RESEARCH

Swiss Finance Institute
Research Paper Series
N°23-01

"If You Don't Know Me by Now ..." Banks' Private Information and Relationship Length

Stijn Claessens
Yale School of Management and CEPR

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

Teng Wang
Board of Governors of the Federal Reserve System

“If You Don't Know Me by Now ...”

Banks’ Private Information and Relationship Length

Stijn Claessens, Steven Ongena, Teng Wang*

May 13, 2024

Does the private information banks generate about their corporate borrowers deepen and change in nature over time, and if so, how? Exploiting the comprehensive Federal Reserve’s supervisory dataset, we distinguish two dimensions to the private information embedded in internal credit ratings: *depth* and *direction (better* or *worse),* which we confirm to correlate with loan terms. Longer firm-bank relationships deepen private information in both directions, with effects often strongly nonlinear and peaking at about five years. Learning effects are particularly salient for smaller and leveraged firms, smaller, leveraged, and illiquid banks, at longer firm-bank distances, and during non-COVID times. (98 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@bf.uzh.ch; phone: +41 44 634 39 54. Wang is with the Board of Governors of the Federal Reserve System; 20th St and Constitution Ave NW, Washington, DC 20551, USA; email: teng.wang@frb.gov; phone: +1 202 452 6436. For comments, we are grateful to Paul Calem, Sheng Li, Hirofumi Uchida, Greg Weitzner, seminar participants at the Bank for International Settlements (Basel), Federal Reserve Board, and conference participants at the Banking and Finance Workshop (Tokyo) and the European Economic Association Annual Meetings. The paper circulated previously titled “Out of Sight, Out of Mind?” Banks’ Private Information, Distance, and Relationship Length.” The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Bank for International Settlements, the Board of Governors, or any staff of these institutions.

What is in the banks' private information set about corporate borrowers? Is this information relevant to loan terms? How does it change over time, and how is it shaped by firm and/or bank characteristics, the physical distance between firms and banks (e.g., Hauswald and Marquez (2003)), and loan characteristics and/or external circumstances that affect the modus communicandi and economic conditions (like COVID-19)?

In this paper, we define two novel proxies of banks' private assessment of corporate borrowers' riskiness, ***depth**,* i.e., banks' assessments of firm risk relative to one based on observables, and its ***direction***, i.e., *better* or *worse* assessments. We then analyze the factors that drive the similarities and differences in the depth and the direction of the banks' private information sets. Collectively, this work aids in analyzing the intricate process of how banks assess the creditworthiness of various corporate borrowers through private information collection over time.

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 gap in research is that unlike public or hard information, which can be easily observed and measured, private or soft information is often obscure and only accessible to the bank making the lending decision (as noted by Liberti and Petersen (2018)).

In our study, we derive proxies of the banks' confidential information set using internal bank credit ratings reported, along with other detailed loan and borrower information, in the Federal Reserve's supervisory Y-14Q quarterly loan-level dataset. Besides much detail, this dataset has a much broader representation than other datasets in the US (e.g., Berger, Bouwman, Norden, Roman, Udell and Wang (2021); Faria-e-Castro, Paul and Sánchez (2021); Beyhaghi, Howes and Weitzner (2023)) as it covers all commercial loans of $1 million or more extended by the largest U.S.-based bank holding companies (BHCs) that are subject to the DFAST/CCAR stress tests.[1] Critically, this dataset uniquely reports banks' internal rating of these loans, which is the basis on which we identify in this study banks' private information. We investigate how the *depth* and *direction* of the lending banks' private information is formed over time through the length of their relationship with firms and is determined by the characteristics of the firm, bank, and/or distance between them.

Our focus on banks' private information as to the risks of their borrowers, derived from banks' internal credit ratings, differs from much of the literature that has focused on outcome variables such as the amounts of lending and the terms thereof. Besides investigating the drivers of depth and direction, we also document the implications of these two dimensions of banks' private information on the terms of loans granted.

In the spirit of the seminal work by Morgan (2002) and Agarwal and Hauswald (2010),[2] we employ a heteroskedastic regression model (Harvey (1976)) to extract from the internal bank credit ratings the part that is not explained by an encompassing set of observables as a measure of banks' private information about the borrowing firm.[3] We employ (the logarithm of) the squared of this residual to capture the *depth* in the banks' private information sets about firms (the larger the dispersion or incongruity, the more depth there is to the private information set). And we use the positive and negative residuals, *better*

---

[1] When it was put in place in 2011:Q3 all banks with over $50 billion in assets were required to report, but the Economic Growth, Regulatory Relief, and Consumer Protection Act (EGRRCPA) increased the reporting size threshold to $100 billion starting in 2019:Q4. Our results are robust, however, to only using banks with over $100 billion in assets in all our analyses.

[2] Using the dispersion in the public credit ratings of banks, Morgan (2002) establishes that banks are more opaque than non-financial institutions. See also 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 matter for explaining 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.

[3] The observables are hard information (public and otherwise), while the banks` private information set will likely include both hard and soft elements (e.g., Liberti and Petersen (2018)).

and worse respectively, as indicators of the direction of the banks' private information.[4] We first relate depth and direction to bank and firm characteristics when the loans get made and show that these measures of banks' private information sets are consistent with priors and the existing literature focusing on drivers of loan terms.

For our main analysis, we first focus on relationship length 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 valuable private information about the firm will be produced, and the banks' internal rating will deviate more from the one based on hard information only, making its depth higher.[5] As to the direction dimensions of private information, the literature is split on the effects of the length of a relationship:[6] we conjecture that valuable private information acquired during a relationship can cause the bank's evaluation to improve from the one based solely on hard information and relative to other banks, resulting in higher *better* and lower *worse* private information.

In the second part of our main analysis, we study how the information impact of relationships varies with firm, bank, and loan characteristics.[7] Consistent with theoretical literature, yet not empirically documented, we expect the various effects for the length of

[4] Figure 1 provides an illustrated example conveying the main intuition of these measures for the variable distance between the bank and the firm, but the intuition also applies to the other firm and bank dimensions.
[5] Appendix 1 provides an illustration of how the two informational components, i.e., a hard component and a soft component, could determine the impact of the distance and the bank-firm relationship length on the depth and the favorability of bank ratings.
[6] In, e.g., Sharpe (1990), Rajan (1992) and von Thadden (2004), loan repayment allows the "inside banks" to distinguish high- from low-quality firms, with low-quality firms more likely to switch to "outside banks".
[7] While Beyhaghi, Howes and Weitzner (2023) study how changes in losses privately expected by banks predict firms' future stock returns, bond returns, and earnings surprises, they do not study the process of learning.

relationship to differ by firm, bank, and loan characteristics, including bank-firm distance, as both information asymmetries as well as abilities and incentives for private information acquisition vary. For example, as a relationship lasts longer, we expect the adverse impact of distance on depth (e.g., Hauswald and Marquez (2003) and Hauswald and Marquez (2006)) to decrease as the bank produces valuable private information from its engagement with the firm, and the likelihood of the internal rating being better (relative to observable factors) to increase and it being worse to decrease. We also expect that private information acquisition will be faster for some types of firms and certain banks, and that the COVID period interrupted the process of learning.

In terms of the paper's sequence, we first estimate depth and directions and show how they vary at the initiation of loans with firm and bank characteristics, including distance. Next, and consistent with our conjectures, we show that the length of the bank-firm relationship contributes to the depth of private information. In other words, as relationships between banks and firms progress, the banks' assessments of the firms deviate more from those based on the available hard information, suggesting that banks learn about the borrower during their relationship. This increase in depth goes hand in hand with an increasingly better assessment of banks of the borrowing firms, providing support for the notion that relationship lending contributes to private information (e.g., Petersen and Rajan (1994)). We refer to this effect as the *"premium through relationship,"* which is built up over time with a peak impact at about five years.

Next in our main analysis, we study how firm, bank, and bank-firm characteristics affect the banks' private information production process. We find that relationships particularly affect depth and direction for smaller and leveraged firms, smaller, leveraged, and illiquid banks, longer firm-bank distances, and during non-COVID times when onsite visits and face-to-face meetings of bankers and customers are possible. But interestingly it does not differ for green or brown firms. Finally, in a validation exercise, we confirm that our three dimensions of banks' private information play a role in setting the terms of their loans. The loan interest rate spread increases in depth and is lower (higher) when the private information is positive (negative). Also, as expected, maturity, amount, and collateralization lengthen, increase, and decrease, respectively, in positive information and in the opposite ways for negative information.

With these 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.

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 tests 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 publicly available information from a credit bureau 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. Our findings on how the depth and direction of ratings changes with physical distance and length of relationships suggest that banks can overcome with longer relationships to some extent the informational challenges posed by firm and bank characteristics,[8] distance, or modus communicandi (e.g., during COVID).

---

[8] 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

Next, 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 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.

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 here 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, how they are affected by relationship length and its interactions with firm, bank, and bank-firm characteristics, including distance, and then confirming that these measures map into the terms of loans.

In this respect, we also contribute to an ever-growing empirical literature that has documented that the intensity of distance-related credit rationing affecting firms may vary by country, period, governmental lending programs, transportation infrastructure, and/or

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 broadening the sample to around 3,400,000 loan-firm-bank-quarter observations over the period 2012M9 to 2021M3 (using the Y-14Q data set that contains all loans above $1 million granted by between 27 and 33 reporting banks).

[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, Nogueira and Ongena (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.

the characteristics of local (bank) competitors.[11] And Degryse, Laeven and Ongena (2009) show that the lending bank's geographical reach is determined not only by its own organizational structure but also by organizational choices made by its rivals. They find that the geographical footprint of the lending bank is smaller when rival banks are relatively larger and more hierarchically organized (and may rely relatively more on hard information). We contribute to this literature by highlighting the impacts of firm, bank, and loan characteristics, including distance, on the quality of private information 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)),[12] 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.

The rest of the paper proceeds as follows. Section I introduces the methodology. Section II introduces the data. Section III reports the main findings, as well as the confirmation of the relevance of the private information measures for loan terms. Section IV concludes.

---

[11] 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)).

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

I. Methodology

To identify the determinants of the congruity of bank credit ratings, we employ a regression model with multiplicative heteroskedasticity as introduced by Harvey (1976). The heteroskedastic version extends the linear regression model by also parametrizing the unexplained variance as a function of exogenous covariates.[13]

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

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

and

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

Equation (1) will be referred to as the "mean equation", while (2) will be labeled the "variance equation". The identifying assumptions are:

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

and

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

y_i is the dependent variable, i.e., the internal bank credit rating, $\boldsymbol{x}_i$ is a vector of explanatory variables in the mean equation that includes a constant, and u_i is a disturbance term. The variance of the error term is an exponential function of a vector of individual-specific attributes denoted by $\boldsymbol{z}_i$. Although other functional forms of heteroskedasticity can be used, the exponential form is particularly convenient because it ensures positive variance.

