(14.90)	(11.46)	(13.85)
Length bank-firm relationship (> 5 years)				0.210***	0.080***	-0.002
				(12.54)	(23.66)	(-0.60)
Static Bank-Firm Variable						
Distance bank HQ to firm	0.004	0.002*	0.000	0.004	0.002**	0.000
	(0.79)	(1.90)	(0.21)	(0.82)	(2.06)	(0.25)
Bank Variables						
Ln(Bank assets)	-0.013**	-0.015***	-0.008***	-0.014**	-0.016***	-0.008***
	(-2.02)	(-9.16)	(-4.98)	(-2.09)	(-9.77)	(-4.77)
Bank equity ratio	-10.405***	-1.313***	-1.917***	-10.598***	-1.589***	-1.703***
	(-24.45)	(-11.66)	(-18.42)	(-25.06)	(-13.83)	(-16.46)
Bank NPL ratio	0.066	0.895***	-0.086	0.090	0.853***	0.041
	(0.16)	(10.28)	(-0.78)	(0.22)	(9.70)	(0.37)
Bank liquid asset ratio	-0.853***	0.063**	-0.075***	-0.908***	0.003	-0.034
	(-7.35)	(1.98)	(-2.68)	(-7.87)	(0.10)	(-1.22)
Bank ROA	-8.035***	1.571***	-0.128	-7.991***	1.002*	0.542
	(-3.61)	(3.07)	(-0.21)	(-3.58)	(1.95)	(0.90)
Firm Variables						
Ln(Firm assets)	0.048***	0.007***	0.007***	0.048***	0.006***	0.008***
	(9.72)	(3.68)	(5.30)	(9.64)	(3.28)	(5.48)
Firm ROA	0.610***	0.089***	0.091***	0.607***	0.091***	0.087***
	(22.33)	(15.55)	(13.87)	(22.22)	(15.92)	(13.39)
Firm leverage	0.140***	-0.134***	-0.151***	0.142***	-0.137***	-0.147***
	(3.49)	(-18.11)	(-8.86)	(3.55)	(-18.53)	(-8.63)
Green	-0.113***	-0.012	-0.011	-0.112***	-0.011	-0.011
	(-2.75)	(-1.31)	(-1.09)	(-2.74)	(-1.19)	(-1.11)
Brown	0.283***	0.046***	0.039*	0.282***	0.045**	0.041*
	(4.19)	(2.58)	(1.68)	(4.18)	(2.52)	(1.73)
Observations	2,715,622	2,715,622	2,715,622	2,715,622	2,715,622	2,715,622
Adjusted R-squared	0.010	0.016	0.007	0.010	0.013	0.006
Impact of an increase by four years (or one class) on the dependent variable as a percent of its standard deviation						
	1%	6%	4%	7%	1%	8%
				10%	8%	9%
				8%	18%	0%

Table 4. Bank-Firm Relationship Length, Distance Bank HQ to Firm and Banks' Private Information

The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to non-listed firms. The number of observations equals 2,994,729. The dependent variables are: the Depth of Private Information, which is the natural log of the squared residuals; and, the Better and Worse Private Information which is equal to the absolute value of the residual, and 0 otherwise, if the residual is larger or smaller, respectively, than zero. In all cases "the residual" is the estimated residual from the first stage equation in Model (1) in Table 1. The definition for each independent variable is given in Table A.1. Coefficients are listed in the first row, t-statistics based on robust standard errors are reported in the row below in parentheses, and the corresponding significance levels are adjacent to the coefficient. Standard errors are clustered at the firm level. *** Significant at 1%, ** significant at 5%, * significant at 10%.

