Economic Inquiry

NUMERACY AND ON-THE-JOB PERFORMANCE: EVIDENCE FROM LOAN OFFICERS

MARTIN BROWN, KAROLIN KIRSCHENMANN and THOMAS SPYCHER*

We examine how the numeracy level of employees influences their on-the-job performance. Based on an administrative dataset of a retail bank we relate the performance of loan officers in a standardized math test to the accuracy of their credit assessments of small business borrowers. We find that loan officers with a high level of numeracy are more accurate in assessing the credit risk of borrowers. The effect is most pronounced during the precrisis credit boom period when it is arguably more difficult to pick out risky borrowers. (*JEL* G21, J24)

I. INTRODUCTION

Employers in a broad range of industries place significant weight on the numerical skills of job applicants when hiring new employees. Numerical skills are also associated with better labor market outcomes among workers (Joensen and Nielsen 2009; Koedel and Tyhurst 2012). These two observations suggest that employees with strong numerical skills are more productive or make better on-the-job decisions. Numerical skills themselves may foster better decision making as employees are better able to draw meaning from numerical information (Peters et al. 2006). Alternatively, numeracy may be correlated with broader cognitive or social skills—which improve decision speed or quality (Burks et al. 2009). While it is plausible that high levels of numeracy are associated with better job-related performance, there is almost no empirical evidence to support this conjecture.

This paper empirically examines the relation between employee numeracy and on-the-job performance. Our analysis focuses on loan officers in a retail bank. A key task of loan officers is the screening of loan applicants, that is, the assessment of the borrowers' creditworthiness.[1] We study how the numeracy of loan officers relates to the accuracy of their credit assessments of small business borrowers: Are loan officers with high numeracy better able to identify those borrowers who ex post turn out to be risky? With a unique dataset provided by a retail bank we are able to match loan officers' performance in a standardized numeracy test with data on all loan applications that they process. The loan-level

*We thank the Editor (Steven Ongena), an anonymous referee, Bo Becker, Tobias Berg, Kristian Blickle, Adrian Bruhin, Geraldo Cerqueiro, Claudia Custodio, Bob DeYoung, Vasso Ioannidou, Tor Jacobson, Markku Kaustia, Matti Keloharju, Samuli Knüpfer, Rafael Lalive, Michael Lechner, Daniel Metzger, Steven Ongena, Matias Ossandon Busch and Christian Zehnder, seminar participants at the Aalto University School of Business, the Center for European Economic Research, KfW Development Bank, TU Munich and the University of St. Gallen, as well as participants at the 10th Swiss Winter Conference on Financial Intermediation and the Bank of Finland Institute for Economies in Transition Workshop on Banking and Institutions for helpful comments. We are especially grateful to the management and employees of the bank that provided us with the data. Spycher gratefully acknowledges financial support from the Swiss National Science Foundation (project no. 100018_169433).

Brown: Professor Doctor Swiss Institute of Banking and Finance, University of St. Gallen, 9000, St. Gallen, Switzerland., Phone 41-71-224-70-62, E-mail martin.brown@unisg.ch

Kirschenmann: Doctor, Research Unit International Finance and Financial Management, Center for European Economic Research (ZEW), 68161, Mannheim, Germany., Phone 49-621-1235-351, E-mail karolinkirschenmann@gmail.com; karolin.kirschenmann@zew.de

Spycher: Doctor Swiss Institute of Banking and Finance, University of St. Gallen, 9000, St. Gallen, Switzerland., Phone 41-71-224-70-42, E-mail thomas.spycher@unisg.ch

1. Apart from client acquisition and advising customers, the U.S. Bureau of Labor Statistics mentions the gathering, verification and analysis of applicants' information and the loan approval decision as typical tasks of a loan officer (see http://www.bls.gov/ooh/business-and-financial/loan-officers.htm).

ABBREVIATIONS

BIS: Bank for International Settlements
CEO: Chief Executive Officer
FE: Fixed Effects
GDP: Gross Domestic Product
RON: Romanian Leu

Economic Inquiry
(ISSN 0095-2583)
Vol. 58, No. 2, April 2020, 998–1022
doi:10.1111/ecin.12873
Online Early publication January 9, 2020

data contain information on the requested loan terms, the borrower, the initial credit assessment by the loan officer, the approval decision and, for the approved loans, the granted loan terms as well as subsequent loan performance. The sample period 2007–2010 further allows for the analysis of a heterogeneous influence of numeracy during a credit boom and bust phase.

Small business lending provides an ideal framework to study the relationship between numeracy and on-the-job performance. The production and processing of information is a core function of financial intermediaries (Diamond 1984). Two key features of small business lending allow us to study the importance of loan officer numeracy in carrying out this function. First, the lending methodology applied by most small business lenders leaves discretion to the individual loan officer in screening potential borrowers (Berger and Udell 1995). The screening process requires loan officers to collect, verify and assess both quantitative and qualitative information. Loan officers' skills can strongly influence the collection or processing of information. Hence, differences in skills across loan officers should translate into a difference in the quality of client screening. Second, loan officers make a large number of comparable lending decisions for which outcomes are quantitatively measurable. By comparison, for most other skilled professionals on-the-job performance is difficult to measure and hardly comparable across employees.

We face two identification challenges when studying the relation between loan officer numeracy and the accuracy of credit assessments: First, the assignment of loan applications to loan officers is hardly random—and is likely to be related to loan officers' numeracy levels. A profit maximizing bank should allocate the most skilled loan officers to those tasks where their skills can generate the highest profit.[2] The detailed loan-level data at hand help us to account for differences in borrower and application characteristics which may confound the relationship between loan officer numeracy and the accuracy of credit assessments. Second, other loan officer characteristics such as education, age, gender, or job experience might be correlated with both loan officers' numeracy level and their screening accuracy. Our estimates may therefore suffer from an omitted variable bias and represent a spurious relationship between numeracy and screening accuracy. Our administrative dataset includes information on education, age, gender and experience which allows us to control for these confounding loan officer characteristics.

Our results show that loan officers with higher numeracy make more accurate credit assessments. Accuracy is hereby measured by the discriminatory power of the ex ante risk scores assigned by loan officers: Those borrowers classified as risky ex ante are more likely to fall into payment arrears ex post than those borrowers classified as less risky. Subsample analyses suggest that numeracy is especially important for accuracy in the precrisis credit boom when information asymmetries seem strongest. Before the crisis, high numeracy loan officers are clearly better able to discriminate borrowers by their creditworthiness than low numeracy loan officers. This difference in accuracy between loan officers with high and low numerical skills decreases in the crisis period due to a considerable improvement in the accuracy of low numeracy loan officers.

