“If You Don't Know Me by Now ...”

Information Collection by Banks in Lending to Private Firms

Stijn Claessens, Steven Ongena, Teng Wang*

July 23, 2025

What constitutes banks’ information collection on private firms? How does it deepen and change over time? Exploiting a comprehensive Federal Reserve’s supervisory dataset, we extract three dimensions of private information from banks’ internal credit ratings — depth and better/worse assessments — with all three related to various loan terms and performance. We then document how private information evolves as firm-bank relationships lengthen, with learning effects non-linear and peaking at about five years. Learning varies also: it is more salient at longer bank-firm distances, for smaller and more leveraged banks, for larger and riskier firms, and during non-COVID times. (96 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 Federal Reserve Bank of Dallas and the University of Texas at Arlington, USA; email: teng.wang@uta.edu; phone: +1 202 725 1865. For comments, we are grateful to Paul Calem, Sheng Li, Mitch Petersen, Hirofumi Uchida, Greg Weitzner, and seminar participants at the Bank for International Settlements, Federal Reserve Board, Federal Reserve Bank of Kansas City, Hitotsubashi University, University of Texas at Arlington 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), the European Economic Association Annual Meetings, the 2025 Northern Finance Association Conference, and the 7th Conference on the Future of Financial information (INSEAD) for comments. The paper circulated previously under the title “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 Federal Reserve Board of Governors, the Federal Reserve System, or any staff of these institutions.

“If You Don't Know Me by Now ...”

Information Collection by Banks in Lending to Private Firms

July 23, 2025

What constitutes banks’ information collection on private firms? How does it deepen and change over time? Exploiting a comprehensive Federal Reserve’s supervisory dataset, we extract three dimensions of private information from banks’ internal credit ratings — depth and better/worse assessments — with all three related to various loan terms and performance. We then document how private information evolves as firm-bank relationships lengthen, with learning effects non-linear and peaking at about five years. Learning varies also: it is more salient at longer bank-firm distances, for smaller and more leveraged banks, for larger and riskier firms, and during non-COVID times. (96 words)

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 United States — 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 publicly listed companies, which 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 banks' ability to assess and manage credit risk. Classic models by Leland and Pyle (1977) and Fama (1985) notably highlight the significance of information collection in mitigating adverse selection and moral hazard. And models stress how 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 embedded in banks' assessment of private firms, it is not well-understood how this information is acquired, how it varies over time, and how it is shaped by various factors, including the physical distance between banks and firms (e.g., Hauswald and Marquez (2003)), firm and bank characteristics, 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.

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

What is in banks' private information about private (i.e., not publicly listed) borrowers? How is private information affected by physical distance? How does it evolve over time, and what bank and firm characteristics 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 are well over 25 million in the U.S. (in comparison to less than 5,000 publicly listed firms).

To quantify the nature, evolution, and implications of private information embedded in banks' assessment 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 and performance. We then analyze the factors that drive their evolution, as well as similarities and differences across banks and firms. Collectively, this research helps in understanding the intricate process by which banks assess the creditworthiness of private borrowers through the collection of their private information.

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 deeply embedded in the internal credit assessment system of the bank that makes the lending decision and is often obscure (as noted by Liberti and Petersen (2019)).

For our analysis, we collect data on U.S. banks' internal credit assessments of private firms, along with detailed loan contracts and borrower information, as recorded in the Federal

Reserve's supervisory Y-14Q quarterly loan-level dataset.[2] This dataset has a much broader representation than other datasets used for the U.S. (e.g., Beyhaghi, Howes and Weitzner (2024); Faria-e-Castro, Paul and Sánchez (2024); Berger, Bouwman, Norden, Roman, Udell and Wang (2025)) 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 their borrowers, which reflect not only the financial position of the firms, but also the banks' scoring models and techniques and 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 any other financial regulators in the United States. 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 evolves over time and 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.

---

[2] Note that the Y-14Q data differs from a traditional credit bureau in that it none of the data individual banks report to it are shared with other banks. At the same time, it contains information of internal credit risk assessment of each loan assigned by each individual bank's respective loan officers.

[3] When the database 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). All our results are robust, however, to using for the full period only those banks with over $100 billion in assets.

[4] We drop the publicly listed firms as information production for them is very different and largely not done by banks. In earlier versions of our paper, we did include the small share of publicly listed firms, but there are no (meaningful) differences with current results.

To identify the private information embedded in banks' rating, we follow the seminal works by Morgan (2002) and Agarwal and Hauswald (2010).[5] We deploy a two-stage model, which is in a spirit like a heteroskedastic regression model (Harvey (1976)). The core of the methodology involves extracting private information from banks' credit assessments of borrowers. Using this methodology, we aim to capture the portion of a bank's internal credit assessment of one borrower that cannot be explained by the standard set of information, e.g., about the borrower's financial status, sector or location observable to the banker, nor by the characteristics of the bank conducting the assessment. We refer to this as the bank's private information about the borrower.

Importantly, the part of the rating that observable firm characteristics can explain controls for the normal riskiness and profitability of lending to specific firms. 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 only the lending bank that underwrites the loan has access to the information about the borrowing firm.[6] The part that 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 bank size and higher capitalization provide greater scope for diversification benefits and loss absorption, respectively, allowing more risky firms to be rated relatively higher; see, e.g., Plosser and Santos (2018)) is private to the bank.[7]

---

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

[6] 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 information is not shared with or necessarily directly available to all other banks.

[7] The residual could still reflect other differences in how banks combine firm information in their proprietary credit risk modeling to generate the internal ratings. We did test for non-linear and various other effects (and

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 square 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.

We next 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 are made.

Consistent with the existing literature, we confirm that our three novel measures of banks' private information are significant drivers of loan terms, which aligns with prior expectations that internal ratings inherently relate to 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.

We also show that our measures are closely tied to loan performance, as reflected in the actual repayment record in the period ahead. In particular, the likelihood of a loan becoming

report some tests) but no robustness can ever be exhaustive. For this reason, the residual captures both private information and a component due to the bank's idiosyncratic risk modelling, which is also private, at least to the econometrician.

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

90 days overdue four quarters later is higher with greater depth and lower (higher) when private information is positive (negative), all else being equal and controlling for subsequent changes in firm (and bank) characteristics. That our private information proxies predict future loan performance in statistically significant ways with the right signs is another confirmation of their relevance. To further demonstrate robustness, we analyze the terms and performance of loans by distinct banks to the same borrower, which only differ in how the banks privately assess the borrower. We then use the approach of Khwaja and Mian (2008) and control for *all* time-varying dynamics coming from the borrowers' side by using bank and firm-by-year fixed effects (instead of bank and time-varying firm characteristics). Results are similar to our main finding, further confirming the robustness of our proxies.

Next, we investigate how the *depth* and *direction* of the lending banks' private information changes as their relationship with the firm lengthens and how these are 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 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* assessments.

Consistent with our conjectures, we show that the length of the bank-firm relationship contributes to how the depth of private information evolves. As relationships between banks and firms develop, more private information is gathered. This is evidenced by greater divergence between banks' assessments of the firms and assessments solely based on observable, continually updated characteristics. It indicates that banks gain insights about borrowers over the course of their relationships, increasingly better assessing 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, how does bank learning and private information formation vary alongside these dimensions? In line with theoretical literature, yet not empirically documented, we anticipate that the effects related to the length of the relationship will vary by characteristics that indicate different abilities and incentives to acquire private information, as well as levels of information asymmetry. For example, as a relationship lasts longer, we expect the adverse impact of distance on the depth of private information (e.g., Hauswald and Marquez (2003) and Hauswald and Marquez (2006)) not to increase or to possibly even 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 anticipate that the COVID period may have disrupted this learning, since onsite visits and face-to-face meetings between bankers and customers were harder, resulting in less private information being generated. And we expect that more *better* private information is produced over time by those banks that are smaller and less capitalized because their loan officers` and institutional incentives are greater (e.g., Stein (2002)) and for firms that are larger and riskier as interactions over time help more to mitigate the potential initial worries about these borrowers.

To depict the conjectures and understand how bank and firm characteristics affect the banks' private information production process, we provide, besides regression analyses with interaction terms, graphic illustrations of the banks' learning process in 3-D settings. Importantly, these show 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 who are located further away at inception, as the relationship lengthens, this handicap is offset as more private information is collected for distant borrowers. More importantly, the private information collected over time is largely more positive in nature. Similarly, we find that relationship length particularly increases the information depth and makes for *better* assessments for smaller and less capitalized banks and for larger and more leveraged firms. We further perform various additional analyses that consistently support our main findings, confirming their robustness.