[13] Our discussion is based on Cerqueiro, Degryse and Ongena (2013). For other applications, see also Gaul and Stebunovs (2009), Cerqueiro, Degryse and Ongena (2011), Iannotta (2011), Iannotta and Navone (2012), or Baele, De Bruyckere, De Jonghe and Vander Vennet (2014).

The interpretation of γ is crucial for our intended analysis here. Pick one variable from the vector $\mathbf{z}$, say, z^k, and the respective parameter, γ^k. A positive γ^k indicates that the precision of the credit rating model decreases in z^k. One can interpret such a result as evidence of a positive correlation between the variable z^k and the weight of the difficulties in arriving at a precise rating in the rating-setting process. When $\gamma^k = 0$, the error term is homoscedastic and its variance equals σ^2.

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

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

The simplest procedure to estimate the heteroskedastic regression model is to estimate the parameters in the mean equation by OLS and to use the squared errors as raw estimates of the individual variances. Then, one obtains estimates of the parameters in the variance equation by regressing the (logarithm of) squared errors on the set of covariates in the vector $\mathbf{z}$. This procedure is computationally simpler than obtaining the maximum-likelihood estimates in the heteroskedastic regression model. But there is a loss of efficiency in this two-step procedure (Harvey (1976)), which in our application with many observations is a price we are willing to incur.

Alternatively, in case one only has access to fewer observations, one can obtain maximum-likelihood estimates in the heteroskedastic regression model by maximizing the following log-likelihood function with respect to the vector of parameters β and γ:

$$log\mathcal{L}(\beta,\gamma|\boldsymbol{X},\boldsymbol{Z})$$
$$= \frac{N}{2}log(2\pi\sigma^2) - \frac{1}{2}\sum_{i=1}^{N}\boldsymbol{z}_i'\gamma - \frac{1}{2}\sum_{i=1}^{N}e^{-\boldsymbol{z}_i'\gamma}\left(\frac{y_i - \boldsymbol{x}_i'\beta}{\sigma}\right)^2. \quad (6)$$

From a theoretical perspective, the maximum likelihood estimators for the parameters in the mean and variance equations are, in expectation, uncorrelated (see Harvey (1976)). To see this, consider the case in which a single covariate x that affects both the mean and variance of y. The estimator for β that arises from maximizing the log-likelihood function is:

$$\hat{\beta} = \frac{\sum_{i=1}^{N}\left(\frac{x_i y_i}{\hat{\sigma}_i^2}\right)}{\sum_{i=1}^{N}\left(\frac{x_i^2}{\hat{\sigma}_i^2}\right)}, \quad (7)$$

which is the well-known weighted-least squares (WLS) estimator. In this estimation method, the contribution of each observation in the sum of squares is weighted by the inverse of its estimated variance (i.e., its precision):

$$\frac{1}{\hat{\sigma}_i^2} = \frac{1}{\hat{\sigma}^2 e^{\hat{\gamma}x_i}}. \quad (8)$$

In practice, estimation via WLS requires that one specifies the pattern of heteroskedasticity and estimates the individual variances. Virtually, all empirical applications assume the multiplicative heteroskedasticity model. When heteroskedasticity is present in the data, the WLS estimator for β will in general differ from the OLS estimator, because WLS shifts weight from high-variance to low-variance observations. As a result, the difference between OLS and WLS estimators is a direct consequence of heteroskedasticity. However, the parameters in the variance equation, γ, are simply factor loadings capturing variation in the residual variance that otherwise would be averaged out in $\hat{\sigma}^2$. Therefore, γ does not systematically affect β.[14]

II. Data

A. Banks' internal risk rating

As part of its lending (approval) and monitoring processes, a bank typically assesses the credit quality of its borrowers for which is uses an internal credit risk rating scale. The methodology and data used for the development of ratings can vary greatly between banks.

To communicate credit risk externally, including in the context of underwriting or renegotiating loans with borrowers, or to disclose ratings to market participants and supervisory agencies, banks often map these internal credit ratings to an externally comparable, commonly used rating scale. This conversion typically involves comparing the default experience of loans in their own internal credit rating categories to assets with similar default experiences that have public ratings.[15]

The converted, externally comparable rating allows then for cross-bank comparison of credit quality of bank assets.[16] Such a straightforward comparison is not possible in other bank (loan) data sets such as the Call Reports, DealScan, or the Shared National Credit

---

[14] Our two-step procedure can in principle also be applied iteratively, where the mean equation is re-estimated by WLS using as weights the inverse of the estimated variances. Through this iterative process the estimates obtained in both equations converge to the maximum likelihood estimates.

[15] In the context of the Y14 data used in the supervisory stress tests, the Federal Reserve receives banks' concordance maps that translate their internal ratings to a common S&P-like rating scale (the scale is the one required by the Fed, and the Fed can also unilaterally make certain minor adjustments). It is possible that banks use the same concordance map when they communicate with market participants.

[16] 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 probability of defaults and are within the range of default probability of public-rated A ("single A") loans, then the internal ratings of both loans would have a converted external rating of A.

Program. The resultant standardized bank rating of firms we employ ranges from 1 (= best) to 10 (= worst).

B. Data on loans, relationships, and banks

Our primary data source is the Federal Reserve's Y-14Q reports, a quarterly collection of banks' holding of commercial and industrial (C&I) loan data collected by the Federal Reserve beginning in the fall of 2011.[17] Our data cover all C&I loans in size more than $1 million when originated and held by the largest bank holding companies in the U.S. by assets.[18] Loan sizes range from the $1 million reporting threshold (in commitment) to billions of dollars, thus covering the spectrum from loans to SMEs to large, listed corporations. 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 data on firm financials and performance, including total assets, ROA, and leverage. It includes as well as identification of the borrower, allowing us to calculate the distance between bank branch (or HQ) and firm.

The Y-14Q data not only covers 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 September 2012 to March 2021. For each quarter, we consider loans recorded on banks' balance sheets and apply the following filters to provide a clean sample. We eliminate all loans to other financial institutions and governments (NAICS code of 52 and 92). We also drop loans with a committed exposure below $1 million, the official minimum size requirement to be

[17] 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.
[18] The number of reporting banks varies over our sample period between 27 and 33.

included in the Y-14Q. Schedule H.1 explicitly excludes “small business loans” — loans that are evaluated based on borrower, not the firm, credit quality or rated on a different scale than other corporate loans. For consistency, we drop all loans reported with “a small business” as their line of 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.[19] As of 2019:Q4, these loans covered more than 60 percent of the balances of all commercial and industrial (C&I) loans as reported in the Federal Reserve’s Y-9C.[20]

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 statement information — including bank age, size, liquidity, profitability, and capital ratio — for U.S. holding companies and the branches of foreign companies that operate in the U.S.

III. Findings

A. Mean equation

The dependent variable in the mean equation is the *Standardized bank rating of firm* which is defined as the rating given to the firm by the bank transposed to a common scale. This information is sourced from the FR Y-14Q Schedule H.1.

The mean equation is mainly there to predict — as good as possible — the bank rating based on observable information, including when dealing with new customers, defined as those for which the length of the bank-firm relationship is less than a quarter (of a year).

---

[19] Credit ratings frequently change as the bank-firm relationship goes on (and the bank-firm exposure is 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).

[20] Federal Reserve Board, Form FR Y-9C, provides the Consolidated Financial Statements for Bank Holding Companies.

We therefore include a comprehensive set of firm and bank controls,[21] and bank-firm distance, disregarding for example potential multicollinearity and bad control issues. As firm controls, the mean equation includes: *Ln(Firm assets)*, *Firm ROA*, *Firm leverage*, and three indicator variables for whether the firm is *Public*, *Green*, or *Brown* (recall that Appendix Table A1 contains the precise definitions of all variables). As bank controls, *Ln(Bank assets)*, *Bank equity ratio*, *Bank NPL ratio*, *Bank liquid asset ratio*, and *Bank ROA*. Since the fitted estimates only reflect the observable information, any banks' private assessment of the firm is captured in the residual.

The estimated coefficients are in Table 1, in Model (1). The sign and size of these estimated coefficients are straightforward and reasonable though not of our immediate concern (our main findings are unaffected by variations in the set of controls). Better ratings (i.e., lower numbers) are received by large, profitable, lowly leveraged, public or green firms, and granted by small, leveraged, lowly performing, illiquid, or less profitable banks, and by banks in closer proximity to the firm.

B. Variance equation: Dependent Variables

Using the estimated residual between the actual ratings and fitted ratings predicted by observables from the mean equation in Model (1), we construct our three variables to

[21] Because we are interested in how the depth and direction of private information varies over relationship time with firm and bank characteristics including firm or bank interacted with calendar time fixed effects *à la* Khwaja and Mian (2008) will result in excess saturation (as relationship and calendar time are inevitably correlated). In addition, including such excessively saturated effects set then also removes single bank firms from the sample (e.g., Degryse, De Jonghe, Jakovljevic, Mulier and Schepens (2019), De Jonghe, Dewachter, Mulier, Ongena and Schepens (2020), Greenstone, Mas and Nguyen (2020)). To maintain consistency across specifications (as we do not include these fixed effects in the variance equation either) we do not include such fixed effects in the mean equation, but doing so does not alter estimates of the main coefficients of interest much. This should not come as a surprise given that Cerqueiro, Degryse and Ongena (2013) for example points out that the estimates of the coefficients in the variance equation are often surprisingly unaffected by changes in the set of variables included in the mean equation, because the maximum likelihood estimators for the parameters in the mean and variance equations are, from a theoretical perspective in expectation, uncorrelated (see Harvey (1976)). This finding also pertains to the bank and firm variables which, in any case, are rather standard when explaining credit ratings (e.g., Altman (1968)). Altering this set also leaves the variance equation estimates mostly unaffected.

capture the content of banks' private information: *Depth*, and *Better* and *Worse Private Information. Depth* is calculated as the natural log of the squared residuals from the mean equation estimated in Model (1). *Better Private Information* is the (absolute value of the) estimated residual from the mean equation estimated in Model (1) when it is less than 0 and equals 0 otherwise. *Worse Private Information* is the estimated residual when it is more than 0 and equals 0 otherwise. While *Depth* tells the overall magnitude of private information, *Better Private Information* and *Worse Private Information* capture the degree of the (un)favorable nature of the information for each direction.

C. Private Information at Inception

So far, i.e., in the mean equation, we have shown that banks assess firms initially differently depending on observable firm and bank characteristics, as well as the distance to the bank. We next relate the three private information content measures at the beginning of the lending relationship to known factors, thus abstracting from any influences due to learning. We regress the three measures we constructed on several factors that capture firms' and banks' characteristics and the geographic distance between them. The dependent variable in Table 1 Model (2) is *Depth*, and in Models (3) and (4), the "direction equations", the dependent variables are *Better* or *Worse Private Information* respectively. With the residuals directional, i.e., greater *Better* or *Worse* are both higher values, positive estimated coefficients imply that the specific factor adds to the private information, i.e., betters it, relative to that estimated using observable variables, in Model (3) and conversely in Model (4) worsens it.