Previous research has shown that numeracy is correlated with an array of cognitive and social skills which may prove essential in the screening of small and opaque borrowers. Individuals with higher numeracy seem less prone to framing effects (Peters et al. 2006), and seem better able to anticipate social behavior (Burks et al. 2009). Thus, loan officers with higher levels of numeracy can be expected to be more accurate in verifying and interpreting hard information as well as evaluating soft information. Individuals with higher numeracy have also been found to be more patient (Burks et al. 2009; Frederick 2005), which might imply that they are better able to take the longer-term future into account when assessing borrowers' credit risk. Our happenstance data does not allow us to disentangle the effect of pure numerical skills, that is, the ability to understand and work with numbers and to do logical reasoning, from correlated cognitive or social skills. However, our results highlight that a simple test which captures numerical skills and correlated personal traits can be used to identify employees with better on-the-job performance.

Our findings contribute to a broad literature in finance, economics and psychology that analyzes how numerical skills affect corporate and personal[3] decision making as well as labor

2. Fang, Kempf, and Trapp (2014) show that fund families allocate their most skilled managers to less efficient market segments. In less efficient markets skills have the highest reward and the allocation maximizes profits.

3. See Reyna et al. (2009) for an overview on health decisions.

market performance. Experimental research provides evidence that numeracy influences strategies used for decision making and the quality of the decisions taken. Individuals with higher numeracy have superior judgment abilities (Ghazal, Cokely, and Garcia-Retamero 2014) and are more likely to choose the normatively better option with a higher expected value (Pachur and Galesic 2013).

Empirical studies based on field data document that numeracy, cognitive skills and financial literacy are associated with better personal financial decisions. Investors with higher IQ are able to select mutual funds with lower fees (Grinblatt et al. 2015), are less prone to the disposition effect and are able to generate higher returns (Grinblatt, Keloharju, and Linnainmaa 2012). Individuals with lower financial literacy more frequently transact in high-cost manners, for example, they pay higher credit card fees or use more high-cost debt (Lusardi and Tufano 2015). Gerardi, Goette, and Meier (2013) document significantly higher mortgage default rates among individuals who are not able to perform basic mathematical calculations. In a sample of members of the United States military, Agarwal and Mazumder (2013) find that a higher math test score is associated with fewer personal finance mistakes related to credit card use and home equity loans compared to other skills tested in the Armed Forces Qualifying Test.

Labor economics provides evidence that employers value math skills in the hiring process (Koedel and Tyhurst 2012) and that more mathematical education results in better labor market outcomes (Joensen and Nielsen 2009).[4] These findings support the conjecture that employees with high numeracy are more productive and perform better on-the-job. However, to our knowledge, there is only one study connecting a concept related to numerical skills to job performance.[5] Burks et al. (2009) find that truck drivers with higher cognitive skills are more likely to avoid planning mistakes that could lead to performance failures such as arriving late for deliveries. Our study extends the literature by providing unique evidence for the effect of numeracy on on-the-job performance among skilled professionals.

Our findings also contribute to a strand in the empirical banking literature which studies the role of loan officers in bank internal decision making. Recent studies have analyzed the influence of internal organization (e.g., Brown et al. 2015; Hertzberg, Liberti, and Paravisini 2010; Liberti and Mian 2009; Qian, Strahan, and Yang 2015) and incentives (e.g., Agarwal and Ben-David 2018; Berg 2015; Cole, Kanz, and Klapper 2015). Other papers focus on loan officers' characteristics that might explain why certain loan officers perform better within a given organizational and incentive structure. Existing work looks at the influence of loan officers' gender (Beck, Behr, and Guettler 2013), experience (Andersson 2004; Bruns et al. 2008), education (Bruns et al. 2008) and traumatic experiences (Morales-Acevedo and Ongena 2019). We add to this literature by documenting an important role of loan officers' numerical skills for the quality of lending decisions.

Finally, we contribute to the recent literature which examines lending standards over the business cycle (e.g., Beck et al. 2018; Berger and Udell 2004; Dell'Ariccia, Igan, and Laeven 2012; Dell'Ariccia and Marquez 2006). In line with Becker, Bos, and Roszbach (2018), we provide evidence for a lower accuracy of internal risk ratings during the credit boom, pointing toward higher information asymmetries. We add to the literature by showing that loan officer skills are most important during this boom phase with strong information asymmetries.

The remainder of this paper is organized as follows. In Section II, we describe the institutional background. In Section II we describe our data, while we explain our methodology in Section IV. We present our results in Section V and conclude in Section VI.

II. INSTITUTIONAL BACKGROUND

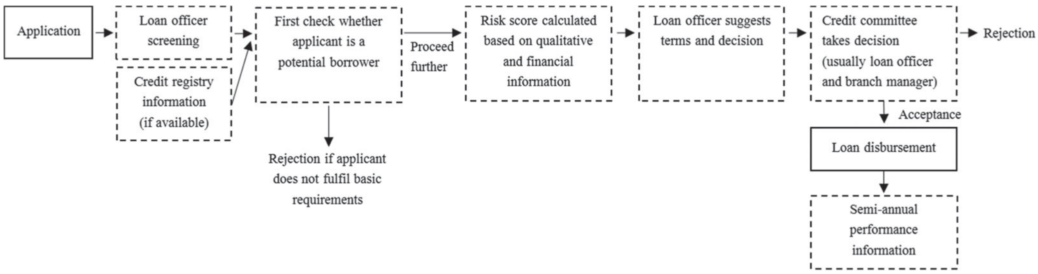
A. The Bank's Lending Process

The bank that provided us with the data is a country-wide retail bank in Romania. It is part of an international banking group and serves micro and small enterprises as well as households. The bank does not substantially differ in terms of business practices and loan

4. Joensen and Nielsen (2009) show that higher earnings are mainly the results of differences in career paths and not of differences in earnings of individuals following a comparable career path.

5. A recent literature analyses the importance of chief executive officer (CEO) traits and skills for performance. Custodio and Metzger (2014) show that CEOs' financial expertise is correlated with differences in firms' financial policies that benefit performance. Kaplan, Klebanov, and Sorensen (2012) study CEOs involved in private equity deals and document a positive correlation between their skills (performance in a general ability test and execution skills) and their performance. Further, a related strand of literature analyzes the impact of fund manager skills on fund performance (e.g., Chevalier and Ellison 1999; Li, Zhang, and Zhao 2011).

FIGURE 1
Lending Process



Notes: This summary of the bank's lending process is based on extensive interviews with loan officers and credit risk managers of the bank.

products from small U.S. or European commercial banks which specialize in relationship lending to small businesses. One potential difference to some commercial lenders is the incentive structure of the bank: The bank regularly agrees with branch managers and loan officers on performance goals. While the achievement of these goals may affect the career path of employees within the bank, goal achievement is *not* financially incentivized through performance pay.