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. Banks' internal credit rating

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. Information on loan repayments, for example, is 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 a very large sample of private firms.

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

[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.

[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.

showing that these measures map into the terms and performance 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 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

[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.

[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)).

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] including 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.

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, we obtain estimates of the

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

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 two directional assessment variables. One we call *Better Private Information*, and is the (absolute value of the) estimated residual from the first stage equation when it is less than 0, and equals 0 otherwise. The second is called Worse Private Information and 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 illustration to convey 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 far from Bank 3 (B3), maybe a bit closer to Bank 2 (B2), but even further from Bank 1 (B1). Across all banks and firms, the rating is increasing (i.e., worsening) in distance (an observable) along an estimated line, depicted in dashed gray. This relationship is captured in the first stage. All banks give the firm F5 a different rating (hence there is a dispersion), 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 a rating based on all included observables. These residuals are investigated in the second stage.

In terms of summary statistics,[14] we find 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

[14] 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.

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 also provide direct and detailed measures of the private information banks gather, making them essential for analyzing how banks mitigate information asymmetries in lending.

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).[15] The internal credit risk rating allows then for cross-bank comparison of private

[15] In the context of the Y-14Q data used in the supervisory stress tests, the Federal Reserve also 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 PDs and are within the PD range of publicly 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.

information embedded. Such a straightforward comparison is not possible in other bank or loan data sets such as 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 the internal ratings, we gather additional loan-level information from the Federal Reserve's Y-14Q quarterly report. This report, which has been in place since the fall of 2011, obtains data from reporting banks regarding their commercial and industrial (C&I) loans. The data cover all C&I loans over $1 million in size when originated and held by the largest bank holding companies in the U.S. by assets.[16] Loan sizes range from the 1 million dollar reporting threshold (in commitment) to billions of dollars, thus covering loans to a wide size spectrum of privately held firms (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 collected by the lending bank(s) on firm financials and performance, including total assets, ROA, and leverage as well as the identification of the private borrower, allowing us to calculate the distance between the bank branch (or HQ) and the firm. As noted (footnote 2), no firm information is shared with other banks.

Importantly, 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

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

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 obtain 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", which are evaluated based on the borrower, not the firm, and credit quality, and are rated on a different scale than other corporate loans. But, 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]

In addition, we collect data on bank characteristics from FR Y-9C data. The Y9-C data, which is publicly available, contains quarterly balance sheet and income 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. Comparing the loans in our sample with all C&I loans reported in Y-9C shows that our sample amounted to more than 70 percent in 2022:Q4.[18]

IV. Findings

[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.

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 there to explain — as well as possible — the bank ratings based on information that is observable to loan officers in principle, including financial information of the borrower firms and the banks that make the loans.[19] We do this for a combined sample of the initial credit assessments, 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 any subsequent assessments, that is, as long as the loan is on the balance sheet of a bank in our sample. We include bank-firm distance and a comprehensive set of bank and firm controls to be as exhaustive as possible (yet disregarding 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). Since the fitted estimate reflects all quantitative information that can, in principle, be collected by any bank, a bank's private assessment of a firm, which reflects the specific methodology it uses to derive ratings, is captured in the residual. We deem this banks' private information, but since, as noted, firm financial information is not shared among the banks, this favors a conservative measure of private information.

[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 (2025) for further discussions.

The estimated coefficients are in Table 1 Model (1). The sign and size of the 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, more leveraged, poor loan quality, 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 intuitive and largely similar to those already established in the literature. For example, weaker banks are shown to give better ratings to firms 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 reported to the Fed data on the specific firm it lends to. However, when we exclude all financial information and just include firm sector and green/brown type dummies for example, our main findings are mostly unaffected by this modulation. To further assess the robustness of our baseline results, we extend our specification in the first stage by incorporating a richer set of controls. In particular, we augment the model with the squared terms of all bank- and firm-level characteristics included in the baseline specification in addition to their raw values. The inclusion of these higher-order terms allows for more flexible functional forms, mitigating concerns that our findings might be driven by non-linearities or omitted interactions between

key covariates. Importantly, our main results remain quantitatively and qualitatively unchanged under this expanded specification.[20]

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 measures to capture banks' private information. We next relate these three measures at the beginning of the lending relationship to initially known factors, thus abstracting from any influences due to learning. We use several factors including the geographic distance between bank and firm and their characteristics, also reported in Table 1.

In Model (2), the dependent variable is *Depth of Private Information*, while in Models (3) and (4) the dependent variables correspond to the "directional equations," namely *Better Private Information* and *Worse Private Information*, respectively. Given that the residuals in these directional models are structured such that larger values of both *Better* and *Worse* represent stronger directional assessments, a positive estimated coefficient can be interpreted as the associated factor adding to the bank's private information. Specifically, in Model (3), a positive coefficient indicates that the factor improves (or "betters") the private signal relative to what is inferred from observable variables, whereas in Model (4), a positive coefficient reflects a worsening of the private signal relative to the benchmark.

[20] Altering the set of variables included in the first stage equation leaves the second stage equation estimates mostly unaffected. This fact should not come as a surprise. As Cerqueiro, Degryse and Ongena (2013) points out, 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 also pertains to the bank and firm variables which, in any case, are rather standard when explaining credit ratings (e.g., Altman (1968)).

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***,[21] 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 (pp) (= 0.007 * 6.461), or about 10 percent of its standard deviation (= 0.045 / 0.451). Hence, physical distance reduces 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, and thus reduced incentives, 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, higher loan quality, 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

[21] As noted in the Tables, we indicate statistical significance as follows: *** p<0.01, ** p<0.05, and * p<0.1.

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 this private information leads banks to rate better when dealing with “good” larger firms (recall that a lower rating is “better”), and to rate worse when dealing with “bad” larger firms.

These results suggest that the banks studied here (which, recall, are the larger ones in the U.S.) when they amass private information on firms and reflect it in their ratings, including their direction, initially assign more “cookie cutter” (and hence more explainable based on observables) ratings for smaller firms (as in, e.g., Cole, Goldberg and White (2004)). However, as we will see below, relatively there is more learning over time for the larger and more leveraged firms, for which it may be more valuable to do.

Another way to assess private information is to focus on the 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.[22] 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 for the firm (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

[22] 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).

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

D. Private information and loan terms

To confirm that our measures reflect banks' credit assessments, we first investigate how banks' loan terms vary with our three measures of bank private information. In Table 2, 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. Note that we use here again all observations after the loan was initiated, to allow for the loan terms to be adjusted as bank and firm characteristics change and the relationship evolves. All the regressions in Panel A 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. In Panel B we include instead of the bank and firm variables of Panel A, bank and firm * year fixed effects, which thus account for all constant bank and time-varying firm characteristics. The overall sample sizes are somewhat smaller for the spread regressions (as Y-14Q does not report the interest rate on undrawn credit) but otherwise they are very similar to the full

sample, so results do not reflect the sample choices.[23] Note also that the use of multi-bank firms only makes for a slightly smaller sample of firms.

The estimates reported in Panel A show that when lending, greater depth comes with a higher interest rate spread. This finding likely 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. Economically, the relations in Panel A are the strongest 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 relations for collateralization, which in Panel A 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 relations for the other dependent variables, loan maturity and amount, are economically much smaller, 2 percent or less equivalently expressed.

Next in Table 2 Panel B, we study only the multi-bank firms, where we effectively compare at a given point in time the implications of private information sets contained in loans to the same borrower by distinct banks, who still differ in their relations. The estimates in Panel B still show consistently that our private information measures relate to lending terms. But the

[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.

statistical and economic significance of the results reported in Panel B are much less than of those in Panel A. Specifically, the estimates now suggest in economic terms that our private information measures explain only around a quarter or so of the effects explained in Panel A. Recalling our deliberations — above — on the excessiveness of saturating with firm-time fixed effects, this is no surprise. When using firm * year fixed effects, we control for any change known and unknown to an econometrician that these multi-bank firms may have experienced in any quarter. This obviously reduces the information a bank privately is assumed to have relative to the other banks, which in turn makes for a lower relation to its lending terms. Directly, the effects are nevertheless the same in both Panels. These regression results are also robust to using our alternative information variables (which adds squared terms of all explanatory variables to the first stage), as reported in Appendix Tables A4 and A5. Overall, these associations confirm that our information measures are a meaningful part of the banks' lending procedures.