Most estimates of the firm, bank, and distance coefficients for the *Depth* and *Direction* equations, Models (2) to (4), are intuitive. Take for example firm size. In Model (2) the estimated coefficient equals 0.083***,[22] which implies that a one standard deviation increase in *Ln(Firm assets)* increases *Depth* by 0.19 (= 0.083 * 2.396) or around 8 percent, in Model (3) the estimate implies that size "lowers" *Better Private Information* by -0.036 (= -0.015 * 2.396) or by 7 percent. And in Model (4) the estimate implies that it also "increases" *Worse Private Information* by 0.093 (= 0.039 * 2.396) or 14 percent. In other words, banks initially attain more (depth in their) private information on larger firms and the banks' extra private information makes them lower their ratings when dealing with the "better" larger firms (recall that a lower rating is "better", i.e., maps into a lower probability of default), and increase their ratings when dealing with the "worse" larger firms. This is an interesting result suggesting that the banks studied here (which, recall, are among the larger ones in the US) initially amass more private information on larger firms, and then act upon it by adjusting rating direction, while they assign more "cookie cutter" (and hence based on observables more explainable) ratings for smaller firms (as in, e.g., Cole, Goldberg and White (2004)). However, as we will see below, relatively there is more subsequent "learning" for the smaller firms than for the larger firms.

Banks are also guided more by private information when dealing with profitable, leveraged, public, or other-than-green firms, for which it may be easier to do. And especially, large, leveraged, low-performing, illiquid, or unprofitable banks display more private information in their ratings; it may be more tempting or important for them to do so.

The estimates on *Distance bank HQ to firm* in Models (2) to (4) equal -0014, -0.006***, and -0.000, respectively, implying that distance makes for lower private information in positive ratings (i.e., ratings lower), with no statistically significant effects on depth and

---

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

negative ratings. These estimates imply that an increase in (log) distance from zero to the median (i.e., 6.5 or 657 miles) reduces better ratings by 3.9 percentage points (pp) (= 0.006 * 6.5), or about 12 percent of the mean residual (= 0.039 / 0.335). Hence, physical distance decreases in economically significant manners initial private information, leading to lower favorable ratings. This finding likely reflects a combination of the greater unfamiliarity of banks with such borrowers and the higher costs of collecting information.[23]

D. 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 2, Models (1) to (6), we provide our full regression results, which include the usual bank and firm controls, as well as the distance between the bank and the firm. 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 bank-firm relationship (0.25 << 3 years; 3 << 5 years; and > 5 years). The latter specification allows for the effects of a relationship to vary by period of length.

The estimates on the simple *Length bank-firm relationship* (Models (1) to (3)) equal 4.937***, 6.950***, and -5.550***, 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 0.20 (= 4.937 * 0.004), which is around 8 percent of its standard deviation (= 0.20 / 2.519), increases the likelihood of a better rating by 0.03 pp (= 6.950 * 0.004), which is around 6

[23] In a final validity section, we show that our three private information measures are meaningfully associated with salient loan terms at initiation and during the life of the loan.

percent of the standard deviation of the residual variable (= 0.03 / 0.494), and reduces the likelihood of a worse rating by 0.02 pp (= -5.550 * 0.004), which is around 3 percent of the standard deviation of the residual variable (= 0.02 / 0.679). Hence, as the relationship lengthens, banks give more weight in their internal ratings to private information and tend to rate firms better.

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

Note that the estimates on *Distance bank HQ to firm* in Table 2, both Models (1) to (3) and Models (4) to (6), now equal 0.009*, 0.003**, and 0.001 respectively, which vary from those in Table 1 (-0014, -0.006***, and -0.000, respectively). This suggests that after the initial lending stage, having a relationship significantly influences how distance affects depth and better and worse ratings, to the point that adverse effects are mitigated or even overcome. In other words, as relationships lengthen, banks start to deviate for distant

[24] The estimated coefficients on the dummies capture the deviation from the bases, which is the impact on the outcome variable when the length of the bank-firm relationship is below or equal to 0.25 years. Notice that the time period potentially spent in each period bin is somewhat different, i.e., 2.75, 2, and 7 years (the maximum is 12 years), which affects the proportion of observations in each bin, which equals 3, 34, 20, and 43 percent, respectively (see Appendix A.2), and the precision of the estimates.
[25] The high share of relationships longer than 5 years (43 percent) may explain why the negative estimated coefficient on that bin mirrors the overall negative coefficient estimate on relationship length in Model (3).
[26] When we split the length of the bank-firm relationship variable more finely into six dummies capturing lengths of one, two, three, four, five, or six or more years, we confirm that the effects for depth and better private information are somewhat larger for the longer relationship lengths, whereas for worse private information, the peak impact of a relationship occurs in the fifth year.

borrowers more from the hard information only they used in their initial assessment and rate these borrowers differentially.

E. Private Information Formation Across Banks and Firms

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

In Table 3 we report the regression results. Specifically, we add to the regression a variable chosen respectively among five firm characteristics (size, leverage, publicly traded versus private status, green and brown industry, with the latter not being complements), three bank characteristics (bank size, capital, liquidity ratio), the distance between the bank and the firm, and a dummy for the COVID period (2020:Q1 - 2021:Q1). We include every time the variable itself, the length of relationship and the interaction of the variable with relationship length. For each of the results, only three estimated coefficients are reported for each variable of interest, that is, the coefficients for the specific bank, firm characteristic or time period, the relationship length, and their interaction. The regressions do include the usual bank, firm and loan controls, but these are not reported. The sample size is the same for all regressions, some 3 million, so any variations in results do not reflect sample choices. Of most interest are the estimated coefficients for the interactions. These terms reflect how various firm and bank characteristics differently affect the formation of private information over time and thus show light on the processes of bank learning. Several findings emerge from the results.

In terms of firm characteristics, the variables most common statistically significant across the three information measures are asset size and leverage. The size result is that as the relationship goes on, banks reduce their *Depth* and have a smaller *Worse* private assessment for firms that are larger. The leverage effect is that over time banks increase the depth of their assessment and raise their positive private view of firms that are more leveraged. Overall, these results suggest that banks are more incentivized to learn about firms and adjust their private assessment when the firm is smaller and more leveraged (and thus riskier).

In term of bank characteristics, the results show that relationship length adds relatively less to *Depth*, worsens *Better* assessments, and adds more to *Worse* assessments for larger, better capitalized and more liquid banks. Note that these effects are, as expected, asymmetrical for firms with *Better* versus with *Worse* assessments, but larger in absolute size for *Better* assessments, making overall for less favorable ratings. This suggests that such 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.

In terms of distance, we find that a longer relationship leads to greater *Depth*, a higher *Better* assessment, and a lower *Worse* assessment for firms that are further away. Since the direct effects of distance are to lower a *Better* and raise a *Worse* assessment, a longer relationship thus offsets these effects to a degree. Finally, during COVID times, the length of the relationship is meaningfully less important for determining the *Depth* of private information, as indicated by the large and highly significant negative coefficient for the dummy. As the size of the coefficient is comparable in absolute size with that for length of the relationship, there appears to have been no learning as to *Depth* during the full COVID period. Directionally, it appears that for firms with a *Better* assessment, ratings did not suffer, but those with a *Worse* assessment got better ratings. This suggests that during the COVID period, the combination of difficulty in meeting with the firm in person with ample general fiscal and monetary support led banks to maintain their rating for firms with a *Better* assessment but upgraded it for the other firms (potentially, also in light of the more ample support, forbearing), even though they had less, or no, interactions and *Depth* declined.

The overall magnitudes of the effects of the combinations of relationship length and the key bank, firm and distance, accounting for their interactions, are displayed in Figure 2, Panels A-C. To benchmark the economic relevance of the impact, the red arrows on the figures display 10 percent of the standard deviation of the respective outcome variable.

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

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

F. Validity: Banks' private information and loan terms

Finally, to validate our measures, we investigate how banks' loan terms vary with our three measures of bank private information.

In Table 4, we regress four key loan terms, i.e., the Loan interest rate spread, the Ln(Loan maturity), the Ln(Loan amount), and d(Collateralized), on our three information measures. All the regressions include the usual firm, bank, and loan characteristics as controls, as well as bank and industry fixed effects,[27] and a constant. Note that we use here 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. The overall sample size is somewhat smaller for the spread regression (as Y-14 does not report the interest rate on undrawn credit) but otherwise is very similar to the full sample, so results do not reflect the sample choices.[28]

The estimates show that greater depth increases the interest rate spread and lowers the loan maturity, while it decreases the likelihood of collateralization. The former likely reflects that uncertainty about the borrower's quality leads to a higher risk premium being charged and a higher demand for collateral. Effects of depth are economically the largest for the spread. The greater depth in information coming with less collateralization could reflect a greater dispersion in lending, whereby in extremis, one bank gives a very large, collateralized loan, rather than multiple banks giving medium sized loans, making overall for less collateralization (as measured with a dummy). But its effect is economically very small. We also see some evidence, but not statistically significant, of rationing in terms of the amount and length of the loan.

Directionally, the measures of private information relate more strongly to loan terms. A more positive assessment comes with a lower spread, longer maturity, larger amount, and less collateralization. And a more negative assessment increases both the spread and

[27] In Appendix Table A3 we show that results vary somewhat in coefficients' size but are directionally similar when we comprehensively saturate the regressions with bank and firm*year:quarter fixed effects.
[28] 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.

likelihood of collateralization, and lowers the loan maturity and amount, all as expected. The effects are economically most meaningful for the spread charged, which decreases (increases) by 9.2 (11.4) percent of its median for a one standard deviation increase in positive (negative) assessment. Next important are the effects for collateralization, with a decline of 6.9 percent and an increase of 5.1 percent upon a similarly defined change in positive and negative assessment respectively. The effects for the other dependent variables, loan maturity and amount, are economically much smaller, 1.5 percent or less equivalently expressed. Overall, these results confirm that our private information measures are meaningful representations of the banks' lending procedures.

IV. 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 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. Second, we show how the length of relationship impacts these dimensions and how impacts vary with bank, firm, and loan characteristics, including distance between bank and firm. In this way, we gain additional insights as to the process of learning through relationships. Finally, we validate our three private information dimensions by relating them to the terms of loans granted.

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 effects of the length of relationships peak at about five years.

Importantly, effects are particularly salient for smaller and leveraged firms, smaller, leveraged, and illiquid banks, at longer firm-bank distances, and during non-COVID times. They are often strongly nonlinear and economically meaningful. This suggests that some banks have specific business models that make them more likely to invest resources to overcome the information asymmetries related to lending to such firms. Specifically, larger, highly capitalized, and highly liquid banks accumulate much less positive information on their borrowers over time, whereas smaller banks, 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 distance and relationship length 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 relationship differ by bank and firm characteristics. Such analysis has not been conducted before.

References

Agarwal, Sumit and Robert Hauswald. 2010. "Distance and Private Information in Lending." *Review of Financial Studies*, 23(7), 2757-88.