	(1)	(2)	(3)
Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables			
Distance bank HQ to firm	0.005	-0.005***	0.006***
	(0.80)	(-3.27)	(3.55)
Length bank-firm relationship	4.528	-0.659	0.325
	(1.10)	(-0.59)	(0.26)
Distance bank HQ to firm * Length bank-firm relationship	-0.134	1.194***	-0.936***
	(-0.21)	(6.22)	(-4.72)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.016	0.007



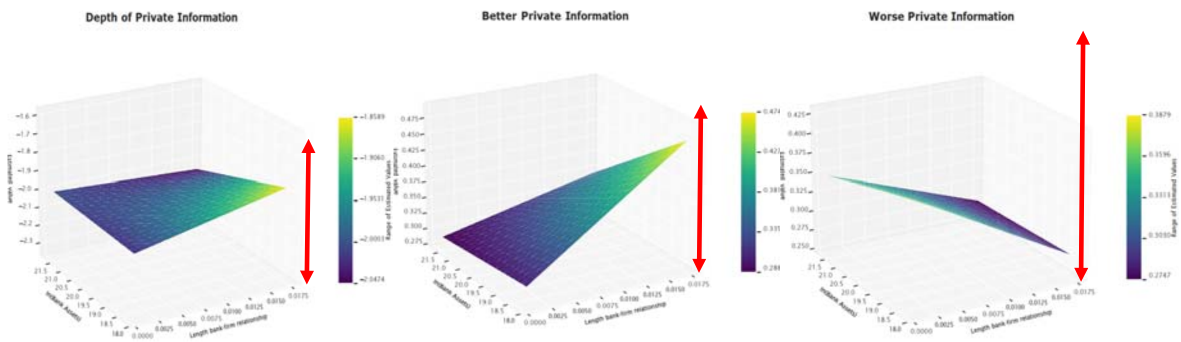
The figures plot the Depth of Private Information and Better and Worse Private Information for the length of the bank-firm relationship and its interactions with distance bank HQ to firm. All other variables are set at their median. The red arrows indicate fifty percent of the standard deviation of the respective outcome variable.

	(1)	(2)	(3)
Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables			
COVID-19	-0.034*	-0.042***	0.050***
	(-1.69)	(-12.16)	(10.26)
Length bank-firm relationship	3.866***	6.750***	-5.417***
	(3.45)	(19.77)	(-20.50)
COVID-19 * Length bank-firm relationship	-4.632**	0.749*	-2.730***
	(-2.41)	(1.87)	(-5.98)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.016	0.007

Table 5. Bank-Firm Relationship Length, Bank Size and Equity and Banks' Private Information

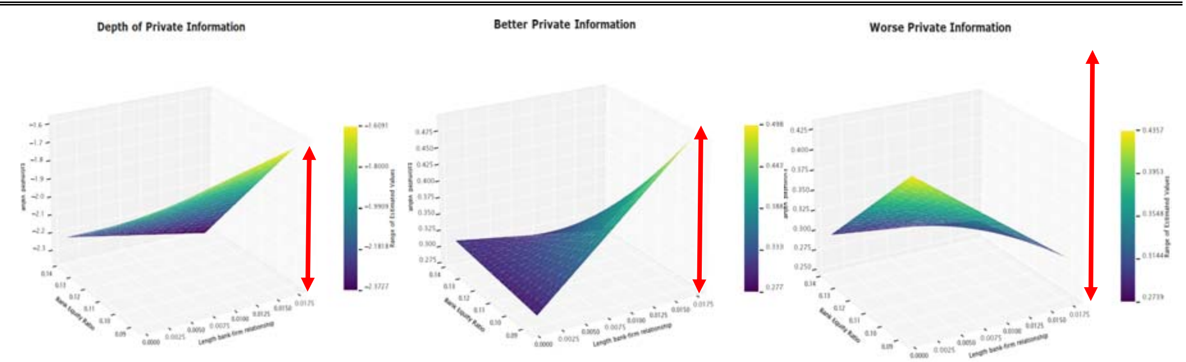
The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to non-listed firms. The number of observations equals 2,994,729. The dependent variables are: the Depth of Private Information, which is the natural log of the squared residuals; and, the Better and Worse Private Information which is equal to the absolute value of the residual, and 0 otherwise, if the residual is larger or smaller, respectively, than zero. In all cases "the residual" is the estimated residual from the first stage equation in Model (1) in Table 1. The definition for each independent variable is given in Table A.1. Coefficients are listed in the first row, t-statistics based on robust standard errors are reported in the row below in parentheses, and the corresponding significance levels are adjacent to the coefficient. Standard errors are clustered at the firm level. *** Significant at 1%, ** significant at 5%, * significant at 10%.