E. Private information and loan performance

We next examine the predictive power of our information measures for subsequent loan performance by estimating regressions on the same sample. Specifically, we relate the three measures of private information to indicators of loan delinquency within the next four quarters, defined as whether the loan becomes overdue by more than 90 days. The 90-day threshold is widely recognized in both banking practice and regulatory frameworks—such as those outlined by the Basel Committee and U.S. bank supervisory agencies—as the standard benchmark for classifying loans as non-performing, with loans past due by more than 90 days typically considered to have a significantly increased risk of default and often written down

or placed into collections.[24] Table 3 presents the regression results, where Panel A employs the usual bank and firm control variables, and Panel B uses bank and firm * year fixed effects to account for all constant (bank) and time-varying (firm) characteristics.

The two sets of regression results are consistent in showing that the information measures relate to loan performance, with the depth of and worse private information positively, and better private information negatively associated with a more likely overdue status of more than 90 days, four quarters in the future. These regression results are also robust to using our alternative information variables (which add squared terms of all explanatory variables to the first stage), as reported in Appendix Tables A4 and A5. This is another confirmation that our information measures are meaningful reflections of how banks privately assess firms.

F. 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 4, 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.

[24] We also used an alternative threshold of 30-day past due as a robustness check in Appendix Table A3.

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 0.6 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.[25] For *depth*, the impact of relationship seems to peak between 3 and 5 years. For *better* information, there seems to be an increasing value of relationship throughout as the estimated coefficients continue to increase. For *worse* information, it appears that the impact peaks between 3 and 5 years, after which ratings marginally improve (i.e., get less worse).[26] Overall, these regression results suggest that banks proactively use their relationship to improve their assessments of the firms they lend to.[27]

---

[25] 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.

[26] 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).

[27] 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.

The estimates on distance (bank HQ to firm) in Table 4, across Models (1)–(3) and Models (4)–(6), are now 0.004, 0.002*, and 0.000, respectively. These differ notably from the corresponding estimates in Table 1 (−0.012, −0.007***, and 0.001, respectively). This shift suggests that, beyond the initial lending stage, the development of a lending relationship significantly alters how distance influences *depth* and the *better* and *worse information*. In particular, the adverse effects of distance observed in the initial assessment are mitigated — or in some cases, even reversed — once a stronger bank–firm relationship is established. In other words, as relationships mature, banks appear more willing to deviate from purely "hard" information when assessing distant borrowers, applying more nuanced or relationship-driven evaluations. Moreover, these regression results are robust to alternative specifications, including the use of expanded information measures that incorporate squared terms of bank and firm characteristics (reported in Appendix Table A6), confirming that the observed attenuation of the distance effect is not driven by modeling choices.

G. Private information formation over time and 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, affect the formation of banks' private information over time through the various kinds of engagement between the bank and the firm.

In Tables 5 to 7 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 one variable each among the distance between the bank

and the firm, a dummy for the COVID period (2020:Q1 - 2021:Q1),[28] bank size and capital, and firm size and leverage. We include every time the variable itself, the relationship length, and the interaction of the specific variable with relationship length. The regressions do include the usual bank-firm, bank, and firm controls but these are not reported. The sample size is kept the same for all regressions, i.e., 2,713,281 observations, so that any variations in results do not reflect sample choices.

For each of the results, only three estimated coefficients are reported, that is, the coefficients for the specific variable, the relationship length, and their interaction. Of most interest are the estimated coefficients for the interactions. These terms reflect how various distance, bank, and firm characteristics and time periods differently affect the formation of private information and thus shed 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 relevance of the impact, the red arrows display 50 percent of the standard deviation of the respective outcome variable.

Several key findings emerge from these analyses. Focusing on distance effects in Table 5, we observe that the length of the lending relationship does not significantly alter the overall *depth* of private information collected for distant borrowers. However, the nature of private information becomes more favorable over time as is reflected by higher amount of *better information* (positive coefficient in Model (2)) and lower amount of *worse information* collected (negative coefficient in Model (3)), as the relationship lengthens. This pattern

[28] To avoid any impact of the COVID crisis, we also employed a sample that excludes observations from the year 2020 on and find that all our earlier results do not change materially.

suggests that a longer relationship substantially mitigates the adverse effects of distance by enabling banks to form sharper, more differentiated assessments of borrower quality.

The graphic representations of the regression analyses further illustrate how the process behind banks' private information formation varies across borrowers located close to versus farther away. The first chart shows that the overall *depth* of banks' private information does not change much over the time of the relationship, as reflected in the largely horizontal plane, with only a limited change from dark blue to yellow as the relationship lengthens. Turning to the direction of private information, however, we observe importantly that the learning curves for both more favorable and unfavorable ratings are steeper for distant borrowers who are located further away. Although initially being assigned a less favorable and more unfavorable rating, banks seem to accumulate *better* information about those borrowers located further away over time. The much higher *better* private information and somewhat less *worse* information content suggest that firms located further away do gain in banks' trust as the relationship lengthens. This result is consistent with the classic banking literature, which finds that geographic distance can initially create significant hurdles for banks to tackle information asymmetry problems (e.g., Dell'Ariccia and Marquez (2004)). 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 during the COVID-19 crisis may have hindered in-person interactions between bank loan officers and borrowing firms, which are essential for collecting private

information. Supportive of this hypothesis, as shown in Table 5 in the lower panel, the length of the relationship is meaningfully less important during COVID times 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 to that for the length of the relationship, there appears to have been on net no change as to *depth* during the COVID period. Directionally, it appears that for firms with *better* assessments, ratings improved only marginally, but those with *worse* assessments got significantly 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, even though they had less, or no, interactions (potentially, also in light of the more ample support, forbearing).[29]

In term of bank characteristics, the results in Table 6 and the graphic analyses 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. This suggests that 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 banks' lending more to information-intensive firms (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

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

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, not reported, are similar to those for well capitalized banks).

Finally, in terms of firm characteristics, firm size and riskiness are typically thought to relate to private information acquisition and learning. The size result in Table 7 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. Using leverage to proxy firms' risks, we find that over time banks increase the *depth* of their assessment and raise their *better* private view of those riskier firms, i.e., those that are more leveraged. Overall, these results suggest that banks are more incentivized to learn about firms and adjust their private assessment more when the firm is larger and more leveraged (and thus riskier).

Overall, the figures derived from Tables 5 to 7 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 firms that are further away, by smaller and less leveraged banks, and for large and highly leveraged firms. In terms of *worse* private information, longer relationships improve scores to a significant degree for firms that are further away, to some extent for smaller and less leveraged banks, and meaningfully for large and more levered firms. These associations suggest that the various types of banks make meaningfully different choices as to how to enhance their private information and vary in their learning process by specific types of borrowers.

Together, these findings suggest that the geographic distance between the banks and firms, a bank's business model, as well as firm size and riskiness affect how strong learning from

relationship affects banks' private assessments of borrowers. A possible common thread is that the acquisition by smaller banks of private information on firms is more relationship-based, i.e., making the setting more conducive to learning, whereas for larger firms, more is at stake, making learning more pressing.

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 and performance 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 larger and more 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., Hauswald, R., 2010. Distance and Private Information in Lending. Review of Financial Studies. 23, 2757-2788.

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

Bannier, C.E., Behr, P., Güttler, A., 2010. Rating Opaque Borrowers: Why Are Unsolicited Ratings Lower? Review of Finance. 14, 263-294.

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

Berg, T., Reisinger, M., Streitz, D., 2021. Spillover Effects in Empirical Corporate Finance. Journal of Financial Economics. 142, 1109-1127.

Berger, A.N., Bouwman, C.H.S., Norden, L., Roman, R.A., Udell, G.F., Wang, T., 2025. Is a Friend in Need a Friend Indeed? How Relationship Borrowers Fare during the COVID-19 Crisis. Journal of Financial Intermediation. 63, 101150.

Berger, A.N., Klapper, L.F., Udell, G.F., 2001. The Ability of Banks to Lend to Informationally Opaque Small Businesses. Journal of Banking and Finance. 25, 2127-2167.

Berger, A.N., Udell, G.F., 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, 850-855.

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

Blickle, K., He, Z., Huang, J., Parlatore, C., 2025. Information-based Pricing in Specialized Lending. Journal of Financial Economics. Forthcoming.

Bonfim, D., Nogueira, G., Ongena, S., 2021. “Sorry, We're Closed" Bank Branch Closures, Loan Pricing and Information Asymmetries. Review of Finance. 25, 1211-1259.

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

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

Carling, K., Jacobson, T., Lindé, J., Roszbach, K., 2007. Corporate Credit Risk Modeling and the Macroeconomy. Journal of Banking and Finance. 31, 845-868.