Our analysis focuses on first-time loans to small businesses with amounts of up to 30,000 Euro. These "micro" loans make up the bulk of the bank's loan portfolio. The credit assessment and approval process for these loans follows a standardized process which is illustrated in Figure 1.[6]

In a first step, prospective borrowers fill in a paper-based application form and submit it to the closest bank branch. For first time borrowers, the application is filled out without loan officer involvement and should therefore not be influenced by loan officer skills. Clients state their requested amount, requested currency (Lei or Euro) and requested maturity and provide information on the loan purpose, other bank relationships as well as the ownership structure and the free cash flow / disposable income of the firm. Each loan application is then assigned to a loan officer within the branch where the borrower submitted the application. The allocation of an application to a loan officer is first and foremost based on loan officers' available capacity. That said, our data reveals that some loan officers do have an industry focus or tend to process predominantly applications for either small or large volumes.

In a second step, the allocated loan officer screens the application and assigns an initial risk score to the borrower. As we measure loan officers' screening performance by the accuracy of the initial risk score, we need to clarify which components of the initial risk score are potentially influenced by numeracy. In general, the initial risk score is based on quantitative and qualitative information: During an on-site visit, the loan officer verifies the quantitative information provided in the application such as accounting data that allow for the computation of disposable income or free cash flow and assesses collateral values. Further, the loan officer collects qualitative information by assessing the entrepreneur's character and overall managerial quality as well as the market outlook for the business. Concurrently, the bank's back office provides credit registry information on the borrower to the loan officer. It is important to note that many of the banks' first-time micro loan applicants have never had another bank loan before and henceforth no credit registry information exists. If information is available it becomes part of the credit risk assessment.[7] The bank has a policy that loans with very negative credit registry information (e.g., the days of arrears within the last 2 years are above a certain threshold) or with clearly poor financial information are rejected as early as possible in the screening process. For all other loan applications the loan officer enters the collected qualitative information and verified quantitative data into a standardized spread sheet

6. Our description of the lending process is based on extensive interviews with loan officers and credit risk managers of the bank.

7. Unfortunately, we do not have access to the credit registry information.

which then automatically calculates the risk score based on an underlying algorithm.[8] This process does not require any manual calculations.

Therefore, any differences in accuracy should originate from differences in the loan officers' input to the rating model rather than from their ability to simply calculate numbers. A first source of heterogeneity could stem from differences in the quality of the financial information verification. Peters et al. (2006) show that higher numeracy individuals are less prone to framing effects and are able to draw stronger and more precise affective meaning from numbers and comparisons using numbers. A second source of heterogeneity could stem from differences in the precision of the qualitative information collection. Again, framing effects and the skill to draw precise affective meaning (Peters et al. 2006) but also the higher likelihood to choose the normatively better option with a higher expected value (Pachur and Galesic 2013) may influence the precision of the market outlook analysis. And the assessment of the borrower's character and managerial quality arguably requires social skills. There is evidence that cognitive skills are useful for social interaction (Burks et al. 2009). Overall, we expect higher numeracy to improve the verification and interpretation of quantitative information as well as the precision of qualitative information.

In a third step, the loan officer suggests loan terms (volume, currency, maturity) and recommends the lending decision to the credit committee.[9] For the majority of loan applications in our sample there are two members in the credit committee: the branch manager and the loan officer. The credit committee evaluates the provided information, verifies the risk score, reviews the loan officer's suggestion and makes a final lending decision.

In case of a positive lending decision (67% of the applications) and if the client accepts the loan terms (95% of the offered loans), the loan is disbursed and the repayment performance reported semi-annually.

8. Generally, the risk score can take on values from 1 (lowest risk class) to 5 (highest risk class). However, the bank's policy is to reject first loan applications with an initial risk score exceeding 3. Accordingly, we only observe initial risk scores from 1 to 3 and treat firms with initial scores other than 1 as risky.

9. Interest rates are largely standardized for the loans in our sample (as is the usual practice with micro loans), that is,. that they are mainly determined by the size of the loan and are not fully risk-adjusted.

B. The Numeracy Test

To perform the credit assessment described above loan officers require diverse cognitive and social skills. We have an indicator of loan officers' numerical skills in the form of a score on a math test conducted in February 2010. All loan officers employed at that date were obliged to take the test at the same time at selected locations in the country. The test was announced on short notice so there was limited time for preparation. Passing the math test (there was an option to retake the test) was a requirement for the continuation of the employment relationship. The math test was prescribed by the international banking group to all its subsidiaries worldwide and thus can be considered as exogenous to the Romanian subsidiary—and its loan officers—which we study. The test measured basic numerical skills on the level of high school math covering percentage calculations, probability theory, logic and geometric understanding and equations.[10] Thus, the test is a comprehensive measure of numeracy comparable to tests discussed in Ginsburg, Manly, and Schmitt (2006).

C. The Economic Environment

Romania experienced a substantial lending boom over the period 2000–2007 during which the stock of credit relative to Gross Domestic Product (GDP) increased from 7% to 35%. Credit to firms and households grew in some years by more than 50%. Lending volumes slowed down significantly and economic growth turned negative in the last quarter of 2008. With the global financial crisis hitting Romania in 2009, the share of nonperforming loans in banks' portfolios rose sharply. These underlying economic conditions had a severe impact on the bank that we study. Figure A1 shows that its total assets, gross loans and total deposits decreased in 2009 while its nonperforming loan ratio increased sharply. After years of branch network expansion, several branches were also closed in 2010.

Our dataset covers both precrisis and crisis years so that we can analyze potential heterogeneities in the effect of numeracy on loan officers' decision quality over a boom and bust cycle.

10. Three example questions from the test are provided in Appendix A. The test was part of a series of tests such as a more advanced math test as well as an accounting test. The additional tests were taken at different dates and only completed by a subgroup of loan officers who took the first math test. Hence, we focus on the first math test as our measure of numerical skills.

Based on the macroeconomic and bank variables, we classify our sample into two subperiods. The precrisis period lasts up to the third quarter of 2008 with positive GDP and credit growth and very low nonperforming loan rates. We classify October 2008 to December 2010 (before the Bank started math training courses in 2011) as the crisis period over which Romania's GDP dropped significantly and nonperforming loan rates increased steadily.

III. DATA

We merge two bank-internal administrative datasets. The loan officer data covers all loan officers that passed the numerical test in February 2010 and contains information on loan officer characteristics including their numeracy score. The credit file data contains information on the loans (and loan applications) that were handled by these loan officers between 2006 and 2013. Table 1 provides definitions and full sample summary statistics of all credit file variables that we employ in our analysis. Table 2 shows summary statistics by subperiod.

A. Loan Officer Data

We have information on the characteristics of the 151 loan officers who obtained the minimum passing score (*Numeracy score*) of 65% or higher in the above described math test. We were not able to obtain information on 38 loan officers with numeracy scores below 65%. This restricts the range of the treatment but still leaves us with considerable variation in the numeracy score. Importantly, this sample restriction does not cause a bias of our estimates since the selection did not occur based on the outcome variable. Overall, the sample restriction should lead to a lower observed treatment effect between the highest and the lowest observed level of numeracy compared to the case where loan officers across all test results and their lending decisions would be observed.