Altman, Edward I. 1968. "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *Journal of Finance*, 23(4), 589-610.

Baele, Lieven; Valerie De Bruyckere; Olivier De Jonghe and Rudi Vander Vennet. 2014. "Do Stock Markets Discipline Us Bank Holding Companies: Just Monitoring, or Also Influencing?" *North American Journal of Economics and Finance*, 29(7), 124-45.

Bannier, Christina E.; Patrick Behr and Andre Güttler. 2010. "Rating Opaque Borrowers: Why Are Unsolicited Ratings Lower?" *Review of Finance*, 14(2), 263-94.

Barone, Guglielmo; Roberto Felici and Marcello Pagnini. 2011. "Switching Costs in Local Credit Markets." *International Journal of Industrial Organization*, 29(6), 694-704.

Berger, Allen N.; Christa H. S. Bouwman; Lars Norden; Raluca A. Roman; Gregory F. Udell and Teng Wang. 2021. *Is a Friend in Need a Friend Indeed? How Relationship Borrowers Fare During the Covid-19 Crisis*. Bloomington IN: Indiana University Kelley School of Business, Research Paper, 1.

Berger, Allen N. and Gregory F. Udell. 2002. "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure." *Economic Journal*, 112(477), 32-53.

Beyhaghi, Mehdi; Cooper Howes and Gregory Weitzner. 2023. *The Information Advantage of Banks: Evidence from Their Private Credit Assessments*. Washington DC: Board of Governors of the Federal Reserve System, Mimeo.

Bonfim, Diana; Gil Nogueira and Steven Ongena. 2021. "“Sorry, We're Closed" Bank Branch Closures, Loan Pricing and Information Asymmetries." *Review of Finance*, 25(4), 1211-59.

Boot, Arnoud W. A. 2000. "Relationship Banking: What Do We Know?" *Journal of Financial Intermediation*, 9(1), 3-25.

Cao, Jin; Emilia Garcia-Appendini and Cédric Huylebroek. 2024. *Banking on Deposit Relationships: Implications for Hold-up Problems in the Loan Market*. Oslo NO: Norges Bank, Working Paper, 4.

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

Carling, Kenneth; Tor Jacobson; Jesper Lindé and Kasper Roszbach. 2007. "Corporate Credit Risk Modeling and the Macroeconomy." *Journal of Banking and Finance*, 31(3), 845-68.

Cavallo, Eduardo; Andrew Powell and Roberto Rigobon. 2013. "Do Credit Rating Agencies Add Value? Evidence from the Sovereign Rating Business." *International Journal of Finance and Economics*, 18(3), 240-65.

Cerqueiro, Geraldo; Hans Degryse and Steven Ongena. 2011. "Rules Versus Discretion in Loan Rate Setting." *Journal of Financial Intermediation*, 20(4), 503-29.

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

Cole, Rebel A.; Lawrence G. Goldberg and Lawrence J. White. 2004. "Cookie-Cutter Versus Character: The Micro Structure of Small Business Lending by Large and Small Banks." *Journal of Financial and Quantitative Analysis*, 39(2), 227-52.

Correa, Ricardo; Kuan-Hui Lee; Horacio Sapriza and Gustavo A. Suarez. 2014. "Sovereign Credit Risk, Banks' Government Support, and Bank Stock Returns around the World." *Journal of Money, Credit and Banking*, 46(s1), 93-121.

De Jonghe, Olivier; Hans Dewachter; Klaas Mulier; Steven Ongena and Glenn Schepens. 2020. "Some Borrowers Are More Equal Than Others: Bank Funding Shocks and Credit Reallocation." *Review of Finance*, 24(1), 1-43.

Degryse, Hans; Olivier De Jonghe; Sanja Jakovljevic; Klaas Mulier and Glenn Schepens. 2019. "Identifying Credit Supply Shocks with Bank-Firm Data: Methods and Applications." *Journal of Financial Intermediation*, Forthcoming.

Degryse, Hans; Vasso Ioannidou and Steven Ongena. 2015. "Bank-Firm Relationships: A Review of the Implications for Firms and Banks in Normal and Crisis Times," W. Tsutomu, I. Uesugi and A. Ono, *The Economics of Interfirm Networks.* Berlin: Springer, 177-89.

Degryse, Hans; Moshe Kim and Steven Ongena. 2009. *Microeconometrics of Banking: Methods, Applications and Results*. Oxford University Press.

Degryse, Hans; Luc Laeven and Steven Ongena. 2009. "The Impact of Organizational Structure and Lending Technology on Banking Competition." *Review of Finance*, 13, 225-59.

Degryse, Hans; Paola Morales-Acevedo and Steven Ongena. 2019. "Competition in Banking," A. N. Berger, P. Molyneux and J. O. S. Wilson, *Oxford Handbook of Banking.* Oxford: Oxford University Press, 776-813.

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

____. 2005. "Distance, Lending Relationships, and Competition." *Journal of Finance*, 60(1), 231-66.

____. 2007. "The Impact of Competition on Bank Orientation." *Journal of Financial Intermediation*, 16, 399-424.

Di, Gong; Steven Ongena; Shusen Qi and Yanxin Yu. 2024. *"Time for a Change of Scenery": Loan Conditions When Firms Switch Bank Branches*. London UK: Centre for Economic Policy Research, Discussion Paper, 18985.

Duqi, Andi; Angelo Tomaselli and Giuseppe Torluccio. 2018. "Is Relationship Lending Still a Mixed Blessing? A Review of Advantages and Disadvantages for Lenders and Borrowers." *Journal of Economic Surveys*, 32(5), 1446-82.

Ederington, Louis H. and Jeremy C. Goh. 1998. "Bond Rating Agencies and Stock Analysts: Who Knows What When?" *Journal of Financial and Quantitative Analysis*, 33(4), 569-85.

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

Faria-e-Castro, Miguel; Pascal Paul and Juan M. Sánchez. 2021. *Evergreening*. St. Louis MO: Federal Reserve Bank of St. Louis, Working Paper, 2021-012.

Flannery, Mark J.; Simon H. Kwan and Mahendrarajah Nimalendran. 2013. "The 2007–2009 Financial Crisis and Bank Opaqueness." *Journal of Financial Intermediation*, 22(1), 55-84.

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

Granja, João; Christian Leuz and Raghuram G. Rajan. 2022. "Going the Extra Mile: Distant Lending and Credit Cycles." *Journal of Finance*, 77(2), 1259-324.

Greenstone, Michael; Alexandre Mas and Hoai-Luu Nguyen. 2020. "Do Credit Market Shocks Affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and "Normal" Economic Times." *American Economic Journal: Economic Policy*, 12(1), 200-25.

Gupta, Manish and Steven Ongena. 2022. *The Impact of the Sba Funding Programs on the Distance and Pricing of Loans to Small Businesses*. Nottingham UK: Nottingham University Business School, Mimeo.

Hand, John, R. M.; Robert W. Holthausen and Richard W. Leftwich. 1992. "The Effect of Bond Rating Agency Announcements on Bond and Stock Prices." *Journal of Finance*, 47(2), 733-52.

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

Hauswald, Robert and Robert Marquez. 2006. "Competition and Strategic Information Acquisition in Credit Markets." *Review of Financial Studies*, 19(3), 967-1000.

____. 2003. "Information Technology and Financial Services Competition." *Review of Financial Studies*, 16(3), 921-48.

Herpfer, Christoph; Cornelius Schmidt and Aksel Mjøs. 2022. "The Causal Impact of Distance on Bank Lending." *Management Science*, 69(2), 723-40.

Hirtle, Beverly. 2006. "Stock Market Reaction to Financial Statement Certification by Bank Holding Company Ceos." *Journal of Money, Credit and Banking*, 38(5), 1263-91.

Iannotta, Giuliano. 2011. "Market Discipline in the Banking Industry: Evidence from Spread Dispersion." *European Journal of Finance*, 17(2), 111-31.

____. 2006. "Testing for Opaqueness in the European Banking Industry: Evidence from Bond Credit Ratings." *Journal of Financial Services Research*, 30(3), 287-309.

Iannotta, Giuliano and Marco Navone. 2012. "The Cross-Section of Mutual Fund Fee Dispersion." *Journal of Banking and Finance*, 36(3), 846-56.

Ioannidou, Vasso P. and Steven Ongena. 2010. ""Time for a Change": Loan Conditions and Bank Behavior When Firms Switch Banks." *Journal of Finance*, 65, 1847-78.

Jones, Jeffrey S.; Wayne Y. Lee and Timothy J. Yeager. 2012. "Opaque Banks, Price Discovery, and Financial Instability." *Journal of Financial Intermediation*, 21(3), 383-408.

Khwaja, Asim Ijaz and Atif Mian. 2008. "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market." *American Economic Review*, 98(4), 1413-42.

King, Michael R.; Steven Ongena and Nikola Tarashev. 2020. "Bank Standalone Credit Ratings." *International Journal of Central Banking*, 16(3), 101-44.

Kliger, Doron and Oded Sarig. 2000. "The Information Value of Bond Ratings." *Journal of Finance*, 55(6), 2879-902.

Kysucky, Vlado and Lars Norden. 2016. "The Benefits of Relationship Lending in a Cross-Country Context: A Meta-Analysis." *Management Science*, 62(1), 90-110.

Liberti, Jose Maria and Mitchell A. Petersen. 2018. "Information: Hard and Soft." *Review of Corporate Finance Studies*, 8(1), 1-41.