Dependent Variable	(1) *Depth of Private Information*	(2) *Better Private Information*	(3) *Worse Private Information*
Independent variables			
Ln(Bank assets)	0.006	-0.004*	-0.012***
	(0.86)	(-1.92)	(-5.73)
Length bank-firm relationship	71.941***	44.588***	-18.009***
	(4.17)	(8.90)	(-4.20)
Ln(Bank assets) * Length bank-firm relationship	-3.434***	-1.903***	0.630***
	(-3.98)	(-7.59)	(2.87)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.017	0.007



The figures plot the Depth of Private Information and Better and Worse Private Information for the length of the bank-firm relationship and its interactions with (the natural logarithm of) bank assets. All other variables are set at their median. The red arrows indicate fifty percent of the standard deviation of the respective outcome variable.

Dependent Variable	(1) *Depth of Private Information*	(2) *Better Private Information*	(3) *Worse Private Information*
Independent variables			
Bank equity ratio	-7.731***	0.490***	-2.774***
	(-14.17)	(3.25)	(-21.08)
Length bank-firm relationship	47.480***	36.291***	-19.521***
	(6.86)	(16.20)	(-14.41)
Bank equity ratio * Length bank-firm relationship	-408.758***	-275.684***	130.964***
	(-6.47)	(-14.78)	(11.00)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.020	0.008

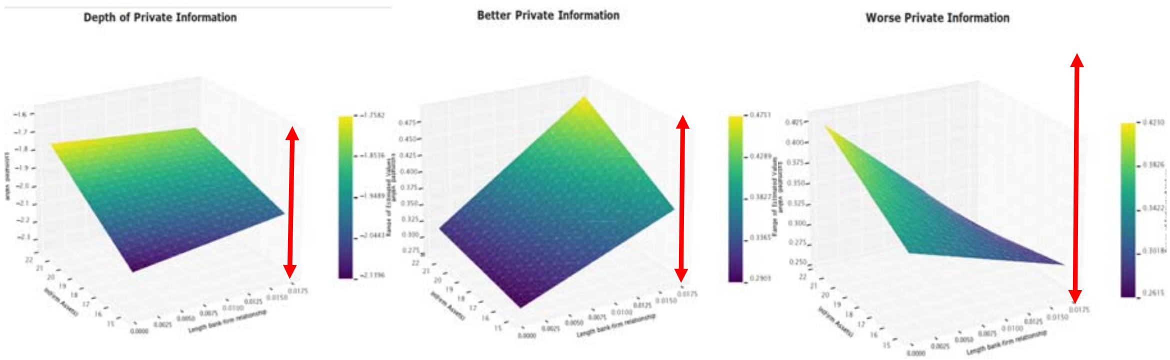


The figures plot the Depth of Private Information and Better and Worse Private Information for the length of the bank-firm relationship and its interactions with the bank equity ratio. All other variables are set at their median. The red arrows indicate fifty percent of the standard deviation of the respective outcome variable.

Table 6. Bank-Firm Relationship Length, Firm Size and Leverage and Banks' Private Information