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

Cavallo, E., Powell, A., Rigobon, R., 2013. Do Credit Rating Agencies Add Value? Evidence from the Sovereign Rating Business. International Journal of Finance and Economics. 18, 240-265.

Cerqueiro, G., Degryse, H., Ongena, S., 2011. Rules versus Discretion in Loan Rate Setting. Journal of Financial Intermediation. 20, 503-529.

Cerqueiro, G., Degryse, H., Ongena, S., Using Heteroskedastic Models to Analyze the Use of Rules versus Discretion in Lending Decisions, in: A. Bell, C. Brooks, M. Prokopczuk, (Eds.), Handbook of Research Methods and Applications in Empirical Finance. Edward Elgar, London, 2013, pp. 216-237.

Cole, R.A., Goldberg, L.G., White, L.J., 2004. Cookie-Cutter versus Character: The Micro Structure of Small Business Lending by Large and Small Banks. Journal of Financial and Quantitative Analysis. 39, 227-252.

Correa, R., Lee, K.-H., Sapriza, H., Suarez, G.A., 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., Dewachter, H., Mulier, K., Ongena, S., Schepens, G., 2020. Some Borrowers Are More Equal than Others: Bank Funding Shocks and Credit Reallocation. Review of Finance. 24, 1-43.

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

Degryse, H., Ioannidou, V., Ongena, S., Bank-Firm Relationships: A Review of the Implications for Firms and Banks in Normal and Crisis Times, in: W. Tsutomu, I. Uesugi, A. Ono, (Eds.), The Economics of Interfirm Networks. Springer, Berlin, 2015, pp. 177-189.

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

Degryse, H., Laeven, L., Ongena, S., 2009. The Impact of Organizational Structure and Lending Technology on Banking Competition. Review of Finance. 13, 225-259.

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

Degryse, H., Ongena, S., 2005. Distance, Lending Relationships, and Competition. Journal of Finance. 60, 231-266.

Degryse, H., Ongena, S., 2007. The Impact of Competition on Bank Orientation. Journal of Financial Intermediation. 16, 399-424.

Degryse, H., Ongena, S., Competition and Regulation in the Banking Sector: A Review of the Empirical Evidence on the Sources of Bank Rents, in: A.V. Thakor, A.W.A. Boot, (Eds.), Handbook of Financial Intermediation and Banking. Elsevier, Amsterdam, 2008, pp. 483-554.

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

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

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

Duqi, A., Tomaselli, A., Torluccio, G., 2018. Is Relationship Lending Still a Mixed Blessing? A Review of Advantages and Disadvantages for Lenders and Borrowers. Journal of Economic Surveys. 32, 1446-1482.

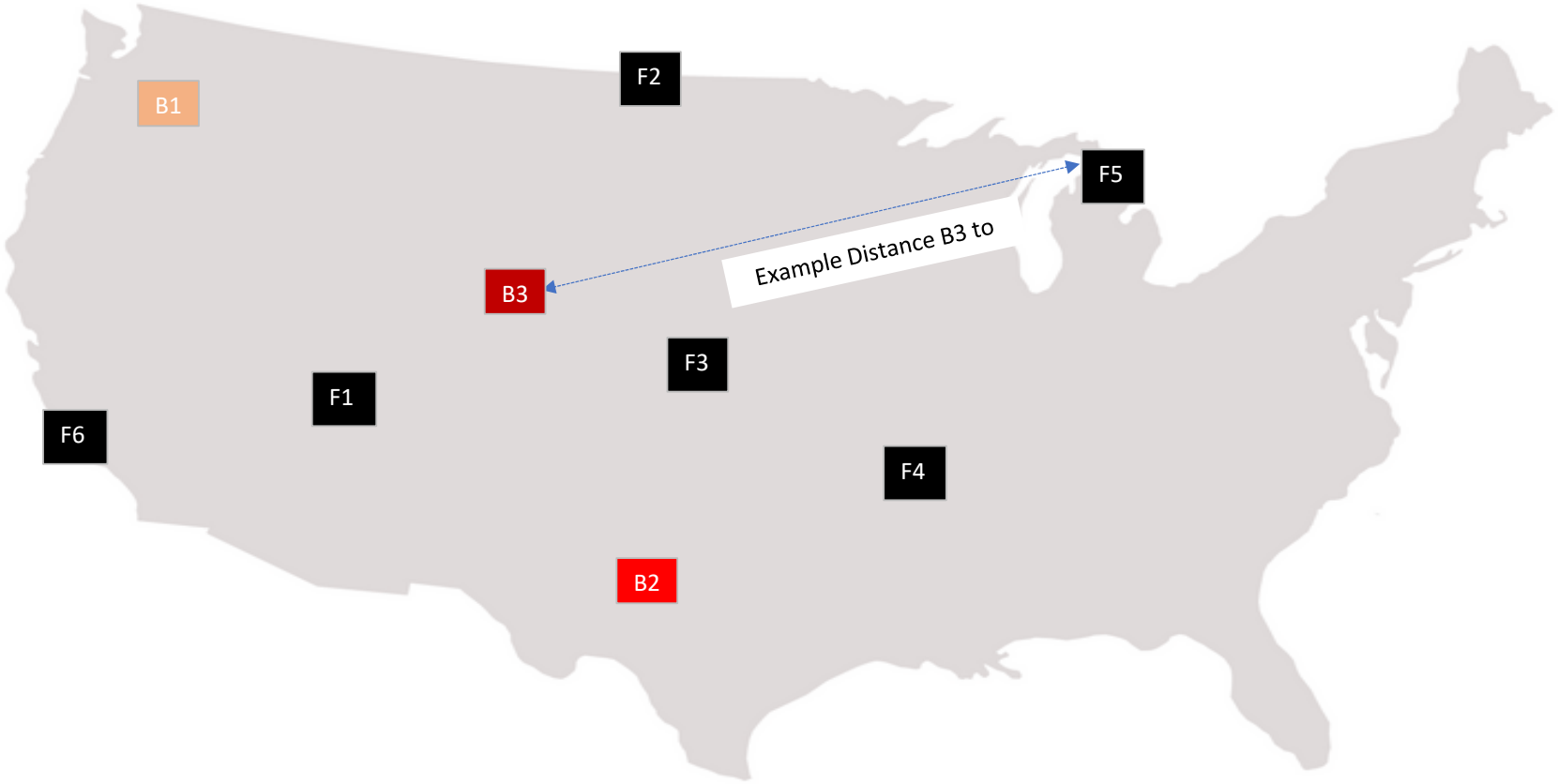
Ederington, L.H., Goh, J.C., 1998. Bond Rating Agencies and Stock Analysts: Who Knows What When? Journal of Financial and Quantitative Analysis. 33, 569-585.
Elyasiani, E., Goldberg, L.G., 2004. Relationship Lending: A Survey of the Literature. Journal of Economics and Business. 56, 315-330.
Fama, E.F., 1985. What's Different about Banks? Journal of Monetary Economics. 15, 5-29.
Faria-e-Castro, M., Paul, P., Sánchez, J.M., 2024. Evergreening. Journal of Financial Economics. 153, 103778.
Fischer, K.H., Banken und unvollkommener Wettbewerb. Empirische Beitrage zu einer Industrieokonomik der Finanzmarkte. Goethe University Frankfurt, Frankfurt, 2001.
Flannery, M.J., Kwan, S.H., Nimalendran, M., 2013. The 2007–2009 Financial Crisis and Bank Opaqueness. Journal of Financial Intermediation. 22, 55-84.
Granja, J., Leuz, C., Rajan, R.G., 2022. Going the Extra Mile: Distant Lending and Credit Cycles. Journal of Finance. 77, 1259-1324.
Greenstone, M., Mas, A., Nguyen, H.-L., 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, 200-225.
Gupta, M., Ongena, S., The Impact of the SBA Funding Programs on the Distance and Pricing of Loans to Small Businesses. Nottingham University Business School, Nottingham UK, 2022.
Hand, J., R. M., Holthausen, R.W., Leftwich, R.W., 1992. The Effect of Bond Rating Agency Announcements on Bond and Stock Prices. Journal of Finance. 47, 733-752.
Harvey, A.C., 1976. Estimating Regression Models with Multiplicative Heteroscedasticity. Econometrica. 44, 461-465.
Hauswald, R., Marquez, R., 2003. Information Technology and Financial Services Competition. Review of Financial Studies. 16, 921-948.
Hauswald, R., Marquez, R., 2006. Competition and Strategic Information Acquisition in Credit Markets. Review of Financial Studies. 19, 967-1000.
Herpfer, C., Schmidt, C., Mjøs, A., 2022. The Causal Impact of Distance on Bank Lending. Management Science. 69, 723-740.
Hirtle, B., 2006. Stock Market Reaction to Financial Statement Certification by Bank Holding Company CEOs. Journal of Money, Credit and Banking. 38, 1263-1291.
Iannotta, G., 2006. Testing for Opaqueness in the European Banking Industry: Evidence from Bond Credit Ratings. Journal of Financial Services Research. 30, 287-309.
Iannotta, G., 2011. Market Discipline in the Banking Industry: Evidence from Spread Dispersion. European Journal of Finance. 17, 111-131.
Ioannidou, V.P., Ongena, S., 2010. "Time for a Change": Loan Conditions and Bank Behavior When Firms Switch Banks. Journal of Finance. 65, 1847-1878.
Jakovljević, S., Degryse, H., Ongena, S., 2015. A Review of Empirical Research on the Design and Impact of Regulation in the Banking Sector. Annual Review of Financial Economics. 7, 423-443.
Jiménez, G., Ongena, S., Peydró, J.-L., Saurina, J., 2012. Credit Supply and Monetary Policy: Identifying the Bank Balance-Sheet Channel with Loan Applications. American Economic Review. 102, 2301-2326.
Jones, J.S., Lee, W.Y., Yeager, T.J., 2012. Opaque Banks, Price Discovery, and Financial Instability. Journal of Financial Intermediation. 21, 383-408.