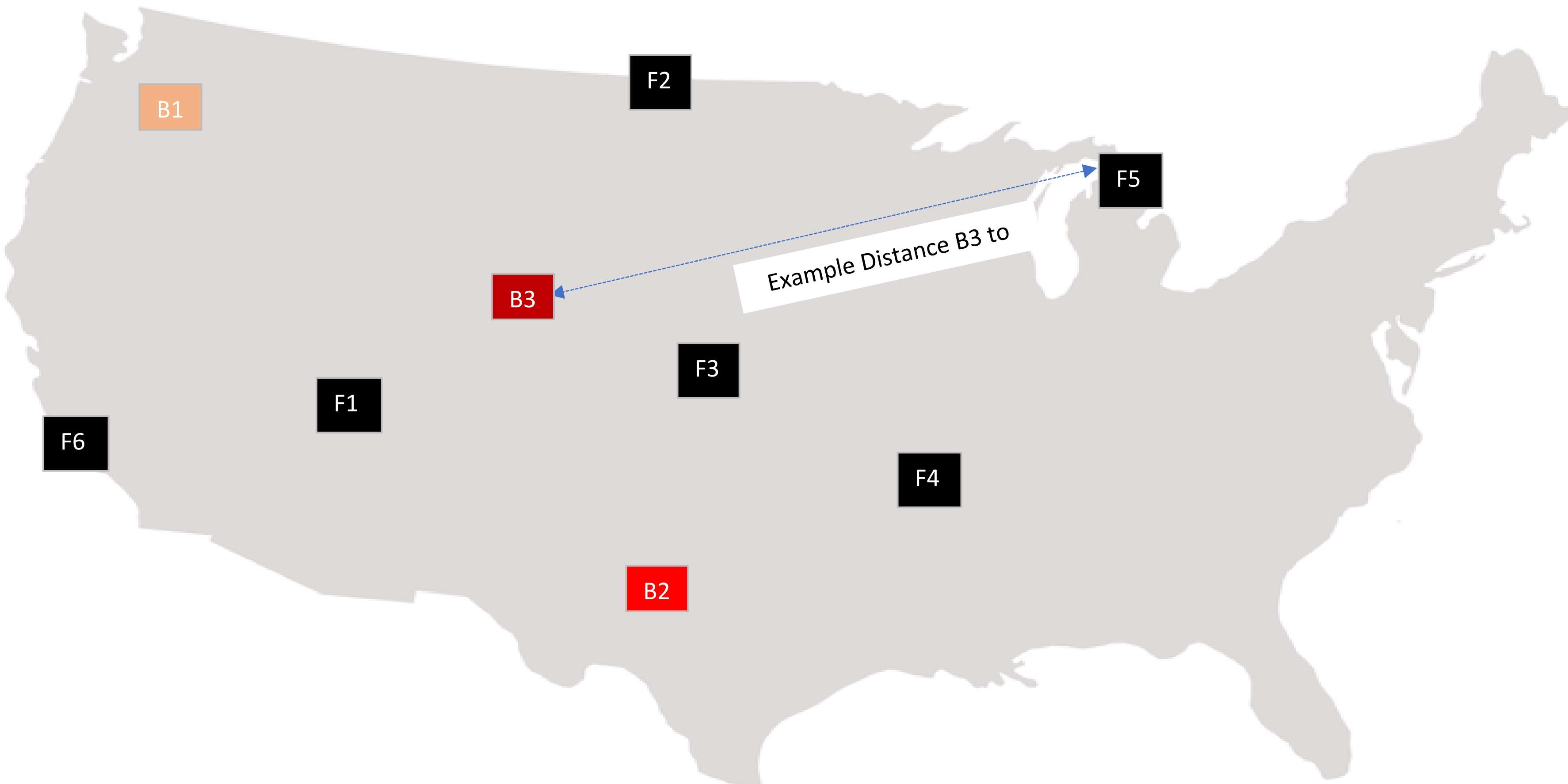
Livingston, Miles; Andy Naranjo and Lei Zhou. 2007. "Asset Opaqueness and Split Bond Ratings." *Financial Management*, 36(3), 49-62.

Morgan, Donald. 2002. "Rating Banks: Risk and Uncertainty in an Opaque Industry." *American Economic Review*, 92, 874-88.

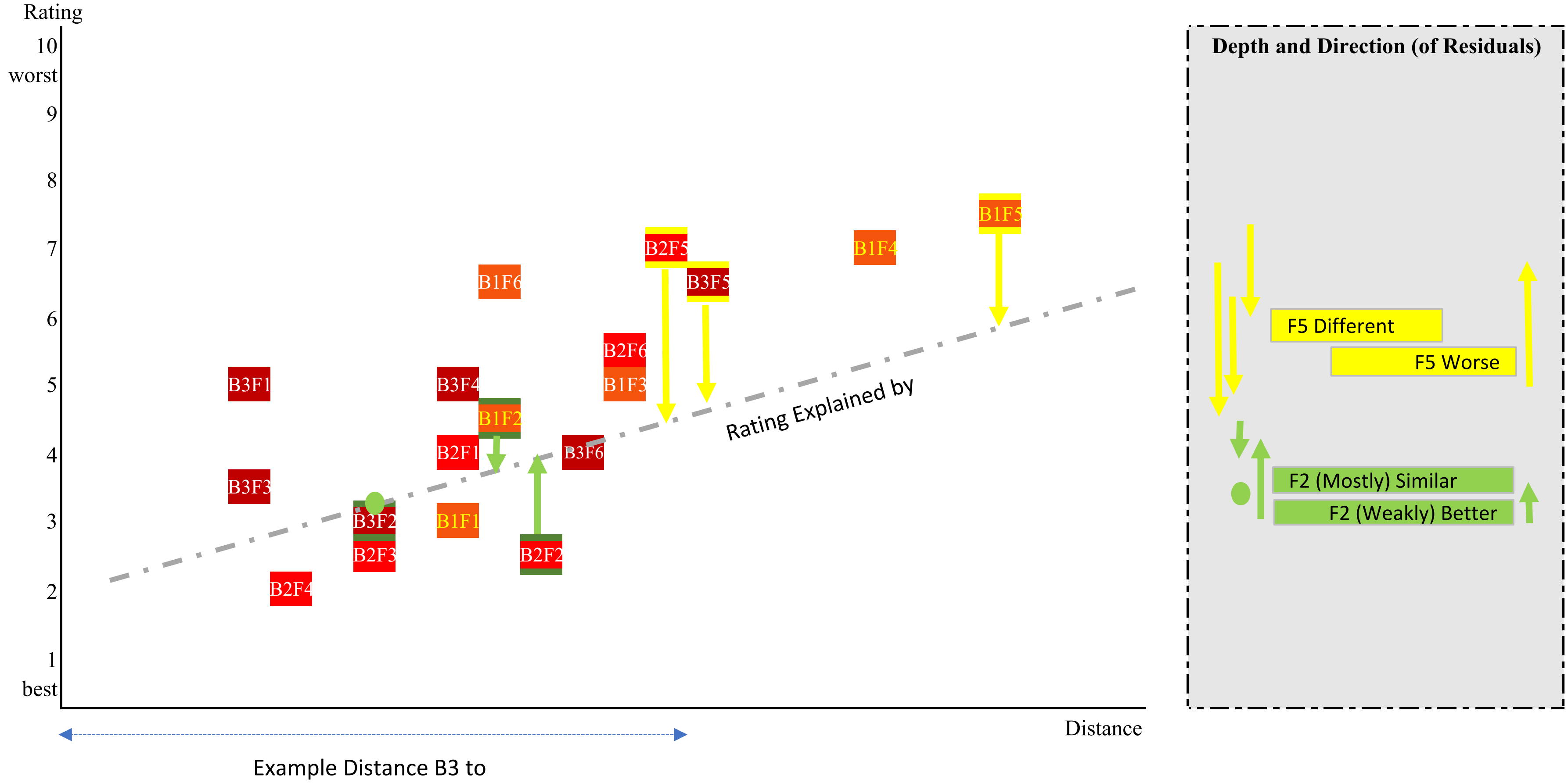
Nakamura, Leonard I. and Kasper Roszbach. 2018. "Credit Ratings, Private Information, and Bank Monitoring Ability." *Journal of Financial Intermediation*, 36, 58-73.
Ongena, Steven and David C. Smith. 2000. "Bank Relationships: A Review," P. Harker and S. A. Zenios, *The Performance of Financial Institutions.* London: Cambridge University Press, 221-58.
Petersen, Mitchell A. and Raghuram G. Rajan. 1994. "The Benefits of Lending Relationships: Evidence from Small Business Data." *Journal of Finance*, 49, 3-37.
____. 2002. "Does Distance Still Matter? The Information Revolution in Small Business Lending." *Journal of Finance*, 57(6), 2533-70.
Plosser, Matthew C. and João A. C. Santos. 2018. "Banks' Incentives and Inconsistent Risk Models." *Review of Financial Studies*, 31(6), 2080-112.
Rajan, Raghuram G. 1992. "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." *Journal of Finance*, 47(4), 1367-400.
Sharpe, Steven A. 1990. "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships." *Journal of Finance*, 45(4), 1069-87.
Sironi, Andrea. 2003. "Testing for Market Discipline in the European Banking Industry: Evidence from Subordinated Debt Issues." *Journal of Money, Credit and Banking*, 35(3), 443-72.
Stein, Ingrid. 2015. "The Price Impact of Lending Relationships." *German Economic Review*, 16(3), 367-89.
Uchida, Hirofumi; Gregory F. Udell and Wako Watanabe. 2008. "Bank Size and Lending Relationships in Japan." *Journal of the Japanese and International Economies*, 22(2), 242-67.
von Thadden, Ernst-Ludwig. 2004. "Asymmetric Information, Bank Lending, and Implicit Contracts: The Winner's Curse." *Finance Research Letters*, 1(1), 11-23.
Xu, Yuqian; Anthony Saunders; Binqing Xiao and Xindan Li. 2020. "Bank Relationship Loss: The Moderating Effect of Information Opacity." *Journal of Banking and Finance*, 118, 105872.

Figure 1. Depth and Direction of Private Information

Panel A. Geographical Distribution of Banks and Firms



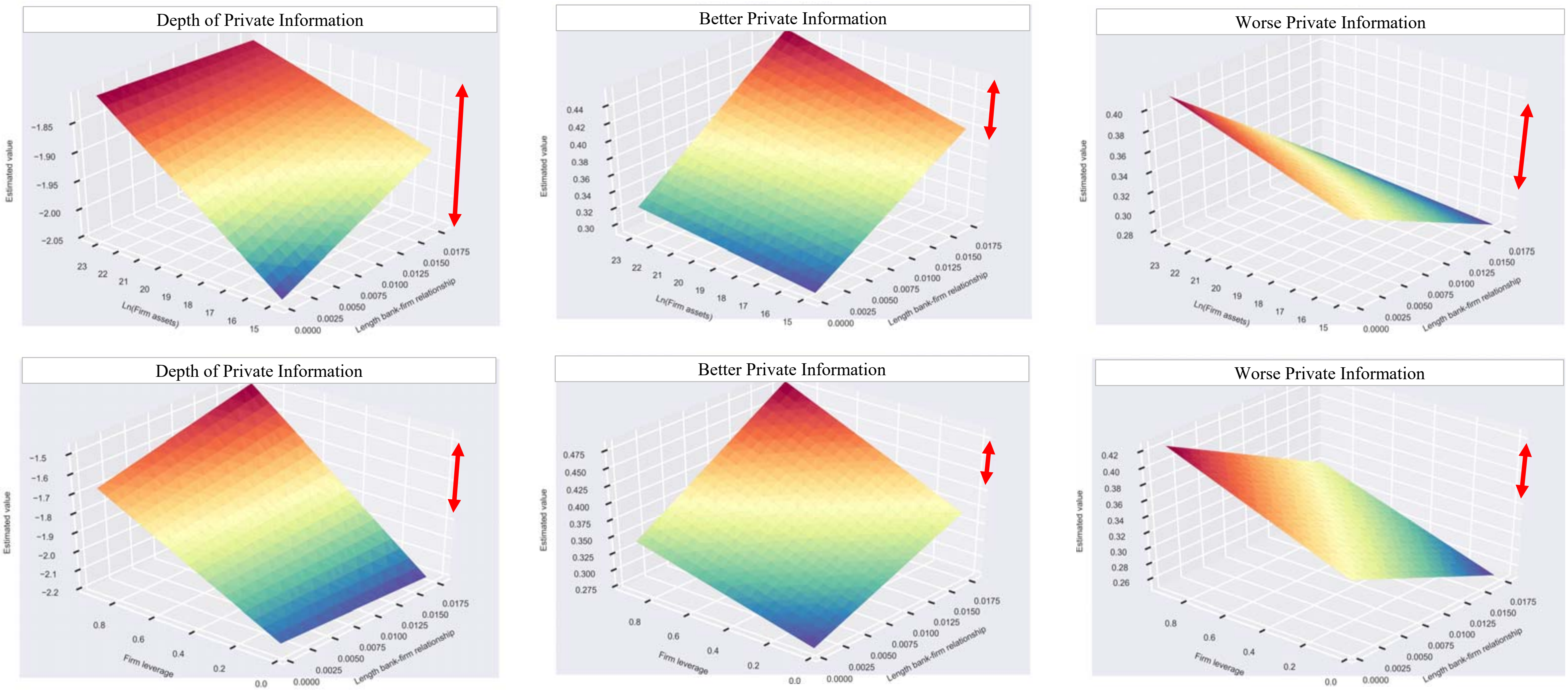
Panel B. Standardized Bank Rating of Firm, and Depth of Private Information and Direction of Private Information