The *Numeracy score* reflects the share of correctly answered questions. We exclude loan officers whose highest degree is not a bachelor degree (21 loan officers) to ensure that a potential effect of numeracy on loan officers' risk score accuracy is not driven by heterogeneity in education.[11] Figure 2 provides a histogram of the *Numeracy score* of the 130 loan officers in our final sample. We use dummy variables to distinguish three levels of numeracy. *Low numeracy* is a dummy variable that is 1 for loan officers with a numeracy score between 65% and 80%, *Medium numeracy* is a dummy variable that is 1 for loan officers with an numeracy score from 80% to 89% and *High numeracy* is a dummy variable that is 1 for all loan officers with a numeracy score of 90%–100%.[12]

Table 3 displays the average numeracy score, gender, age and work experience for our sample of loan officers by numeracy level. Table 3 shows that loan officers with a medium level of numeracy are more often female and more experienced than both high and low numeracy loan officers.

B. Credit File Data

Our initial credit file dataset consists of all 33,918 loan applications submitted over the period 2006–2013 to the bank and processed by loan officers who passed the numeracy test in February 2010. Focusing on loan officers from similar educational backgrounds reduces the sample to 29,474 loan applications. Out of these applications, 4,902 did not enter the screening stage due to formal errors, very negative credit registry information or because the client did not want to proceed further. We therefore observe 24,572 loan applications which were processed, out of which the bank made 16,540 loan offers (67%). In 856 cases, the client did not accept the loan offer leaving 15,684 granted loans in the sample.

We focus our analysis on the period July 2007 to December 2010. Since our sample contains only loan applications processed by loan officers that took the numeracy test in February 2010, there are very few loan applications in the sample for 2006 and early 2007. We begin our sample in July 2007 to ensure a sufficient number of loan applications per quarter and to cover a long enough precrisis period (five quarters). In order to rule out any influence of the mandatory math training courses that the Bank implemented in 2011 and 2012), we exclude all loan applications

11. In robustness tests we use the full sample of loan officers and control for their educational background. Results (available upon request) remain qualitatively unchanged.

12. The thresholds ensure that roughly one third of the loan applications in our final analysis sample are handled by loan officers in each numeracy level. In robustness tests we set the thresholds so that one third of loan officers are in each numeracy category and we use the linear numeracy score. In both cases results remain qualitatively unchanged.

TABLE 1
Summary Statistics and Variable Definitions

	Obs	Mean	*SD*	Min	Max	Mean Low	Mean Medium	Mean High	Description
Panel A: Granted Loans									
Dependent variable									
Arrears	6,498	0.08	0.27	0.00	1.00	0.06	0.07	0.10	Dummy = 1 if 30 day payment arrear within first 24 months
Variables of interest									
Risky	6,498	0.23	0.42	0.00	1.00	0.19	0.25	0.26	Dummy = 1 if initial score > 1
Low numeracy	6,498	0.33	0.47	0.00	1.00	1.00	0.00	0.00	Dummy = 1 if low numeracy loan officer; score < 0.8
Medium numeracy	6,498	0.32	0.47	0.00	1.00	0.00	1.00	0.00	Dummy = 1 if medium numeracy loan officer; score 0.8–0.89
High numeracy	6,498	0.35	0.48	0.00	1.00	0.00	0.00	1.00	Dummy = 1 if high numeracy loan officer; score 0.9–1
Numeracy score	6,498	0.84	0.11	0.65	1.00	0.71	0.85	0.95	Numeracy score as measured in the test
Transformed numeracy score	6,498	0.53	0.30	0.00	1.00	0.16	0.56	0.85	Transformed numeracy score: (Numeracy score − 0.65)/0.35
Basic controls									
Ln(Requested amount)	6,498	8.34	0.94	4.68	10.31	8.14	8.45	8.42	Ln(requested amount in EUR)
Requested amount in Euro	6,498	6,242	5,840	108	30,000	5,198	6,688	6,803	Requested amount in EUR
Request Euro	6,498	0.12	0.33	0.00	1.00	0.10	0.11	0.15	Dummy = 1 if requested loan in Euro
Time relationship	6,498	1.86	1.84	0.00	8.18	1.88	1.85	1.85	Years since bank account at bank; 0 if no account
Extended controls									
Leverage	6,221	1.35	11.39	0.01	807.43	0.99	1.72	1.36	
Ln(Sales)	6,305	7.35	1.44	3.24	13.14	7.02	7.60	7.43	Ln(Sales in EUR)
Young firm	6,498	0.23	0.42	0.00	1.00	0.17	0.30	0.23	Dummy = 1 if firm age < 5
Agriculture	6,498	0.48	0.50	0.00	1.00	0.63	0.35	0.47	Dummy = 1 if agricultural firm
Total assets/requested amount	6,440	5.42	10.57	0.00	443.31	5.13	5.23	5.85	(Fixed assets and chattel items)/requested amount
Loan officer controls									
Female	6,498	0.55	0.50	0.00	1.00	0.47	0.62	0.56	Dummy = 1 if loan officer female
Experienced	6,498	0.58	0.49	0.00	1.00	0.50	0.67	0.57	Dummy = 1 if loan officer experience > median at test date
Age	6,498	32.1	2.86	27.00	41.00	32.26	32.02	32.02	Age in years
Panel B: Loan applications									
Dependent variable									
Granted	10,562	0.62	0.49	0.00	1.00	0.67	0.60	0.58	Dummy = 1 if application granted by the bank
Variables of interest									
Low numeracy	10,562	0.30	0.46	0.00	1.00	1.00	0.00	0.00	Dummy = 1 if low numeracy loan officer; score < 0.8
Medium numeracy	10,562	0.33	0.47	0.00	1.00	0.00	1.00	0.00	Dummy = 1 if medium numeracy loan officer; score 0.8–0.89
High numeracy	10,562	0.37	0.48	0.00	1.00	0.00	0.00	1.00	Dummy = 1 if high numeracy loan officer; score 0.9–1
Control variables									
Ln(Requested amount)	10,562	8.48	0.95	4.68	10.31	8.27	8.58	8.55	Ln(requested amount in EUR)
Requested amount in Euro	10,562	7,126	6,307	108	30,000	5,922	7,557	7,706	Requested amount in EUR
Request Euro	10,562	0.14	0.35	0.00	1.00	0.12	0.13	0.17	Dummy = 1 if requested loan in Euro
Time relationship	10,562	1.19	1.72	0.00	8.18	1.31	1.15	1.12	Years since bank account at bank; 0 if no account