Khwaja, A.I., Mian, A., 2008. Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market. American Economic Review. 98, 1413-1442.
King, M.R., Ongena, S., Tarashev, N., 2020. Bank Standalone Credit Ratings. International Journal of Central Banking. 16, 101-144.
Kliger, D., Sarig, O., 2000. The Information Value of Bond Ratings. Journal of Finance. 55, 2879-2902.
Kysucky, V., Norden, L., 2016. The Benefits of Relationship Lending in a Cross-Country Context: A Meta-Analysis. Management Science. 62, 90-110.
Leland, H.E., Pyle, D.H., 1977. Informational Asymmetries, Financial Structure, and Financial Intermediation. Journal of Finance. 32, 371-387.
Liberti, J.M., Petersen, M.A., 2019. Information: Hard and Soft. Review of Corporate Finance Studies. 8, 1-41.
Livingston, M., Naranjo, A., Zhou, L., 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, 874-888.
Nakamura, L.I., Roszbach, K., 2018. Credit Ratings, Private Information, and Bank Monitoring Ability. Journal of Financial Intermediation. 36, 58-73.
Ongena, S., Smith, D.C., Bank Relationships: a Review, in: P. Harker, S.A. Zenios, (Eds.), The Performance of Financial Institutions. Cambridge University Press, London, 2000, pp. 221-258.
Petersen, M.A., Rajan, R.G., 1994. The Benefits of Lending Relationships: Evidence from Small Business Data. Journal of Finance. 49, 3-37.
Petersen, M.A., Rajan, R.G., 2002. Does Distance Still Matter? The Information Revolution in Small Business Lending. Journal of Finance. 57, 2533-2570.
Plosser, M.C., Santos, J.A.C., 2018. Banks' Incentives and Inconsistent Risk Models. Review of Financial Studies. 31, 2080-2112.
Rajan, R.G., 1992. Insiders and Outsiders: the Choice between Informed and Arm's-Length Debt. Journal of Finance. 47, 1367-1400.
Sharpe, S.A., 1990. Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships. Journal of Finance. 45, 1069-1087.
Sironi, A., 2003. Testing for Market Discipline in the European Banking Industry: Evidence from Subordinated Debt Issues. Journal of Money, Credit and Banking. 35, 443-472.
Stein, I., 2015. The Price Impact of Lending Relationships. German Economic Review. 16, 367-389.
Stein, J.C., 2002. Information Production and Capital Allocation: Decentralized versus Hierarchical Firms. Journal of Finance. 57, 1891-1922.
Uchida, H., Udell, G.F., Watanabe, W., 2008. Bank Size and Lending Relationships in Japan. Journal of the Japanese and International Economies. 22, 242-267.
Veronis, N., The Importance of Private Markets. iCapital Institutional Capital Network, Inc., New York NY, 2023.
von Thadden, E.-L., 2004. Asymmetric Information, Bank Lending, and Implicit Contracts: the Winner's Curse. Finance Research Letters. 1, 11-23.
Xu, Y., Saunders, A., Xiao, B., Li, X., 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

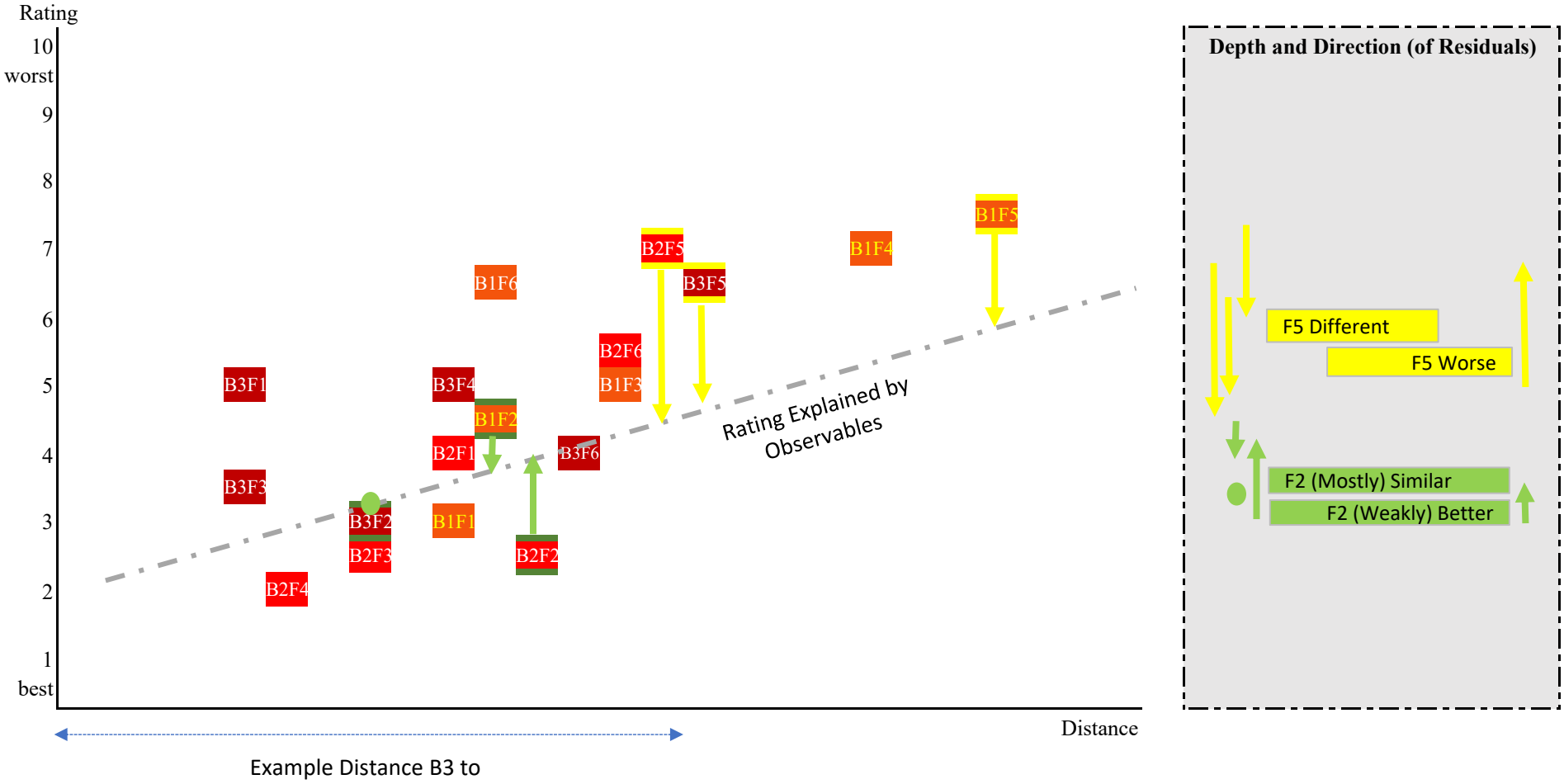


Table 1. Main Results: Bank Rating of Firm, Depth and Direction of Private Information

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 mean 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	*Internal 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.