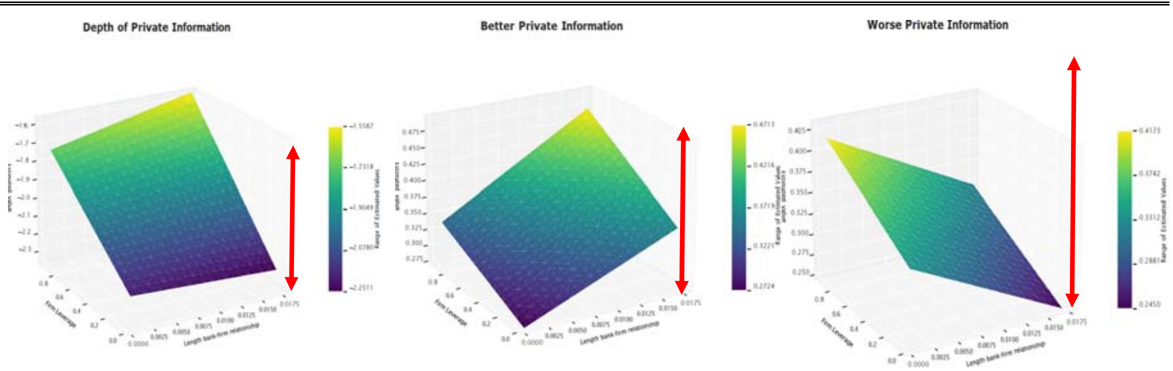
The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to non-listed firms. The number of observations equals 2,994,729. The dependent variables are: the Depth of Private Information, which is the natural log of the squared residuals; and, the Better and Worse Private Information which is equal to the absolute value of the residual, and 0 otherwise, if the residual is larger or smaller, respectively, than zero. In all cases "the residual" is the estimated residual from the first stage equation in Model (1) in Table 1. The definition for each independent variable is given in Table A.1. Coefficients are listed in the first row, t-statistics based on robust standard errors are reported in the row below in parentheses, and the corresponding significance levels are adjacent to the coefficient. Standard errors are clustered at the firm level. *** Significant at 1%, ** significant at 5%, * significant at 10%.

Dependent Variable	(1) *Depth of Private Information*	(2) *Better Private Information*	(3) *Worse Private Information*
Independent variables			
Ln(Firm assets)	0.058***	0.003	0.012***
	(10.27)	(1.59)	(7.34)
Length bank-firm relationship	32.224***	-3.543	10.157***
	(2.63)	(-0.75)	(2.82)
Ln(Firm assets) * Length bank-firm relationship	-1.671**	0.603**	-0.916***
	(-2.24)	(2.12)	(-4.21)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.016	0.007



The figures plot the Depth of Private Information and Better and Worse Private Information for the length of the bank-firm relationship and its interactions with (the natural logarithm of) firm assets. All other variables are set at their median. The red arrows indicate fifty percent of the standard deviation of the respective outcome variable.

Dependent Variable	(1) *Depth of Private Information*	(2) *Better Private Information*	(3) *Worse Private Information*
Independent variables			
Firm leverage	0.521***	0.074***	0.092***
	(15.93)	(9.49)	(11.24)
Length bank-firm relationship	-3.600*	5.520***	-5.412***
	(-1.84)	(9.72)	(-10.64)
Firm leverage * Length bank-firm relationship	15.044***	2.557**	-0.164
	(4.41)	(2.38)	(-0.19)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.016	0.007



The figures plot the Depth of Private Information and Better and Worse Private Information for the length of the bank-firm relationship and its interactions with firm leverage. All other variables are set at their median. The red arrows indicate fifty percent of the standard deviation of the respective outcome variable.

Appendix

Appendix A1. Methodology

Given a cross-section of N observations (i.e., credit ratings of loan contracts) indexed by $i=1,...,N$, the regression model with multiplicative heteroskedasticity formalizes as the two following equations:[1]

$$y_i = \boldsymbol{x}_i'\beta + u_i, \tag{1}$$

and

$$\sigma_i^2 = \sigma^2 e^{\boldsymbol{z}_i'\gamma}. \tag{2}$$

Equation (1) is the first-stage "mean equation", while (2) is the second-stage "variance equation". The identifying assumptions are:

$$E[u_i|\boldsymbol{x}_i] = 0, \tag{3}$$

and

$$Var[u_i|\boldsymbol{z}_i] = \sigma_i^2 = \sigma^2 e^{\boldsymbol{z}_i'\gamma}. \tag{4}$$

y_i is the dependent variable, i.e., the internal bank credit rating, $\boldsymbol{x}_i$ is a vector of explanatory variables in the mean equation that includes a constant, and u_i is a disturbance term.

The variance of the error term is an exponential function of a vector of individual-specific attributes denoted by $\mathbf{z}_i$. Although other functional forms of heteroskedasticity can be used, the exponential form is particularly convenient because it ensures positive variance.