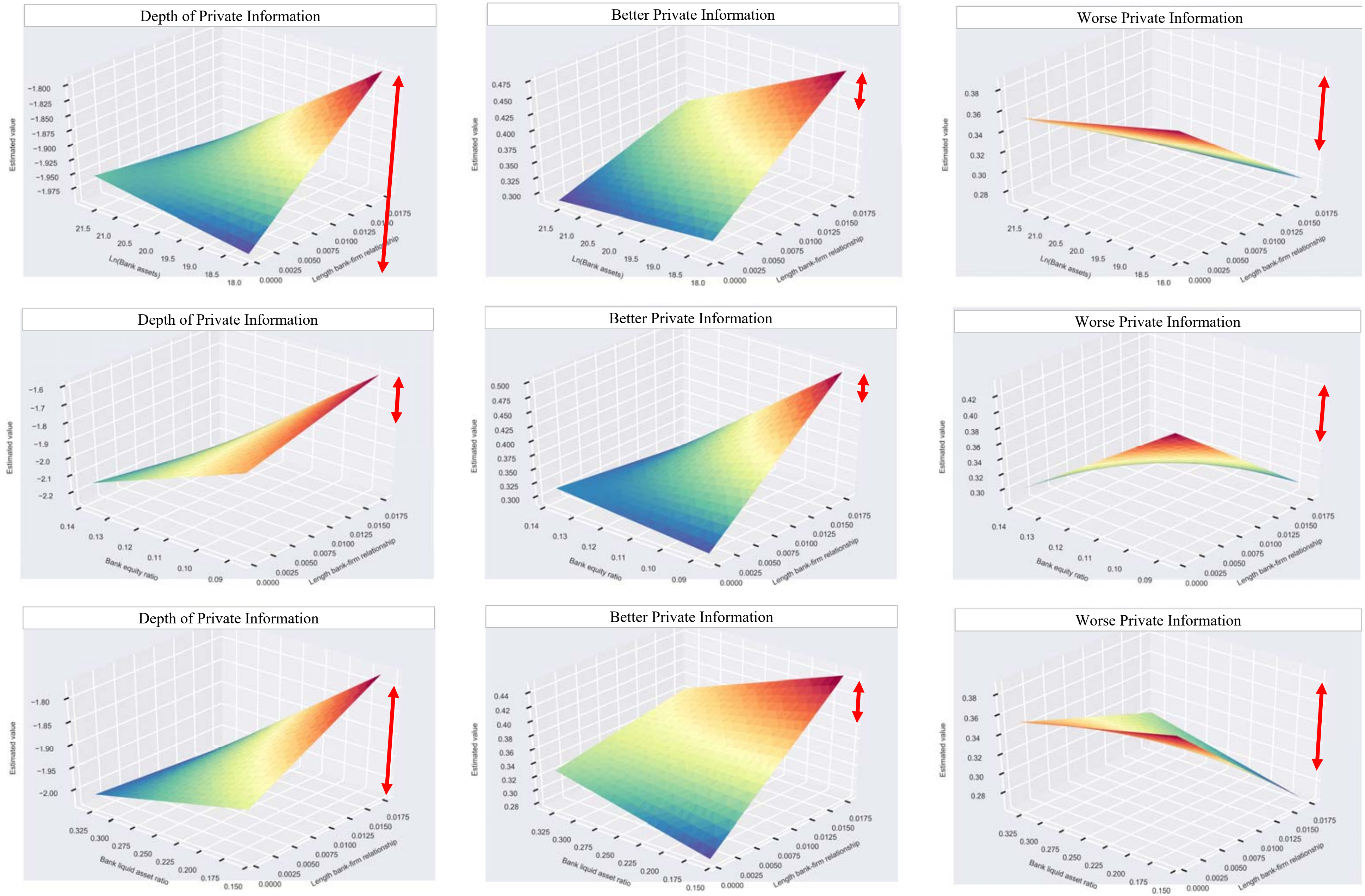
Panel A depicts example geographical distances between three banks (in various shades of red) and six firms (in black). Panel B plots on the horizontal axis the distance between banks and firms, and on the vertical axis example ratings. For the Distance from Bank 3 to Firm 5 and example arrow is placed 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.

Figure 2. Depth and Direction of Private Information: Length of the Bank-Firm Relationship and Its Interactions with Firm, Bank Characteristics

Firm

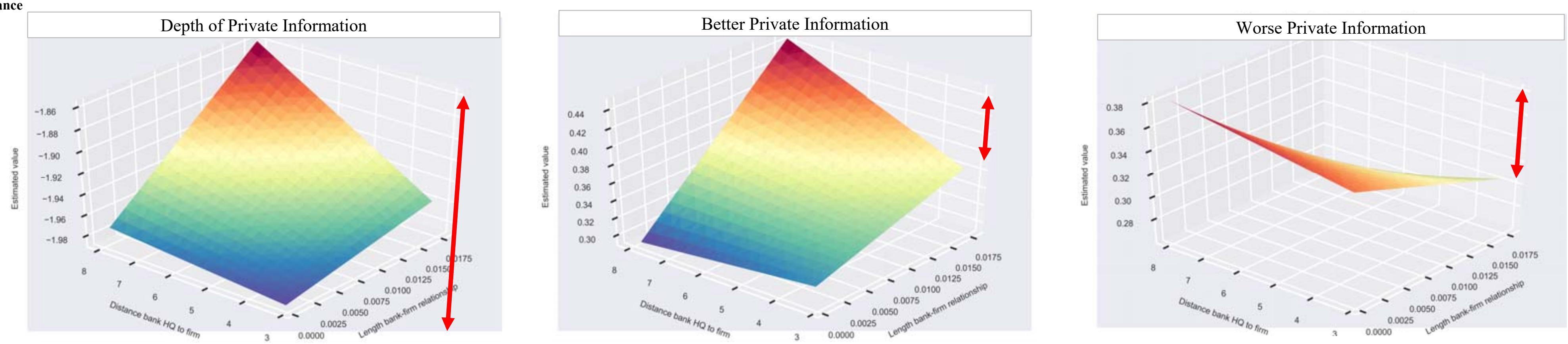


Bank



Depth of Private Information
Better Private Information
Worse Private Information
Estimated value
Ln(Bank assets)
Length bank-firm relationship
Depth of Private Information
Better Private Information
Worse Private Information
Estimated value
Bank equity ratio
Length bank-firm relationship
Depth of Private Information
Better Private Information
Worse Private Information
Estimated value
Bank liquid asset ratio
Length bank-firm relationship

Distance



The figure plots the variance and positive and negative residual equation estimates from Table 4 for length of the bank-firm relationship and its interactions with various firm and bank characteristics such as firm and bank ln(Total assets). All other variables are set at their median. The red arrows indicate ten percent of the standard deviation of the respective outcome variable.

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

Model	(1)	(2)	(3)	(4)
Sample	***All bank ratings of firms***	***Length bank-firm relationship < 0.25 Years***		
Definition	**Standardized bank rating of firm (1 = best, 10 = worst)**	**Ln(Residual squared)**	**Residual if Residual < 0**	**Residual if Residual > 0**
Dependent Variable Name	*Bank Rating of Firm*	*Depth of Private Information*	*Better Private Information*	*Worse Private Information*
Independent variables				
Firm Variables				
Ln(Firm assets)	-0.121***	0.083***	-0.015***	0.039***
	(-36.84)	(12.83)	(-8.64)	(19.05)
Firm ROA	-2.065***	1.314***	-0.333***	0.550***
	(-85.77)	(18.72)	(-29.81)	(22.44)
Firm leverage	0.469***	0.593***	0.171***	-0.016
	(41.84)	(13.50)	(19.21)	(-1.40)
Public	-0.169***	0.302***	0.025*	0.024
	(-8.48)	(4.68)	(1.86)	(1.16)
Green	-0.052**	-0.151**	-0.038***	-0.003
	(-2.44)	(-2.28)	(-3.30)	(-0.20)
Brown	-0.004	0.060	0.076***	-0.054**
	(-0.10)	(0.45)	(2.95)	(-2.27)
Bank Variables				
Ln(Bank assets)	0.097***	0.083***	0.001	-0.001
	(31.56)	(6.82)	(0.49)	(-0.20)
Bank equity ratio	3.378***	-9.242***	1.082***	-4.195***
	(16.38)	(-9.96)	(6.33)	(-20.18)
Bank NPL ratio	-4.830***	-2.333***	1.722***	-0.940***
	(-29.70)	(-2.76)	(11.99)	(-4.28)
Bank liquid asset ratio	0.392***	-1.329***	0.420***	-0.544***
	(6.48)	(-5.87)	(8.03)	(-9.23)
Bank ROA	15.236***	-23.256***	1.570	-3.199
	(14.88)	(-2.92)	(1.22)	(-1.48)
Static Bank-Firm Variable				
Distance bank HQ to firm	0.009***	-0.014	-0.006***	-0.000
	(3.78)	(-1.60)	(-3.74)	(-0.16)
Observations	2,996,502	64,061	64,061	64,061
Adjusted R-squared	0.168	0.019	0.035	0.047

The table reports estimates from ordinary least squares regressions. The sample in Model (1) includes all bank ratings given to 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 Standardized 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%.