Table 2. Panel A.	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Dependent Variable		*Loan Interest Rate Spread*				*Ln(Loan Maturity)*		
Independent variables								
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	
	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Dependent Variable		*Ln(Loan Amount)*				*d(Collateralized)*		
Independent variables								
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	

Table 2. Panel B.	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*
Dependent Variable		*Loan Interest Rate Spread*				*Ln(Loan Maturity)*		
Independent variables								
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	
	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Dependent Variable		*Ln(Loan Amount)*				*d(Collateralized)*		
Independent variables								
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. Impact of Private Information on Loan Performance

This table reports OLS regression estimates to assess how Depth of Private Information and Better or Worse Private Information affect loan performance. The 90 Days Past Due in 4 Quarters 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 column also report the impact of a one standard deviation change in depth or direction of private information on the loan performance, in percent scaled by the standard deviation of this loan performance. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3. Panel A.	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Dependent Variable	*90 Days Past Due in 4 Quarters*			
Independent variables				
Depth of Private Information	0.001***			*1.76%*
	(7.04)			
Better Private Information		-0.002***		*-0.63%*
		(-6.64)		
Worse Private Information			0.008***	*3.23%*
			(8.34)	
Bank-Firm, Bank and Firm Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes	
Observations	1,069,515	1,069,515	1,069,515	
Adjusted R-squared	0.007	0.005	0.019	

Table 3. Panel B.	(1)	(2)	(3)	*Impact of one standard deviation increase on dependent variable as percent of its standard deviation*
Dependent Variable	*90 Days Past Due in 4 Quarters*			
Independent variables				
Depth of Private Information	0.000***			*0.00%*
	(3.80)			
Better Private Information		-0.000*		*0.00%*
		(-1.76)		
Worse Private Information			0.002***	*0.81%*
			(4.33)	
Bank-Firm, Bank and Firm Controls, and Bank and Firm * Year Fixed Effects	Yes	Yes	Yes	
Observations	1,006,071	1,006,071	1,006,071	
Adjusted R-squared	0.721	0.721	0.721	

Table 4. Bank-Firm Relationship Length and Banks' Private Information

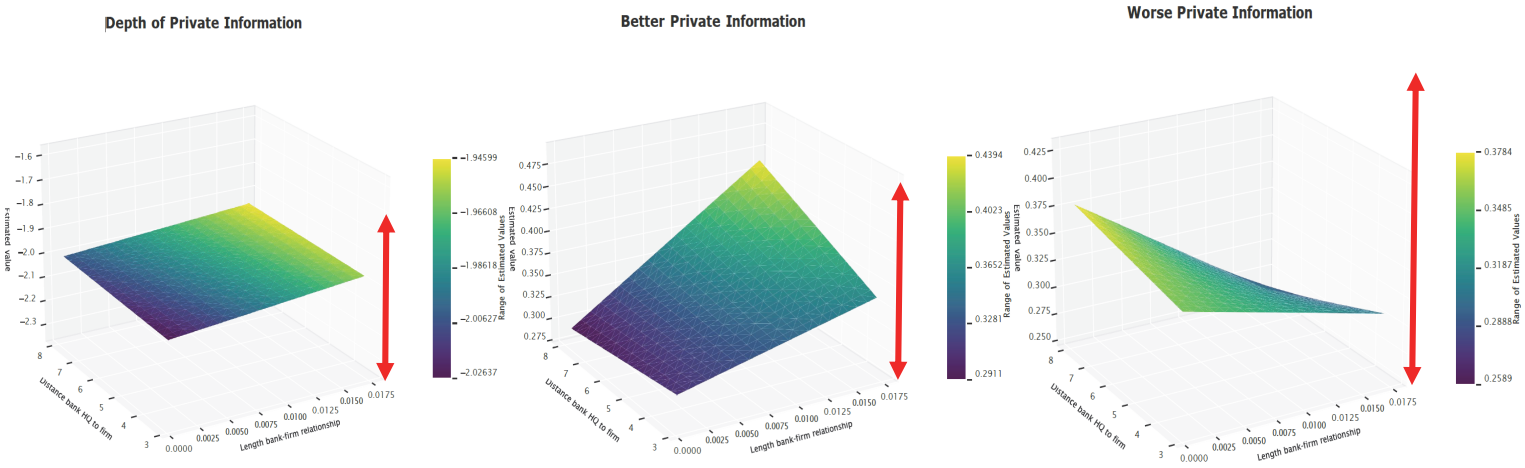
The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to non-listed firms. The number of observations equals 2,715,622. The dependent variables are: in Models (1) and (4) the Depth of Private Information, which is the natural log of the squared residuals; and, in Models (2), (3), (5) and (6) 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 mean 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						
	0.6%	6.0%	-3.8%	7.0%	1.3%	8.5%
				9.7%	8.0%	9.4%
				8.3%	17.7%	-0.3%

Table 5. 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 2,713,281. 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 mean 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%.

Table 5. Panel A.	(1)	(2)	(3)
Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables			
Distance bank HQ to firm	0.004	-0.005***	0.006***
	(0.71)	(-3.23)	(3.46)
Length bank-firm relationship	4.151	-0.568	0.142
	(1.00)	(-0.51)	(0.12)
Distance bank HQ to firm * Length bank-firm relationship	-0.075	1.182***	-0.909***
	(-0.11)	(6.12)	(-4.58)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.016	0.007



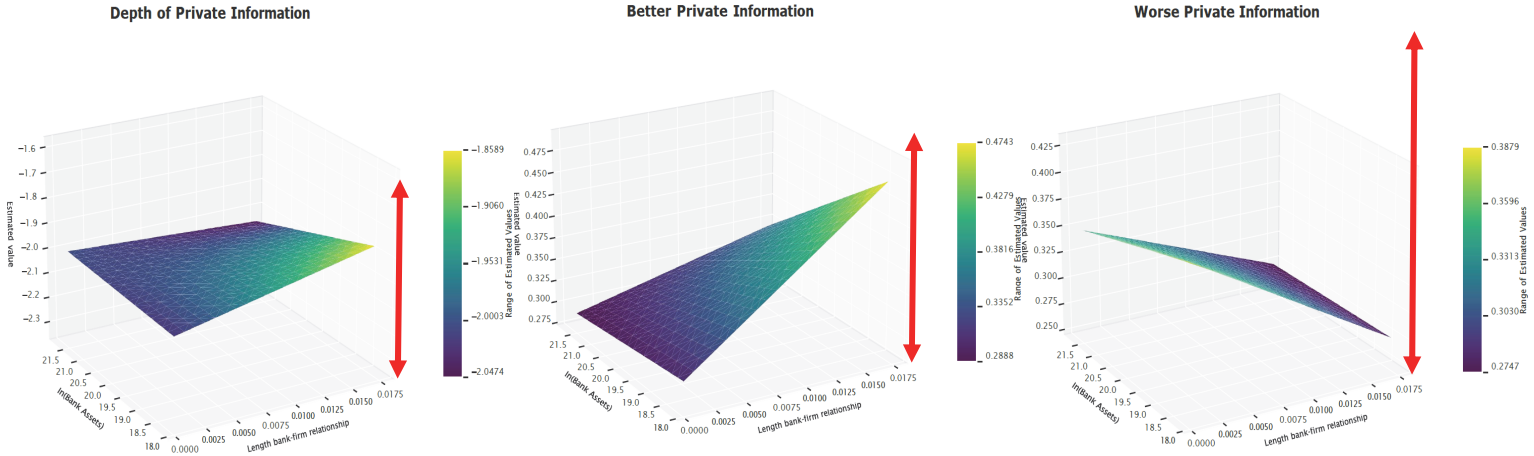
The figures plot the Depth of Private Information and Better and Worse Private Information for the length of the bank-firm relationship and its interactions with distance bank HQ to firm. All other variables are set at their median. The red arrows indicate fifty percent of the standard deviation of the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

Table 5. Panel B.	(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***
	(-1.67)	(-12.12)	(10.27)
Length bank-firm relationship	3.859***	6.774***	-5.435***
	(3.44)	(19.79)	(-20.55)
COVID-19 * Length bank-firm relationship	-4.666**	0.727*	-2.714***
	(-2.43)	(1.81)	(-5.95)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.016	0.007