TABLE 2
Variable Mean by Period and Numeracy Level

	Precrisis: 2007 July–2008 September			Crisis: 2008 October–2010 December		
	Low	Medium	High	Low	Medium	High
Panel A: Granted loans						
Obs	163	425	325	1953	1,673	1956
Number of loan officers	12	29	35	34	42	53
Average number of loans per loan officer	13.58	14.66	9.29	57.53	39.83	36.91
Dependent variable						
Arrears	0.10	0.05	0.09	0.06	0.08	0.11
Variables of interest						
Risky	0.13	0.09	0.15	0.19	0.29	0.28
Numeracy score	0.72	0.85	0.95	0.7	0.85	0.95
Transformed numeracy score	0.21	0.56	0.84	0.16	0.56	0.85
Basic controls						
Ln(Requested amount)	8.74	8.36	8.7	8.09	8.47	8.37
Requested amount in Euro	9,138	6,816	8,455	4,870	6,656	6,528
Request Euro	0.08	0.02	0.06	0.1	0.14	0.16
Time relationship	1.07	1.02	1.01	1.94	2.06	1.99
Extended controls						
Leverage	0.73	0.95	1.35	1.01	1.9	1.37
Ln(Sales)	7.88	7.4	7.93	6.96	7.65	7.36
Young firm	0.33	0.32	0.42	0.16	0.29	0.2
Agriculture	0.48	0.45	0.26	0.65	0.33	0.5
Total assets/requested amount	5.63	4.59	4.38	5.09	5.4	6.1
Loan officer controls						
Female	0.58	0.6	0.66	0.47	0.63	0.55
Experienced	0.93	0.96	0.96	0.47	0.6	0.51
Age	33.3	32.81	32.88	32.17	31.82	31.87
Panel B: Loan applications						
Obs	262	691	522	2,881	2,791	3,415
Number of loan officers	15	30	38	34	42	53
Average number of applications per loan officer	17.47	23.03	13.74	84.74	66.45	64.43
Granted	0.62	0.62	0.62	0.68	0.60	0.57
Control variables						
Ln(Requested amount)	8.70	8.49	8.76	8.23	8.60	8.52
Requested amount in Euro	8,891	7,708	8,954	5,652	7,519	7,515
Request Euro	0.08	0.03	0.06	0.12	0.16	0.18
Time relationship	0.69	0.66	0.66	1.37	1.27	1.19

made after December 2010.[13] For our period of interest we observe 13,115 loan applications of which 8,126 are granted.[14]

Furthermore, we only include applications for loans up to 30,000 Euros into our analysis. Applications for larger volumes are less frequent and most often processed by credit analysts whose job description differs from the job description of loan officers. Our loan sample contains only first-time borrowers. Since no information from previous loans is available for first time borrowers, screening is most difficult and any effect from numeracy should be most prevalent. Also, the focus on first-time borrowers ensures that the assignment of loan applications to loan officers is not influenced by past loan performance.[15] Our final dataset contains 10,562 loan applications and 6,498 loans granted to firms without prior credit relationships with the bank.[16] These loan applications were screened by 130 loan officers at 31 bank branches over the period July 2007 to December 2010.

For each loan application, we know which loan officer handled it and can therefore match loan application and loan officer data. For loan

13. To rule out that the accuracy of risk assessments is affected by the math test itself we conduct a robustness test in which we restrict the sample to those loans granted before the test date only. Our findings are confirmed.

14. Of the dropped loan applications, 1,156 are from the period before July 2007 and 9,445 are from the period after December 2010.

15. This comes at the disadvantage that we cannot observe differences over a client relationship as for example documented for credit rationing by Kirschenmann (2016).

16. Overall, 1,102 loan applications for loans larger than 30,000 Euros and 1,451 loan applications of repeat borrowers are dropped.

FIGURE 2
Distribution of Numeracy Score in Loan Officer Sample

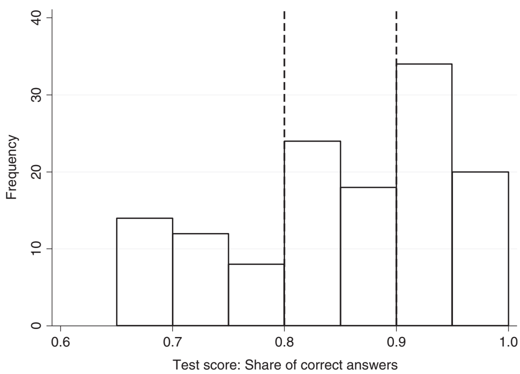


TABLE 3
Loan Officer Summary Statistics

Numeracy Score Range (%)	Low 65–79	Medium 80–89	High 90–100	Total
Nr Loan officers	34	42	54	130
Initial numeracy score	0.72	0.85	0.94	0.85
Female	0.53	0.76	0.65	0.65
Experienced	0.35	0.60	0.52	0.50
Age	31.82	32.07	32.44	32.16

applications, the dataset further contains information on the requested amount, the requested currency,[17] the opening date of the client's account with the bank as well as the involved bank branch. For granted loans, the dataset contains additional information on the borrowing firm at application date (financial information, industry, and firm age), the granted loan terms (volume, currency, interest rate, collateral, maturity) and the initial internal risk rating (which ranges from 1 [lowest risk] to 3). In our final sample used in the empirical analysis, we have 4,999 loans with an initial risk score of 1 and 1,388 (111) loans with an initial score of 2 (3). In our analysis, the variable *Risky* reflects the initial risk rating at loan disbursement. Given the low number of loans with risk score 3, we construct *Risky* as a binary variable that takes on the value 1 if a loan is assigned an initial risk score of 2 or 3 and zero if the loan is assigned a risk score of 1.[18]

We observe semi-annual information on the performance of granted loans as measured by the days in payment arrear. We construct the variable *Arrears* which captures the performance of each loan during the first 24 months after the loan was disbursed. We focus on the first 24 months since initial credit assessment processes in commercial banks are designed to capture potential loan defaults in the first years after disbursement.[19] For each loan, the days in arrear are reported for end of June and end of December. Hence, we can retrace when exactly each loan falls into arrears for at least 30 days for the first time. The binary variable *Arrears* then takes on the value 1 if a loan falls into arrears for at least 30 days within the first 24 months. On average, 8% of the loans in our final sample fall into arrears for at least 30 days during the first 24 months of their maturity. Figure 3 displays the share of nonrisky (solid line) and risky (dashed line) loans that have *not* fallen into 30-day arrears over the first 24 months after loan disbursement. At each point in time, the share of nonrisky loans that is not in arrears is higher than the share of risky loans not in arrears, with the difference between the two increasing steadily. The figure also highlights that the incidence of falling into arrears occurs quite evenly distributed over time for both risky and nonrisky loans.