Table 2. Bank-firm Relationship Length and Banks' Private Information

Dependent Variable / Independent variables	(1) *Depth of Private Information*	(2) *Better Private Information*	(3) *Worse Private Information*	(4) *Depth of Private Information*	(5) *Better Private Information*	(6) *Worse Private Information*
Dynamic Bank-Firm Variable						
Length bank-firm relationship	4.937***	6.950***	-5.550***			
	(4.40)	(19.68)	(-18.54)			
Length bank-firm relationship (0.25 << 3 years)				0.156***	0.008***	0.042***
				(13.10)	(3.71)	(14.42)
Length bank-firm relationship (3 << 5 years)				0.232***	0.037***	0.047***
				-14.66	(12.10)	(11.94)
Length bank-firm relationship (> 5 years)				0.213***	0.085***	-0.009**
				(13.11)	(24.97)	(-2.25)
Firm Variables						
Ln(Firm assets)	0.024***	0.005**	0.006***	0.023***	0.004**	0.006***
	(5.01)	(2.34)	(3.68)	(4.86)	(2.00)	(3.81)
Firm ROA	0.101**	-0.129***	-0.144***	0.103**	-0.132***	-0.140***
	(2.42)	(-15.24)	(-7.84)	(2.49)	(-15.52)	(-7.64)
Firm leverage	0.647***	0.094***	0.097***	0.646***	0.097***	0.094***
	(24.25)	(15.83)	(13.89)	(24.19)	(16.03)	(13.33)
Public	0.345***	0.011	0.017	0.348***	0.014	0.015
	(9.15)	(1.04)	(1.44)	(9.23)	(1.27)	(1.27)
Green	-0.133***	-0.017*	-0.013	-0.131***	-0.015	-0.015
	(-3.54)	(-1.82)	(-1.14)	(-3.47)	(-1.53)	(-1.18)
Brown	0.191***	0.039	0.034	0.190***	0.037	0.035
	(2.99)	(1.64)	(1.41)	(2.97)	(1.58)	(1.46)
Bank Variables						
Ln(Bank assets)	-0.009	-0.015***	-0.008***	-0.010	-0.016***	-0.008***
	(-1.39)	(-8.98)	(-4.28)	(-1.57)	(-9.70)	(-4.05)
Bank equity ratio	-9.117***	-1.219***	-1.814***	-9.352***	-1.503***	-1.597***
	(-21.63)	(-9.99)	(-16.93)	(-22.17)	(-11.96)	(-14.70)
Bank NPL ratio	0.318	0.984***	-0.075	0.367	0.944***	0.053
	(0.81)	(10.71)	(-0.68)	(0.93)	(10.37)	(0.48)
Bank liquid asset ratio	-0.827***	0.051	-0.073**	-0.882***	-0.001	-0.037
	(-6.59)	(1.51)	(-2.04)	(-7.08)	(-0.04)	(-1.06)
Bank ROA	-8.763***	1.102*	-0.543	-8.850***	0.467	0.148
	(-3.90)	(1.84)	(-0.88)	(-3.92)	(0.78)	(0.24)
Static Bank-Firm Variable						
Distance bank HQ to firm	0.009*	0.003**	0.001	0.009*	0.003**	0.001
	(1.85)	(2.13)	(0.59)	(1.89)	(2.20)	(0.69)
Observations	2,994,729	2,994,729	2,994,729	2,994,729	2,994,729	2,994,729
Adjusted R-squared	0.012	0.016	0.007	0.012	0.013	0.006

The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to firms. The number of observations equals 2,994,729. The dependent variables are: in Panel A the Depth of Private Information, which is the natural log of the squared residuals; and, in Panels B and C the Better and Worse Private Information which is equal to the absolute value of the residual, and 0 otherwise, if the residual is larger or smaller, respectively, than zero. In all cases "the residual" is the estimated residual from the 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 3. Depth and Better or Worse Private Information: Estimates for Length of the Bank-Firm Relationship and Its Interactions with a Variable of Interest

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample										
Variable of interest	***Ln(Firm assets)***	***Firm leverage***	***Public***	***Green***	***Brown***	***Ln(Bank assets)***	***Bank equity ratio***	***Bank liquid asset ratio***	***Distance bank HQ to firm***	***COVID period***
Panel A										
Dependent Variable	*Depth of Private Information*									
Independent variables										
Variable of interest	0.031***	0.547***	0.340***	-0.080	0.160**	0.015*	-6.619***	-0.455***	0.003	-0.020
	(6.03)	(17.13)	(7.27)	(-1.63)	(1.98)	(1.85)	(-12.44)	(-3.49)	(0.48)	(-1.05)
Length bank-firm relationship	23.965**	-2.910	4.857***	5.190***	4.884***	84.274***	45.745***	19.148***	-1.518	5.117***
	(2.26)	(-1.37)	(4.49)	(4.61)	(4.32)	(5.08)	(7.08)	(4.34)	(-0.39)	(4.52)
Variable of interest * Length bank-firm relationship	-1.075*	16.584***	0.737	-8.629	5.280	-3.973***	-380.903***	-61.965***	1.047*	-4.873***
	(-1.71)	(4.80)	(0.14)	(-1.05)	(0.58)	(-4.76)	(-6.52)	(-3.07)	(1.65)	(-2.68)
Firm, Bank, and Bank-Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
Panel B										
Dependent Variable	*Better Private Information*									
Independent variables										
Variable of interest	0.003	0.075***	0.003	-0.008	0.021	-0.005*	0.557***	0.328***	-0.005***	-0.039***
	(1.62)	(9.30)	(0.24)	(-0.64)	(0.75)	(-2.44)	(3.49)	(9.09)	(-3.13)	(-10.16)
Length bank-firm relationship	3.785	5.466***	6.829***	6.996***	6.920***	39.637***	35.982***	17.558***	-1.063	6.946***
	(1.05)	(8.93)	(19.97)	(19.67)	(19.68)	(7.57)	(16.50)	(13.33)	(-1.08)	(19.56)
Variable of interest * Length bank-firm relationship	0.179	3.137***	1.116	-1.565	3.067	-1.637***	-270.984***	-46.255***	1.300***	0.553
	(0.84)	(2.84)	(0.72)	(-0.77)	(0.93)	(-6.22)	(-14.85)	(-8.10)	(7.36)	(1.42)
Firm, Bank, and Bank-Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.016	0.016	0.016	0.016	0.016	0.017	0.020	0.018	0.017	0.016
Panel C										
Dependent Variable	*Worse Private Information*									
Independent variables										
Variable of interest	0.008***	0.103***	0.022*	0.002	0.035	-0.010***	-2.732***	-0.231***	0.006***	0.052***
	(4.37)	(11.81)	(1.66)	(0.19)	(1.34)	(-4.46)	(-20.52)	(-6.78)	(3.84)	(10.84)
Length bank-firm relationship	0.113	-5.114***	-5.466***	-5.474***	-5.549***	-13.206***	-20.550***	-11.597***	-0.044	-5.472***
	(0.04)	(-9.09)	(-20.98)	(-18.81)	(-18.63)	(-2.95)	(-15.01)	(-9.90)	(-0.04)	(-18.11)
Variable of interest * Length bank-firm relationship	-0.320*	-0.921	-0.771	-2.586	-0.067	0.383*	140.011***	26.366***	-0.893***	-2.786***
	(-1.69)	(-1.00)	(-0.55)	(-1.17)	(-0.02)	(1.68)	(11.57)	(4.77)	(-5.07)	(-6.27)
Firm, Bank, and Bank-Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.007	0.007	0.007	0.007	0.007	0.007	0.008	0.007	0.007	0.007

The table reports estimates from ordinary least squares regressions. The sample includes all bank ratings given to firms. The number of observations equals 2,994,729. The dependent variables are: in Panel A the Depth of Private Information, which is the natural log of the squared residuals; and, in Panels B and C the Better and Worse Private Information which is equal to the absolute value of the residual, and 0 otherwise, if the residual is larger or smaller, respectively, than zero. In all cases "the residual" is the estimated residual from the 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 4. Impact of Private Information on Loan Terms

Dependent Variable	(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*
Independent variables		*Loan Interest Rate Spread*				*Ln(Loan Maturity)*		
Depth of Private Information	0.025***			*2.69%*	-0.001			*-0.15%*
	(4.53)				(-1.07)			
Better Private Information		-0.422***		*-8.84%*		0.036***		*1.08%*
		(-24.08)				(2.79)		
Worse Private Information			0.395***	*11.35%*			-0.038***	*-1.57%*
			(26.12)				(-6.69)	
Firm, Bank and Loan Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	2,276,100	2,276,100	2,276,100		2,991,730	2,991,730	2,991,730	
Adjusted R-squared	0.109	0.117	0.125		0.261	0.262	0.262	

Dependent Variable	(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*
Independent variables		*Ln(Loan Amount)*				*d(Collateralized)*		
Depth of Private Information	-0.001			-	-0.002***			*-1.54%*
	(-0.62)				(-3.73)			
Better Private Information		0.040***		*1.51%*		-0.045***		*-6.74%*
		(2.73)				(-6.15)		
Worse Private Information			-0.025***	*-1.30%*			0.025***	*5.13%*
			(-3.44)				(10.84)	
Firm, Bank and Loan Controls, and Bank and Industry Fixed Effects	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	2,994,729	2,994,729	2,994,729		2,994,729	2,994,729	2,994,729	
Adjusted R-squared	0.459	0.459	0.459		0.265	0.269	0.268	

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, firm variables, bank variables, the distance between the bank headquarters and the firm, and bank and industry fixed effects. The sample includes corporate loans 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 one standard deviation of Depth of Private Information or information on the loan term, in percent scaled by the median of this loan term. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix

Appendix 1. Illustration of Depth and Direction of Private Information

For illustrative purposes consider the standardized bank rating given by a loan officer to the firm to be based on two informational components, i.e., a hard component and a soft component. For our illustration we consider the hard component to contain only public information, while the soft component will be based solely on private information.[1]

Consider the distinguishing characteristic of both components to be their non-random ("constant") versus random nature. Hence, the standardized internal bank rating $\tilde{R}$ is defined as:

$$Internal\ bank\ rating = \tilde{R} = hH + s\tilde{S} \qquad (1)$$

with H is the non-random part of the rating based on hard information, while $\tilde{S} \sim N(\mu_S, \sigma_S^2)$ is the random component based on soft information, with h and s the relative weights of their contribution to the standardized rating ($0 \leq h, s \ \leq 1,\ \ h + s = 1$), which for now we will take as equal across banks (see the discussion below).[2] Notice that in this simple illustration and for ease of interpretation (and in contrast to our empirical variable definition) a higher $\tilde{R}$ will imply a better rating for the firm (and potentially better credit conditions). We are also ignoring in this illustration the fact that in reality ratings are discrete grades and constrained in their support.

If the loan officer is averse to the randomness of the soft information component, then she may reduce the rating she actually grants accordingly to a rating we then actually can observe, for example like:

---

[1] For a treatise on the differences between hard and soft information, see, e.g., Liberti and Petersen (2018). Hard information can also be private if only the bank can collect additional statistics on the firm`s operations. By its very nature it is harder to consider soft information to be public.