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

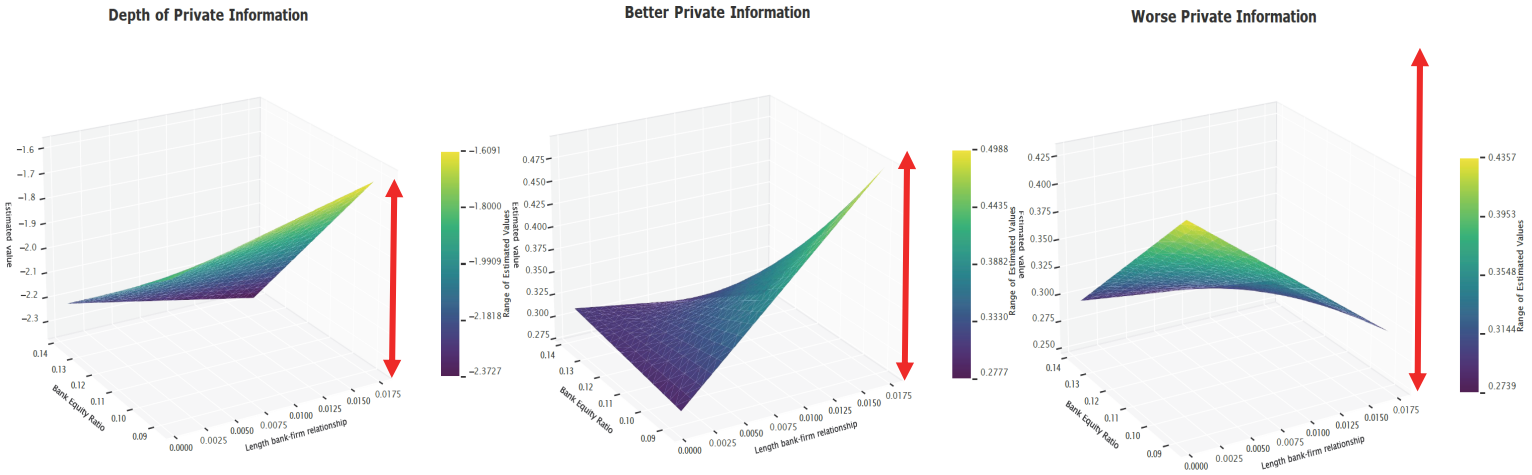
The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to non-listed firms. The number of observations equals 2,713,281. The dependent variables are: the Depth of Private Information, which is the natural log of the squared residuals; and, the 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 mean 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%.

Table 6. Panel A.	(1)	(2)	(3)
Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables			
Ln(Bank assets)	0.006	-0.003*	-0.012***
	(0.82)	(-1.81)	(-5.90)
Length bank-firm relationship	70.644***	45.411***	-19.112***
	(4.08)	(9.04)	(-4.46)
Ln(Bank assets) * Length bank-firm relationship	-3.368***	-1.943***	0.684***
	(-3.88)	(-7.73)	(3.13)
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.

Table 6. Panel B.	(1)	(2)	(3)
	Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables			
Bank equity ratio	-7.739***	0.482***	-2.762***
	(-14.10)	(3.17)	(-20.96)
Length bank-firm relationship	48.337***	36.306***	-19.288***
	(6.92)	(16.08)	(-14.16)
Bank equity ratio * Length bank-firm relationship	-417.194***	-275.836***	128.730***
	(-6.54)	(-14.64)	(10.74)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes
Adjusted R-squared	0.010	0.020	0.008

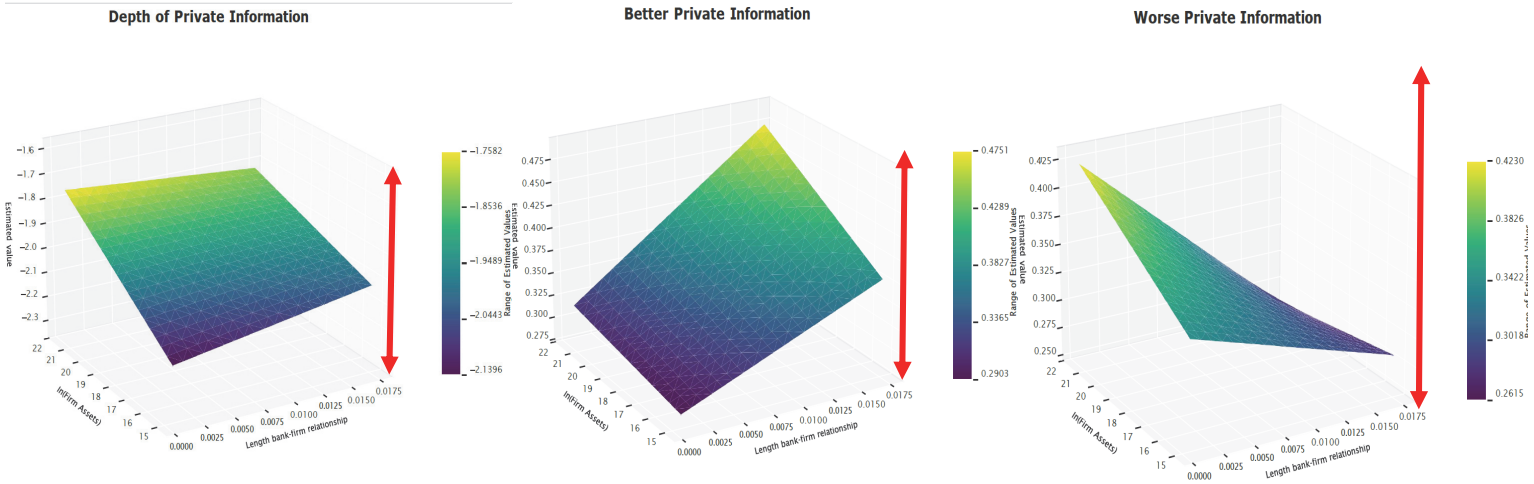


The figures plot the Depth of Private Information and Better and Worse Private Information for the length of the bank-firm relationship and its interactions with the bank equity ratio. All other variables are set at their median. The red arrows indicate fifty percent of the standard deviation of the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

Table 7. Bank-Firm Relationship Length, Firm Size and Leverage and Banks' Private Information

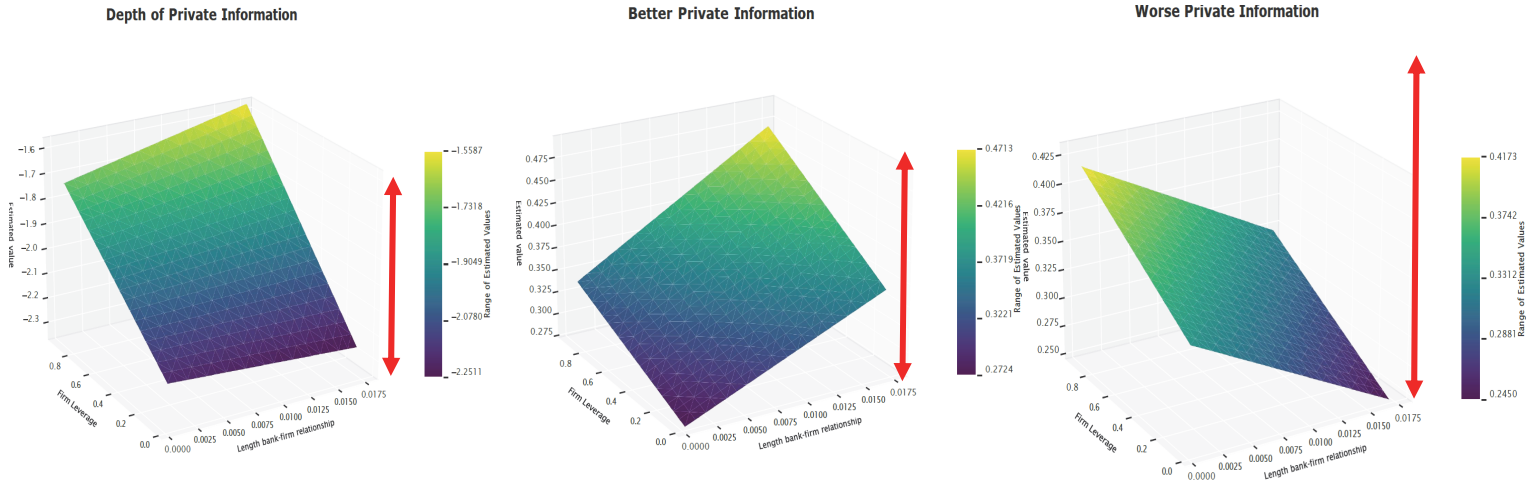
The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to non-listed firms. The number of observations equals 2,713,281. The dependent variables are: the Depth of Private Information, which is the natural log of the squared residuals; and, the 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 mean 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%.

Table 7. Panel A.	(1)	(2)	(3)
Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables			
Ln(Firm assets)	0.057***	0.003	0.012***
	(10.03)	(1.51)	(7.28)
Length bank-firm relationship	31.655**	-3.810	10.107***
	(2.56)	(-0.80)	(2.79)
Ln(Firm assets) * Length bank-firm relationship	-1.638**	0.620**	-0.915***
	(-2.18)	(2.17)	(-4.18)
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.