The interpretation of γ is crucial for our intended analysis here. Pick one variable from the vector $\mathbf{z}$, say, z^k, and the respective parameter, γ^k. A positive γ^k indicates that the precision of

[1] The heteroskedastic version extends the linear regression model by also parametrizing the unexplained variance as a function of exogenous covariates. For a discussion, see for example Cerqueiro, Degryse, and Ongena (2013), and for other applications, see also Gaul and Stebunovs (2009), Cerqueiro, Degryse, and Ongena (2011), Iannotta (2011), Iannotta and Navone (2012), and/or Baele, De Bruyckere, De Jonghe, and Vander Vennet (2014).

the credit rating model decreases in z^k. One can interpret such a result as evidence of a positive correlation between the variable z^k and the weight of the difficulties in arriving at a precise rating in the rating-setting process. When $\gamma^k = 0$, the error term is homoscedastic and its variance equals σ^2.

In this setting, our interest lies only in the first two moments of the conditional distribution of y. It is therefore plausible to assume that the error term follows a normal distribution. Under this assumption, the conditional distribution of y is given by:

$$y_i|\boldsymbol{x}_i, \boldsymbol{z}_i \xrightarrow{d} N\left(\boldsymbol{x}_i'\beta, \sigma^2 e^{\boldsymbol{z}_i'\gamma}\right). \tag{5}$$

The simplest procedure to estimate this heteroskedastic regression model is to estimate the parameters in the first stage mean equation by OLS and to use the squared errors as raw estimates of the individual variances. Then, one obtains estimates of the parameters in the second stage variance equation by regressing the (logarithm of) squared errors on the set of covariates in the vector $\boldsymbol{z}$.

This procedure is computationally simpler than obtaining the maximum-likelihood estimates in the heteroskedastic regression model. But there is a loss of efficiency in this two-step procedure (Harvey (1976)), which in our application with many observations is a price we are willing to incur.

References to the Appendix

Baele, L., V. De Bruyckere, O. De Jonghe, and R. Vander Vennet. 2014. Do Stock Markets Discipline US Bank Holding Companies: Just Monitoring, or also Influencing? *North American Journal of Economics and Finance* 29: 124-45.

Cerqueiro, G., H. Degryse, and S. Ongena. 2011. Rules versus Discretion in Loan Rate Setting. *Journal of Financial Intermediation* 20: 503-29.

———. 2013. Using Heteroskedastic Models to Analyze the Use of Rules versus Discretion in Lending Decisions. In *Handbook of Research Methods and Applications in Empirical Finance.* Eds. A. Bell, C. Brooks and M. Prokopczuk. London: Edward Elgar, 216-37.

Gaul, L., and V. Stebunovs. 2009. Ownership and Asymmetric Information Problems in the Corporate Loan Market: Evidence from a Heteroskedastic Regression, Office of the Comptroller of the Currency, Washington DC, Economics Working Paper, 1.

Harvey, A. C. 1976. Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica* 44: 461-65.

Iannotta, G. 2011. Market Discipline in the Banking Industry: Evidence from Spread Dispersion. *European Journal of Finance* 17: 111-31.

Iannotta, G., and M. Navone. 2012. The Cross-section of Mutual Fund Fee Dispersion. *Journal of Banking and Finance* 36: 846-56.

Appendix Table A1. Variable Definitions and Sources

The table provides variable descriptions and their sources. The data are merged using their most recent available values. All continuous variables are winsorized at the 1 pct and 99 pct level. The main sample includes corporate loans to non-listed firms reported in the Y-14Q by bank holding companies between Sep 30, 2012, and March 31, 2021.