[2] Though stylized, this composition of bank ratings to consist of quantitative ("hard") and qualitative ("soft") elements is adequately realistic (e.g., Machauer and Weber (1998); Nakamura and Roszbach (2018); Berg, Puri and Rocholl (2019)).

$$Observed\ bank\ rating = \ \hat{R} = E(\tilde{R}) - \frac{\gamma}{2}Var(\tilde{R}) = hH + s\mu_S - \frac{\gamma}{2}s^2\sigma_S^2 \qquad (2)$$

With γ the officer`s aversion to using soft information that is uncertain so as not to give the firm too high a rating and too good credit conditions. Notice that one can benignly surmise that the loan officer can only "take back" less than half the soft component she herself "started with", i.e., $\frac{s\mu_S}{2} \geq \frac{\gamma}{2}s^2\sigma_S^2$, or $\mu_S \geq \gamma s\sigma_S^2$, and that the loan officer cannot affect the hard component.[3]

The dispersion or incongruity of ratings (for a representative bank), or depth in private information, in this illustration is then the square of the difference between the hard information component of the internal bank rating and the observed bank rating, while unfavorability, which is the worsening direction of this private information, is the (absolute) difference.

$$Incongruity = \left\{hH - \left[hH + \ s\mu_S - \frac{\gamma}{2}s^2\sigma_S^2\right]\right\}^2$$

$$= s^2\mu_S^2 + \frac{\gamma^2}{4}s^4\sigma_S^4 - \gamma s^3\mu_S\sigma_S^2 = s^2\left\{\mu_S^2 + \frac{\gamma^2}{4}s^2\sigma_S^4 - \gamma s\mu_S\sigma_S^2\right\} \qquad (3)$$

$$= \ s^2\left\{\mu_S^2 + \frac{\gamma^2}{4}s^2\sigma_S^4 + \mu_S[\mu_S - \gamma s\sigma_S^2]\right\}$$

$$Unfavourability = -\ s\mu_S + \frac{\gamma}{2}s^2\sigma_S^2 = s\left\{-\mu_S + \frac{\gamma}{2}s\sigma_S^2\right\} \qquad (4)$$

In sum, depth in private information increases in the weight *h* attributed to the expected value and the variance in the soft information component, σ_S^2, and in the aversion of the loan officer to its randomness, γ; while the worse direction of the private information increases in the weight attributed to and the variation in the soft information component,

[3] We ignore for the sake of simplicity the manipulation possibilities in the way hard information can be entered by the loan officer (Berg, Puri and Rocholl (2019)).

and the aversion of the loan officer to its randomness, but it decreases in the expected value of the soft information component.

Notice that for different banks all these elements may take different values, and that our estimated models try to assess how bank characteristics as a consequence determine both incongruity and unfavorability accordingly.

Finally, let us reflect on how the distance (d) between the bank and the firm and the length of their relationship (l) may affect the collection of soft information. Distance may increase the variance of the soft information component (e.g., Hauswald and Marquez (2003)), in which case both depth and negative direction of private information strengthen. On the other hand, the length of the relationship may increase the mean value of the soft information component the loan officer collects (as firms deemed of good character may well be less likely to be poached by other banks),[4] and decrease its variance (as the collected soft information becomes more precise). Hence negative direction decreases, while the depth of private information will increase if the former effect dominates (i.e., if loan officers do not offset more than half their own increase in the mean value of the soft component corresponding the increase in relationship length).

[4] In Sharpe (1990), Rajan (1992), von Thadden (2004), and Hauswald and Marquez (2006), among others, private information about repayment is being collected by the "inside" bank and especially low-quality firms are poached by the uninformed "outside" banks.

Appendix 1. References

Berg, Tobias; Manju Puri and Jörg Rocholl. 2019. "Loan Officer Incentives, Internal Rating Models, and Default Rates." *Review of Finance*, 24(3), 529-78.
Hauswald, Robert and Robert Marquez. 2006. "Competition and Strategic Information Acquisition in Credit Markets." *Review of Financial Studies*, 19(3), 967-1000.
____. 2003. "Information Technology and Financial Services Competition." *Review of Financial Studies*, 16(3), 921-48.
Liberti, Jose Maria and Mitchell A. Petersen. 2018. "Information: Hard and Soft." *Review of Corporate Finance Studies*, 8(1), 1-41.
Machauer, A. and M. Weber. 1998. "Bank Behavior Based on Internal Credit Ratings of Borrowers." *Journal of Banking and Finance*, 22, 1355-83.
Nakamura, Leonard I. and Kasper Roszbach. 2018. "Credit Ratings, Private Information, and Bank Monitoring Ability." *Journal of Financial Intermediation*, 36, 58-73.
Rajan, Raghuram G. 1992. "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." *Journal of Finance*, 47(4), 1367-400.
Sharpe, Steven A. 1990. "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships." *Journal of Finance*, 45(4), 1069-87.
von Thadden, Ernst-Ludwig. 2004. "Asymmetric Information, Bank Lending, and Implicit Contracts: The Winner's Curse." *Finance Research Letters*, 1(1), 11-23.

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 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	Standardized 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
Loan Outcome	Loan Interest Rate Spread	%	Interest rate spread over the rate of a constant maturity Treasury bond with a similar maturity.	FR Y-14Q Schedule H.1
	Ln(Loan Maturity)	ln years	The log of one plus the number of years from the date of origination to the date of maturity.	FR Y-14Q Schedule H.1
	Ln(Loan Amount)	ln mln $	The log of one plus the size of the loan in $ million.	FR Y-14Q Schedule H.1
	d(Collateralized)	0/1	= 1 if the loan is collateralized, = 0 otherwise.	FR Y-14Q Schedule H.1
Level of Variables	**INDEPENDENT VARIABLES**			
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
	Public	0/1	= 1 if the firm is publicly listed, = 0 otherwise	Compustat
	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
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
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
	Distance bank branch to firm	ln miles	The log of the distance between the bank`s closest branch 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	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	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	FR Y-14Q Schedule H.1
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 reported in the Y-14Q by bank holding companies between Sep 30, 2012, and March 31, 2021. The number of observations equals 2,994,729.

Variables		Unit / Split	Mean	Median	Minimum	Maximum	Standard Deviation
Equation	**DEPENDENT VARIABLES**						
Bank Rating	Standardized bank rating of firm	1 (best) -10 (worst)	4.922	5	1	10	1.06
	Residual	-	0	-0.067	-5.244	6.855	0.966
Residual	*Depth of Private Information*	-	-1.905	-1.398	-29.237	3.85	2.519
	Better Private Information	-	0.338	0.067	0	5.244	0.494
	Worse Private Information	-	0.338	0	0	6.855	0.679
Loan Outcome	Loan Interest Rate Spread	%	2.184	2.315	-3.69	8.1	1.994
	Ln(Loan Maturity)	ln years	1.452	1.611	-5.9	3.23	0.954
	Ln(Loan Amount)	ln $	15.227	14.908	13.816	24.189	1.282
	d(Collateralized)	0/1	0.881	1	0	1	0.324
Level of Variables	**INDEPENDENT VARIABLES**						
Firm							
	Firm assets	tn $	2.298	0.03	0.001	98.598	11.038
	Ln(Firm assets)	ln $	17.781	17.209	13.816	25.314	2.396
	Firm ROA	-	0.077	0.051	-0.239	0.84	0.132
	Firm leverage	-	0.414	0.38	0	1	0.266
	Public	0/1	0.093	0	0	1	0.291
	Green	0/1	0.035	0	0	1	0.185
	Brown	0/1	0.016	0	0	1	0.124
Bank							
	Bank assets	bn $	1,036.269	391.67	20.454	2,808.396	1,014.865
	Ln(Bank assets)	ln thd $	20.049	19.786	16.834	21.756	1.295
	Bank equity ratio	-	0.115	0.112	0.052	0.207	0.015
	Bank NPL ratio	-	0.02	0.015	0	0.079	0.015
	Bank liquid asset ratio	-	0.249	0.246	0.083	0.787	0.06
	Bank ROA	-	0.002	0.003	-0.033	0.016	0.002
Main Bank-Firm							
	Distance bank HQ to firm	ln miles	6.249	6.489	0.047	8.124	1.389
	Distance bank branch to firm	ln miles	1.973	1.566	0	7.733	1.717
	Length bank-firm relationship	years	0.006	0.004	0	0.12	0.006
	Length bank-firm relationship (0.25 << 3 years)	0/1	0.341	0	0	1	0.474
	Length bank-firm relationship (3 << 5 years)	0/1	0.203	0	0	1	0.403
	Length bank-firm relationship (> 5 years)	0/1	0.434	0	0	1	0.496
Loan							
	d(Loan is not a syndicate)	0/1	0.939	1	0	1	0.24
	d(Loan is a term loan)	0/1	0.307	0	0	1	0.461
	d(Loan is a revolver)	0/1	0.463	0	0	1	0.499
	d(Loan is floating rate)	0/1	0.566	1	0	1	0.496
	d(Loan is mixed rate)	0/1	0.185	0	0	1	0.388
	d(Loan purpose is miscellaneous)	0/1	0.241	0	0	1	0.427
	d(Loan purpose is M&A or capital expenditure)	0/1	0.1	0	0	1	0.299
	d(Loan purpose is general)	0/1	0.502	1	0	1	0.5
	d(Loan purpose is real estate)	0/1	0.156	0	0	1	0.363

Appendix Table A3. Impact of Private Information on Loan Terms

Dependent Variable	(1)	(2) *Loan Interest Rate Spread*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*	(1)	(2) *Ln(Loan Maturity)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*
Independent variables								
Depth of Private Information	-0.004			-	0.004			-
	(-0.84)				(2.29)			
Better Private Information		-0.196***		*-4.11%*		0.018		-
		(-5.42)				(1.55)		
Worse Private Information			0.123***	*3.53%*			0.007	-
			(3.83)				(0.92)	
*Firm, Bank and Loan Controls, and Bank and Firm*year_quarter Fixed Effects*	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	1,125,531	1,125,531	1,125,531		1,651,427	1,651,427	1,651,427	
Adjusted R-squared	0.519	0.519	0.519		0.403	0.403	0.403	

Dependent Variable	(1)	(2) *Ln(Loan Amount)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*	(1)	(2) *d(Collateralized)*	(3)	*Impact of one standard deviation increase on dependent variable as percent of its median*
Independent variables								
Depth of Private Information	-0.002			-	0.003***			*2.31%*
	(-0.59)				-2.82			
Better Private Information		0.076***		*2.88%*		-0.016***		*-2.40%*
		(4.25)				(-2.57)		
Worse Private Information			-0.058***	*-3.01%*			0.037***	*7.59%*
			(-4.63)				(7.19)	
*Firm, Bank and Loan Controls, and Bank and Firm*Year:Quarter Fixed Effects*	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	1,653,779	1,653,779	1,653,779		1,653,779	1,653,779	1,653,779	
Adjusted R-squared	0.505	0.505	0.505		0.55	0.55	0.551	

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, firm variables, bank variables, the distance between the bank headquarters and the firm, and bank and industry fixed effects. The sample includes corporate loans 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 one standard deviation of Depth of Private Information or information on the loan term, in percent scaled by the median of this loan term. ***, **, and * denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Swiss Finance Institute

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

swiss:finance:institute

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