Table 7. Panel B.	(1)	(2)	(3)
	Dependent Variable	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables			
Firm leverage	0.521***	0.074***	0.091***
	(15.91)	(9.48)	(11.18)
Length bank-firm relationship	-3.624*	5.541***	-5.441***
	(-1.85)	(9.74)	(-10.68)
Firm leverage * Length bank-firm relationship	15.076***	2.563**	-0.143
	(4.42)	(2.38)	(-0.17)
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 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 mean equation	Own calculations
Residual	*Depth of Private Information*	-	Natural log of the squared residuals from the mean equation estimated	Own calculations
	Better Private Information	-	Negative residual from the mean equation	Own calculations
	Worse Private Information	-	Positive residual from the mean equation	Own calculations
	Depth of Private Information (Plus Squared Terms)	-	Natural log of the squared residuals from the mean equation estimated (with squared terms added)	Own calculations
	Better Private Information (Plus Squared Terms)	-	Negative residual from the mean equation (with squared terms added)	Own calculations
	Worse Private Information (Plus Squared Terms)	-	Positive residual from the mean equation (with squared terms added)	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
Loan Performance	*90-days Past Due in 4-Quarters*	0/1	= 1 if the loan is 90-days past due, = 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. For operational reasons the summary statistics for the 90-Days Past Due in 4-Quarters are approximated based on public statistics collected from https://en.macromicro.me/.

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	-	0.000	-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	4.945	0.451
	Worse Private Information	-	0.298	0	0	6.436	0.577
	Depth of Private Information (Plus Squared Terms)	-	-2.16	-1.539	-33.365	3.595	2.757
	Better Private Information (Plus Squared Terms)	-	0.33	0.041	0	4.825	0.48
	Worse Private Information (Plus Squared Terms)	-	0.33	0	0	6.036	0.719
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
Loan Performance	*90-days Past Due in 4-Quarters*	0/1	0.021	0	0	1	0.143
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

Table A3. Impact of Private Information on Loan Performance, 30 Days Past Due in 4 Quarters

This table reports OLS regression estimates to assess how Depth of Private Information and Better or Worse Private Information affect loan performance. The 30 Days Past Due in 4 Quarters 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. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A3. Panel A.	(1)	(2)	(3)
Dependent Variable	*30 Days Past Due in 4 Quarters*		
Independent variables			
Depth of Private Information	0.002***		
	(7.30)		
Better Private Information		-0.003***	
		(-3.69)	
Worse Private Information			0.015***
			(10.40)
Bank-Firm, Bank and Firm Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes
Observations	1,069,515	1,069,515	1,069,515
Adjusted R-squared	0.008	0.005	0.018

Table A3. Panel B.	(1)	(2)	(3)
Dependent Variable	*30 Days Past Due in 4 Quarters*		
Independent variables			
Depth of Private Information	0.000***		
	(2.62)		
Better Private Information		-0.002***	
		(-2.62)	
Worse Private Information			0.003***
			(4.71)
Bank-Firm, Bank and Firm Controls, and Bank and Firm * Year Fixed Effects	Yes	Yes	Yes
Observations	1,006,071	1,006,071	1,006,071
Adjusted R-squared	0.684	0.684	0.684

Appendix Table A4. Impact of (alternative measure of) Private Information on Loan Terms (Adding Squared 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. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A4. Panel A.	(1)	(2)	(3)	(1)	(2)	(3)
Dependent Variable	*Loan Interest Rate Spread*			*Ln(Loan Maturity)*		
Independent variables						
Depth of Private Information (squared term in the first stage)	0.021***			-0.001		
	(4.76)			(-0.73)		
Better Private Information (squared term in the first stage)		-0.413***			0.052***	
		(-15.89)			(3.90)	
Worse Private Information (squared term in the first stage)			0.372***			-0.041***
			(24.41)			(-6.45)
Bank-Firm, Bank and Firm Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,751,495	1,751,495	1,751,495	2,267,750	2,267,750	2,267,750
Adjusted R-squared	0.121	0.129	0.139	0.231	0.231	0.232
	(1)	(2)	(3)	(1)	(2)	(3)
Dependent Variable	*Ln(Loan Amount)*			*d(Collateralized)*		
Independent variables						
Depth of Private Information (squared term in the first stage)	0.003			-0.003***		
	(1.58)			(-8.41)		
Better Private Information (squared term in the first stage)		0.024			-0.043***	
		(1.48)			(-7.55)	
Worse Private Information (squared term in the first stage)			-0.007			0.015***
			(-0.90)			(8.73)
Bank-Firm, Bank and Firm Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,269,573	2,269,573	2,269,573	2,269,573	2,269,573	2,269,573
Adjusted R-squared	0.404	0.404	0.404	0.149	0.153	0.149

Table A4. Panel B.	(1)	(2)	(3)	(1)	(2)	(3)
Dependent Variable	*Loan Interest Rate Spread*			*Ln(Loan Maturity)*		
Independent variables						
Depth of Private Information (squared term in the first stage)	0.002			0.000		
	(1.29)			(0.52)		
Better Private Information (squared term in the first stage)		-0.085***			0.012**	
		(-4.99)			(2.11)	
Worse Private Information (squared term in the first stage)			0.086***			-0.002
			(5.91)			(-2.11)
Bank-Firm, Bank and Firm Controls, and Bank and Firm * Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,658,326	1,658,326	1,658,326	2,171,738	2,171,738	2,171,738
Adjusted R-squared	0.778	0.778	0.779	0.703	0.703	0.703
	(1)	(2)	(3)	(1)	(2)	(3)
Dependent Variable	*Ln(Loan Amount)*			*d(Collateralized)*		
Independent variables						
Depth of Private Information (squared term in the first stage)	0.000			0.000		
	(0.57)			(0.48)		
Better Private Information (squared term in the first stage)		0.037***			-0.011***	
		(5.19)			(-4.22)	
Worse Private Information (squared term in the first stage)			-0.021***			0.009***
			(-6.99)			(7.22)
Bank-Firm, Bank and Firm Controls, and Bank and Firm * Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,173,487	2,173,487	2,173,487	2,173,487	2,173,487	2,173,487
Adjusted R-squared	0.708	0.709	0.709	0.688	0.688	0.689

Appendix Table A5. Impact of (alternative measure of) Private Information on Loan Performances (Adding Squared Terms)

This table reports OLS regression estimates to assess how Depth of Private Information and Better or Worse Private Information affect loan performance. The indicated loan performance 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. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A5. Panel A.	(1)	(2)	(3)
Dependent Variable	*90 Days Past Due in 4 Quarters*		
Independent variables			
Depth of Private Information	0.001***		
	(7.09)		
Better Private Information		-0.001***	
		(-5.95)	
Worse Private Information			0.007***
			(7.98)
Bank-Firm, Bank and Firm Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes
Observations	943,396	943,396	943,396
Adjusted R-squared	0.007	0.005	0.017

Table A5. Panel B.	(1)	(2)	(3)
Dependent Variable	*90 Days Past Due in 4 Quarters*		
Independent variables			
Depth of Private Information	0.000***		
	(3.62)		
Better Private Information		-0.000	
		(-0.90)	
Worse Private Information			0.001***
			(3.99)
Bank-Firm, Bank and Firm Controls, and Bank and Firm * Year Fixed Effects	Yes	Yes	Yes
Observations	887,531	887,531	887,531
Adjusted R-squared	0.719	0.719	0.719

Table A6. Bank-Firm Relationship Length and Banks' Private Information (Adding Squared Terms)

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 2,715,622. The dependent variables are: in Models (1) and (4) the Depth of Private Information, which is the natural log of the squared residuals; and, in Models (2), (3), (5) and (6) 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 mean 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	2.123*	3.760***	-3.381***			
	(1.69)	(10.96)	(-10.23)			
Length bank-firm relationship (0.25 << 3 years)				0.210***	0.005**	0.059***
				(15.61)	(2.04)	(19.19)
Length bank-firm relationship (3 << 5 years)				0.319***	0.033***	0.072***
				(17.90)	(10.14)	(17.06)
Length bank-firm relationship (> 5 years)				0.267***	0.056***	0.033***
				(14.33)	(14.82)	(7.58)
Bank-Firm, Bank and Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,269,573	2,269,573	2,269,573	2,269,573	2,269,573	2,269,573
Adjusted R-squared	0.03	0.02	0.046	0.03	0.02	0.046