IV. METHODOLOGY

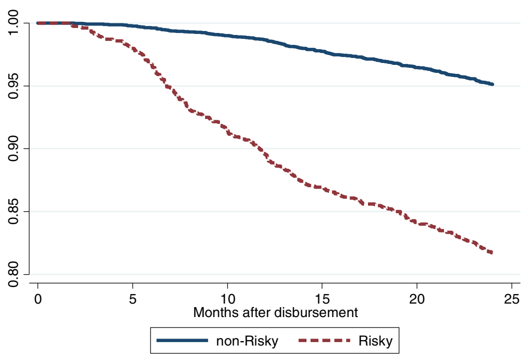
Our objective is to analyze how the level of loan officer numeracy is related to the accuracy of

17. Only 2% of loans were granted in a currency different from the requested currency (for 1% of loans, the application was in Euro and the granted loan in RON and for 1% of loans the application was in RON and the granted loan in Euro). There is no evidence that adjustments of the loan currency substantially differ by the level of loan officer numeracy and that bank-wide changes influencing the loan currency (e.g. the funding structure (Brown, Kirschenmann, and Ongena 2014)) would affect loan officers with different numeracy differently.

18. Table 2 shows that the likelihood to assign a loan as risky is very similar across loan officer numeracy levels in the pre-crisis period: 0.13 for low numeracy loan officers and 0.15 (0.09) for high (medium) numeracy loan officers. Table 2 further shows that loan officers are more likely to classify borrowers as risky during the crisis period. However, the increase in the share of loans classified as risky is much lower for low-numeracy loan officers than for loan officers with high or medium numeracy. Apart from that, the risk classification of a borrower is similarly related to observable characteristics of the borrower and his application across loan officers' numeracy levels both in the pre-crisis and in the crisis periods (results are available from the authors upon request). These findings are a first indication that low numeracy loan officers are less accurate in their risk assignments and, in the pre-crisis period in particular, not just more reluctant to classify loans as risky.

19. In small business lending, banks typically update their credit assessment annually, when new financial statement data on the firm becomes available through its annual accounts.

FIGURE 3
30 Day Arrears over the First 24 Months



Notes: The graph displays the share of loans falling into 30 day arrear over the first 24 months. The lines display the share of loans that have not been in 30 day arrear at any time after disbursement.

their credit assessments. Consider a bank which is recruiting loan officers from a population of interest, that is, in our case college graduates. The bank is interested in how the accuracy of its credit assessments will change if it hires college graduates with high numerical skills rather than college graduates with lower numerical skills. Our analysis provides an estimate of how such a change in recruiting standards may affect the precision of the bank's borrower screening process.

For a given portfolio of loan applications L the bank is interested in estimating the average treatment effect of replacing a low numeracy loan officer with a high numeracy loan officer. We define A as the accuracy level and N as the numeracy level of the loan officer employed by the bank. The average treatment effect is then given by:

$$(1)\quad ATE = E[A(N = high, L) - A(N = low, L)]$$

In order to estimate the average treatment effect in Equation (1) one possible experiment would be the following: First, the bank randomly chooses loan officers from the population of interest (e.g., college graduates). The bank then randomly assigns loan applications to these loan officers. We would then measure the accuracy of the credit assessments A for each loan officer and compare the average accuracy of loan officers with a high numeracy level to the average accuracy of loan officers with a low numeracy level.

Our empirical analysis of the administrative data presented above deviates from this ideal experiment in two crucial dimensions: measurement and identification. First, due to the small number of observations the available data does not allow us to measure the accuracy of credit assessments at the loan officer level, but only for groups of loan officers. Second, loan officers in our sample are hardly randomly chosen, and loan applications are hardly randomly assigned to loan officers. In the following, we first discuss how we measure the accuracy of credit assessments. We then discuss our identification strategy.

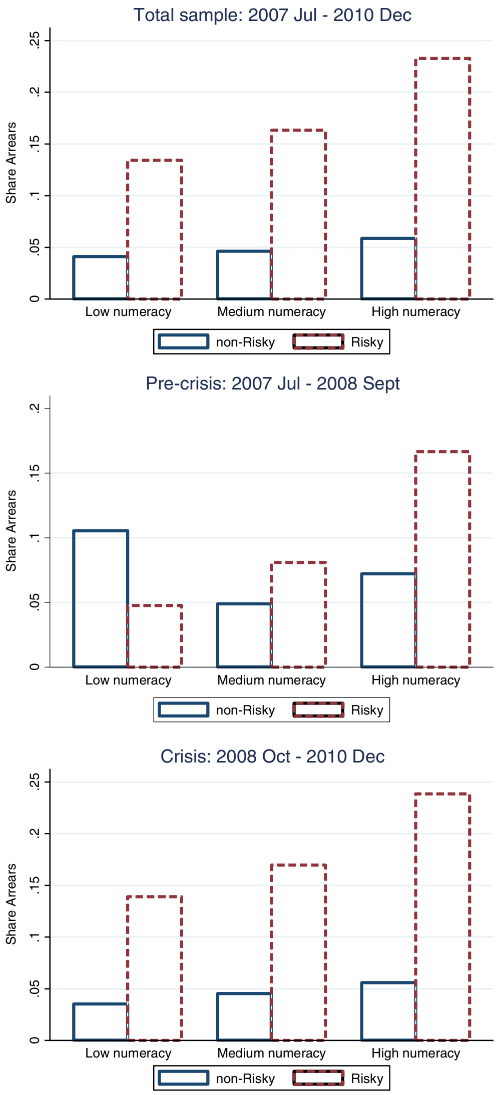
A. Measuring and Comparing Loan Officer Accuracy

We measure the accuracy of loan officers' credit assessments by comparing their ex ante risk assessment of a borrower to the ex post performance of that borrower's loan. This approach follows the methodology applied to assess the discriminatory power of internal rating systems, that is, the system's ability to discriminate ex ante between defaulting and nondefaulting borrowers (Bank for International Settlements [BIS] 2005).

For each granted loan in our sample we observe the initial risk rating as assigned before loan disbursement by the loan officer. We hereby distinguish *Risky* (initial risk score = 2 or 3) from nonrisky (initial risk score = 1) loans. We also observe whether a loan falls into *Arrears* within 24 months of disbursement. A loan officer who is very accurate in assessing the creditworthiness of borrowers would classify most loans as nonrisky which ex post are not in arrears, while he would classify most loans as risky that fall into arrears. Thus, in the portfolio of loans handled by an accurate loan officer we should see that the share of defaulting loans among those classified as risky is much higher than the share of defaulting loans among those classified as nonrisky. By contrast, the portfolio of a loan officer who is not accurate at all would display a similar share of defaulting loans, irrespective of whether the loan was rated as risky or nonrisky.