Variables		Unit / Split	Description	Source
Equation	**DEPENDENT VARIABLES**			
Bank Rating	Internal bank rating of firm	1 (best) -10 (worst)	The rating given by the bank to the firm transferred to a common scale	FR Y-14Q Schedule H.1
	Residual	-	Estimated residual from the first stage equation	Own calculations
Residual	*Depth of Private Information*	-	Natural log of the squared residuals from the first stage equation estimated	Own calculations
	Better Private Information	-	Negative residual from the first stage equation	Own calculations
	Worse Private Information	-	Positive residual from the first stage equation	Own calculations
Loan Outcome	Loan Interest Rate Spread	%	Interest rate spread over the rate of a constant maturity Treasury bond with a similar maturity.	FR Y-14Q Schedule H.1
	Ln(Loan Maturity)	ln years	The log of one plus the number of years from the date of origination to the date of maturity.	FR Y-14Q Schedule H.1
	Ln(Loan Amount)	ln mln $	The log of one plus the size of the loan in $ million.	FR Y-14Q Schedule H.1
	d(Collateralized)	0/1	= 1 if the loan is collateralized, = 0 otherwise.	FR Y-14Q Schedule H.1
Level of Variables	**INDEPENDENT VARIABLES**			
Bank-Firm				
	Distance bank HQ to firm	ln miles	The log of the distance between the bank`s headquarters and the firm`s location	FR Y-14Q Schedule H.1
	Length bank-firm relationship	0.001 years	The number of years since the borrower had the first loan with the bank / 1000	FR Y-14Q Schedule H.1
	Length bank-firm relationship (0.25 << 3 years)	0/1	= 1 if the number of years since the borrower had the first loan with the bank is between 0.25 and 3 years, = 0 otherwise	FR Y-14Q Schedule H.1
	Length bank-firm relationship (3 << 5 years)	0/1	= 1 if the number of years since the borrower had the first loan with the bank is between 3 and 5 years, = 0 otherwise	FR Y-14Q Schedule H.1
	Length bank-firm relationship (> 5 years)	0/1	= 1 if the number of years since the borrower had the first loan with the bank is longer than 5 years, = 0 otherwise	FR Y-14Q Schedule H.1
Bank				
	Bank assets	mln $	Bank total assets in million US$	FR Y9-C
	Ln(Bank assets)	ln mln $	Log of one plus bank total assets	FR Y-9C
	Bank equity ratio	-	Equity ratio, calculated as total equity / total assets	FR Y-9C
	Bank NPL ratio	-	Non-performing loan ratio, calculated as: loans at least 90 days past due or in nonaccrual status / total assets	FR Y-9C
	Bank liquid asset ratio	-	Liquid asset ratio, calculated as: cash + marketable securities / total assets	FR Y-9C
	Bank ROA	-	Return on assets, calculated as: net income / total assets	FR Y-9C
Firm				
	Firm assets	mln $	Firm current assets in million US$	FR Y-14Q Schedule H.1
	Ln(Firm assets)	ln mln $	Natural log of one plus the total amount of firm's current assets	FR Y-14Q Schedule H.1
	Firm ROA	-	Return on assets of the firm, calculated as net income / total assets	FR Y-14Q Schedule H.1
	Firm leverage	-	Leverage ratio of the firm	FR Y-14Q Schedule H.1
	Green	0/1	= 1 if the firm is in a green industry, = 0 otherwise	BLS
	Brown	0/1	= 1 if the firm is in a brown industry, = 0 otherwise	BLS
Loan				
	d(Loan is not a syndicate)	0/1	= 1 if the loan is not a syndicated loan, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan is a term loan)	0/1	= 1 if the loan is a term loan, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan is a revolver)	0/1	= 1 if the loan is a revolver, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan is floating rate)	0/1	= 1 if the loan is a floating-rate loan, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan is mixed rate)	0/1	= 1 if the loan is a mixed-rate loan, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan purpose is miscellaneous)	0/1	= 1 if loan purpose is related to activities other than M&A or capital expenditures, general purpose, or commercial real estate, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan purpose is M&A or capital expenditure)	0/1	= 1 if loan purpose is related to M&A or capital expenditures, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan purpose is general)	0/1	= 1 if loan purpose is general purpose, = 0 otherwise	FR Y-14Q Schedule H.1
	d(Loan purpose is real estate)	0/1	= 1 if loan purpose is related to commercial real estate, = 0 otherwise	FR Y-14Q Schedule H.1

Appendix Table A2. Variable Summary Statistics

The table provides variable summary statistics. The data are merged using their most recent available values. All continuous variables are winsorized at the 1 pct and 99 pct level. The main sample includes corporate loans to non-listed firms reported in the Y-14Q by bank holding companies between Sep 30, 2012, and March 31, 2021. The number of observations equals 2,994,729.

Variables		Unit / Split	Mean	Median	Minimum	Maximum	Standard Deviation
Equation	**DEPENDENT VARIABLES**						
Bank Rating	Internal bank rating of firm	1 (best) -10 (worst)	4.984	5	1	10	1.029
	Residual	-	0	-0.059	-5.165	6.732	0.960
Residual	*Depth of Private Information*	-	-2.127	-1.620	-30.398	3.724	2.522
	Better Private Information	-	0.303	0.026	0.000	4.945	0.451
	Worse Private Information	-	0.298	0.000	0.000	6.436	0.577
Loan Outcome	Loan Interest Rate Spread	%	2.201	2.350	-3.69	8.100	1.984
	Ln(Loan Maturity)	ln years	1.449	1.611	-5.900	3.230	0.963
	Ln(Loan Amount)	ln $	15.096	14.809	13.816	23.180	1.122
	d(Collateralized)	0/1	0.918	1	0	1	0.275
Level of Variables	**INDEPENDENT VARIABLES**						
Main Bank-Firm							
	Length bank-firm relationship	years	1.537	0.729	0	7.723	1.887
	Length bank-firm relationship (0.25 << 3 years)	0/1	0.352	0	0	1	0.477
	Length bank-firm relationship (3 << 5 years)	0/1	0.206	0	0	1	0.404
	Length bank-firm relationship (> 5 years)	0/1	0.420	0	0	1	0.494
	Distance bank HQ to firm	ln(miles)	6.230	6.461	0.047	8.124	1.379
	Distance bank HQ to firm in miles	miles	913.575	638.424	0.048	3374.220	827.423
Bank							
	Bank assets	bn $	989.155	381.560	20.454	2808.396	1007.859
	Ln(Bank assets)	ln thd $	19.982	19.760	16.834	21.756	1.293
	Bank equity ratio	-	0.115	0.113	0.052	0.207	0.015
	Bank NPL ratio	-	0.02	0.015	0	0.079	0.014
	Bank liquid asset ratio	-	0.247	0.244	0.083	0.721	0.057
	Bank ROA	-	0.002	0.003	-0.033	0.016	0.002
Firm							
	Firm assets	tn $	0.806	0.024	0.001	98.598	6.114
	Ln(Firm assets)	ln $	17.326	16.997	13.816	25.314	1.932
	Firm ROA	-	0.081	0.052	-0.239	0.840	0.136
	Firm leverage	-	0.420	0.392	0	1	0.270
	Green	0/1	0.033	0	0	1	0.177
	Brown	0/1	0.011	0	0	1	0.103
Loan							
	d(Loan is not a syndicate)	0/1	0.954	1	0	1	0.210
	d(Loan is a term loan)	0/1	0.324	0	0	1	0.468
	d(Loan is a revolver)	0/1	0.469	0	0	1	0.499
	d(Loan is floating rate)	0/1	0.594	1	0	1	0.491
	d(Loan is mixed rate)	0/1	0.170	0	0	1	0.375
	d(Loan purpose is miscellaneous)	0/1	0.233	0	0	1	0.423
	d(Loan purpose is M&A or capital expenditure)	0/1	0.100	0	0	1	0.300
	d(Loan purpose is general)	0/1	0.495	0	0	1	0.500
	d(Loan purpose is real estate)	0/1	0.171	0	0	1	0.376

Swiss Finance Institute

Swiss Finance Institute (SFI) is the national center for fundamental research, doctoral training, knowledge exchange, and continuing education in the fields of banking and finance. SFI's mission is to grow knowledge capital for the Swiss financial marketplace. Created in 2006 as a public–private partnership, SFI is a common initiative of the Swiss finance industry, leading Swiss universities, and the Swiss Confederation.

swiss:finance:institute

c/o University of Geneva, Bd. Du Pont d'Arve 42, CH-1211 Geneva 4
T +41 22 379 84 71, rps@sfi.ch, www.sfi.ch