Figure 4 displays the share of loans falling into arrears by risk rating, loan officer numeracy and sub-period. Starting with the total sample in the top panel, the graph shows that borrowers initially classified as risky (dashed bars) are more likely to fall into arrears than borrowers initially classified as nonrisky (solid bars), and the discriminating power is largest for the high numeracy loan officers. The same pattern holds for the crisis period. For the precrisis sample we find that for *low* numeracy loan officers a higher share of nonrisky loans falls into payment arrear than of risky loans. Hence, during this period the initial

FIGURE 4
Accuracy of Initial Rating



Notes: The figure displays the share of loans in 30 day payment arrear within 24 months after loan issuance by initial risk rating and numeracy.

rating of these loan officers is unable to discriminate borrowers by creditworthiness.[20]

20. Of the 12 low numeracy loan officers who grant loans in the pre-crisis period, 5 loan officers only classify loans as nonrisky and none of these 31 loans are subsequently in arrears. 7 loan officers classify at least 1 loan as risky. Among these, 2 loan officers experience no arrears among their granted loans. The remaining 5 loan officers classify loans both as risky and nonrisky and experience at least 1 loan in arrears. Each of these 5 loan officers is subject to classification error: they all display a higher share of defaults among those loans which they classify as nonrisky.

To formally measure and compare the accuracy of credit assessments across loan portfolios processed by loan officers with different numeracy scores we choose the following methodology: Consider a portfolio consisting of $l = 1 \ldots L$ loans and the following linear probability model:

$$(2) \qquad Arrears_l = \alpha + \beta \cdot Risky_l + \epsilon_l$$

The estimated coefficient β from this regression provides us with an indicator of the discriminatory power of the initial risk rating for the underlying portfolio of loans. If the risk rating cannot discriminate between those loans which fall into arrears and those that do not, we would yield an estimated coefficient of β=0.[21] If the risk rating perfectly discriminates between those loans which fall into arrears and those that do not, we would yield an estimated coefficient of β=1.[22]

Applying Equation (2) we can formally compare the discriminatory power of the risk rating across two portfolios of loans l and l'. Specifically, we can estimate β within portfolio l and β' within portfolio l'. We can then compare the estimated coefficients β and β' with a Chow test. This is the methodology we pursue in this paper to measure and compare the accuracy of credit assessments by loan officer numeracy. We split our sample of 6,498 loans into three portfolios based on whether the loan was processed by a high, medium or low numeracy loan officer. Applying Equation (2) to each subsample separately we estimate β^{high}, β^{medium} , and β^{Low}.We then compare these estimated coefficients applying a Chow test.[23]

Two choices regarding our estimation strategy warrant discussion: (i) the choice of a categorical numeracy variable rather than a continuous numeracy score and (ii) the choice to conduct sample splits by numeracy level and report Chow

21. In this case the estimated constant *a* would equal the average default rate in the portfolio.

22. In this case the estimated constant *a* would equal zero and *Risky* would be perfectly collinear with *Arrears*.

23. An alternative approach for measuring the discriminatory power of risk ratings is to calculate the accuracy ratio (see e.g., BIS 2005; Engelmann, Hayden, and Tasche 2003; Moody's Investor Services 2003). The accuracy ratio compares the ratio of the correctly classified loans within a loan portfolio to the classification of a perfect model and a random model. However, a major drawback of using the accuracy ratio for our purpose is that there is no method for formally comparing the measure across loan portfolios, that is, for loans processed by low numeracy as opposed to high numeracy loan officers.

tests rather than estimating full sample regressions with interaction terms. We employ a categorical variable indicating low, medium and high numeracy because nonlinear effects of numeracy on accuracy seem likely. For instance, the difference in accuracy might well be higher between low and medium numeracy loan officers compared to the difference between medium and high numeracy loan officers. Arguably, borrowers of low numeracy loan officers are different from those of high numeracy loan officers (see the summary statistics in Table 2). To allow for such differences and to allow all of the variables to have separate coefficients we estimate separate regressions by numeracy level in our main specifications. We resort to sample splits and Chow tests rather than a fully interacted model because the latter suffers from a high degree of multicollinearity due to the many interaction terms. However, in robustness tests we confirm our main results in full sample regressions where all the explanatory variables are interacted with either the numeracy levels or a continuous numeracy score.

In principle, we could estimate Equation (2) separately for each loan officer. We would then obtain a measure of individual loan officer accuracy as presented in Equation (1). However, with the administrative data at hand it is not feasible to estimate accuracy indicators at the loan officer level with reasonable precision. The precision of the estimated coefficient β in the linear regression (2) depends on the size of the underlying loan portfolio and the share of loans which actually default. A crucial limitation to studies of bank credit risk is that only a small share of loans actually defaults. In our sample 8% of the loans enter into payment arrears within 24 months of loan disbursement. Our sample consists of 6,489 granted loans handled by 130 loan officers and thus an average of 50 loans per loan officer. With a default rate of 8% this implies that on average just 4 loans fall into arrears per loan officer. Given the limited number of loans handled by each loan officer and the low default rate it is thus not feasible to precisely measure the accuracy ratio at the loan officer level.

B. Identification

We apply regression (2) to measure the accuracy of the initial risk ratings separately for the portfolios of loans processed by high numeracy, medium numeracy and low numeracy loan officers, respectively. The comparison of β^{high}, β^{medium}, and β^{Low} provide us with an estimate of how loan officer numeracy is related to accuracy if (i) observed numeracy is orthogonal to other loan officer characteristics which may affect the accuracy of their credit assessments and (ii) loan applications are randomly assigned to loan officers. It is unlikely that either of these assumptions hold. Our analysis thus faces two main identification challenges. First, other loan officer characteristics such as education, age, gender or job experience might be correlated with both, loan officers' numeracy levels and the accuracy of their credit assessments. Second, the assignment of loan applications to loan officers is likely to be influenced by numeracy or related characteristics and therefore the unobserved counterfactual accuracy is not equal to the observed outcomes.

To address these challenges, we augment Equation (2) with two vectors of control variables that capture loan officer characteristics LO_j and loan application characteristics X_i. We estimate the following linear probability model for each numeracy level n separately[24]:

$$(3) \quad Arrears_{i,j} = \alpha + \beta_n \cdot Risky_i + \delta \cdot LO_j + \gamma \cdot X_i + \epsilon_{i,j}$$

As discussed above, the coefficient of primary interest in Equation (3) is β_n. It captures the discriminatory power of the initial rating *Risky* within the portfolios of loans processed by loan officers with numeracy level n.

LO_j is a vector of observable loan officer characteristics that are likely to be correlated with numeracy and the accuracy of loan officers' credit assessments. Beck, Behr, and Guettler (2013) find that the loan portfolios of female loan officers perform better than those of male loan officers. Since the effect is most pronounced when female loan officers handle loans of female borrowers, they conclude that female loan officers are better in building trust relationships with their clients. *Female* thus is a dummy that is 1 of the loan officer is female and 0 if male. Andersson (2004) and Bruns et al. (2008) show that job experience or specific human capital might matter for loan officers' lending decisions and the decision process. We therefore include

24. The comparison of coefficients across groups comes with very strong assumptions in nonlinear models. We therefore prefer a linear probability model that comes at the cost of mis-specifying the function form of the dependent variable. In robustness tests we estimate the same effect in a nonlinear logit model but without applying Chow tests. The results confirm our main findings.

Experienced which is a dummy variable that is 1 for loan officers who have worked with the bank for more years than the median of work years at the math test date (2.13 years). *Age* captures the age of the loan officer in years to control for the general life experience of the loan officers.

X_i is a vector of loan-level covariates controlling for factors that could potentially influence the assignment of a loan application to a high numeracy loan officer and be correlated with the potential accuracy of the credit assessment, that is, the difficulty of assessing the creditworthiness of the borrower. A profit-maximizing bank should employ the most skilled loan officers where their skills can generate the highest profit. Intuitively, we would expect banks to allocate those loan applications which are most difficult to assess to their best loan officers. However, it is also feasible that the allocation of loan applications is driven by borrower characteristics that most strongly influence the bank's profit but that, at the same time, make the assessment easier. For instance, the more able loan officers might be assigned to deal with the larger clients, which also have more accurate financial information.

We would like to control for all loan-level or firm-level characteristics which may confound the relationship between loan officer numeracy and the potential accuracy of credit assessments. At the same time we should avoid using endogenous control variables, that is, firm-level or loan-level variables which may be influenced by the numeracy level of the loan officer processing the application. We therefore employ two sets of application and firm control variables. Basic controls contain loan and firm characteristics elicited in the loan application form: The measurement of these variables is thus arguably independent of the loan officer's numeracy level. *Ln(Requested amount)* controls for the volume of the application and *Request Euro* for the requested currency. *Time relationship*, a variable reflecting the years that a firm has an account at the bank, controls for the level of information about the firm that is available within the bank and thus is also a measure of opaqueness.

Extended firm-level controls include variables which are elicited or verified during the credit assessment process: *Leverage*, *ln(Sales)*, *Young firm*, *Agriculture and Total assets/requested amount*. These variables allow controlling for firm size, riskiness, industry and opaqueness in more detail. However, these variables are also potentially influenced by the loan officer's verification procedure and are therefore potentially endogenous control variables. *Ln(Sales)* controls for the size of the applicant and *Total assets/requested amount* for the relative size of the loan application. *Leverage*, defined as the debt capital and the applied loan amount over equity, should provide some obvious signals about the riskiness of the loan application. *Agriculture* is a dummy variable taking on the value 1 if a firm is active in agriculture. *Young firm*, a binary variable capturing firms that were founded less than 5 years prior to the loan application, controls for the firm's opaqueness.

We further include branch fixed effects and quarter fixed effects. The branch fixed effects control for the general local economic environment as well as branch-specific practices. The branch fixed effects are also important to control for the time-invariant characteristics and the numeracy of the branch manager as he / she forms part of the credit committee that checks the credit score and makes the final lending decision.[25] The quarter fixed effects control for the changing macroeconomic conditions during the boom and bust cycle.[26]

Regarding the interpretation of our results, we note that our observable measure of numeracy is very likely correlated with unobservable personal traits of loan officers such as general cognitive ability and social skills. This implies that our estimated "effect" captures the combined effect of numerical skills and the broader set of correlated cognitive and social skills. Our results can therefore not be interpreted as the potential gain to a bank (or other employers) of promoting the numerical skills of employees, for example, through an education intervention. Rather, as hinted at the beginning of this section our results can inform us about the potential gain to a firm of hiring staff with high observable numerical skills

25. Unfortunately, we do not have comprehensive and detailed information on the branch manager characteristics and the credit committee. We have information on the composition of the credit committee from mid-2010 onwards and for 80% of the loans the credit committee consists of the loan officer and the branch manager. For the other 20% the credit committee consists of the branch manager and of a credit risk officer located at the bank's headquarter. Therefore, the branch dummies do not fully capture the influence of the credit committee or the branch manager.

26. For example, in the first quarter of 2009 more than 95% of issued loans in the sample were classified as risky compared to 10%–20% in the quarters before and after. Obviously, the bank made some short-term adjustments to its policies at the beginning of the crisis; however these adjustments apply to all loan officers independent of their numeracy level.

(and related, but less observable, cognitive and social abilities).

V. RESULTS

A. Numeracy and Accuracy

Table 4 presents our baseline estimates. In each column the coefficient of *Risky* reflects the degree to which loan officers in that subsample are able to discriminate borrowers by their creditworthiness. Hence, a higher estimate for *Risky* reflects more accurate credit decisions. Results of the Chow test comparing the coefficients across numeracy levels are presented in the bottom panel of the table. Columns 1–3 display results of the estimation controlling only for basic control variables, loan officer controls and branch fixed effects. In columns 4–6 we add quarter fixed effects and in columns 7–9 extended control variables. Standard errors are heteroscedasticity robust and clustered at the loan officer level.

Considering the specification with basic controls and branch fixed effects only, the magnitude of the estimated coefficient of *Risky* is substantially larger in the sample of loans processed by high numeracy loan officers (column 3: 0.219) as compared to loans processed by low numeracy loan officers (column 1: 0.107) or medium numeracy loan officers (column 2: 0.130). Chow tests reported in the bottom part of the table confirm that the credit assessments of high numeracy loan officers are significantly more accurate than those of low and medium numeracy loan officers. We yield almost identical results in the specifications including quarter fixed effects (columns 4–6) and extended controls (columns 7–9).

In Table 5 we present separate results for the subsample of loans granted in the precrisis and crisis periods. For both subperiods the difference between the estimates of *Risky* for low and high numeracy loan officers is statistically significant at the 1%-level. The difference is, however, larger in the precrisis period (−0.239 vs. -0.108).

In the precrisis period (column 1) the predictive power of the risk rating of loans processed by low numeracy loan officers is even worse than a random assignment. The ability to discriminate borrowers by quality improves significantly for all numeracy levels in the crisis period with the improvement being largest for the low numeracy loan officers. These findings are in line with Becker, Bos, and Roszbach (2018) who show that it is most difficult to accurately sort borrowers according to their riskiness during boom periods in which informational frictions are highest.[27]

Several alternative rationales exist that might explain the improved accuracy in the crisis period. First, it could be that the hiring policy at the bank changed once the crisis unfolded. Table A4 reports results for the subsample of only those loan officers who worked at the bank already before the crisis and we find our main results confirmed. The improved accuracy in the crisis period therefore does not stem from the hiring of better loan officers after the start of the crisis. Second, it could be that low and high numeracy loan officers experience arrear events of the loans that they granted before the crisis at different points in time, which could systematically influence their screening behavior during the crisis. When we compare Kaplan–Meier survival estimates (available upon request) for loans disbursed in the precrisis period by low, medium and high numeracy loan officers, we do not find systematic differences in the timing when arrears occur. For instance, independent of the loan officer's numeracy level almost no arrear events occur during the first 6 months after a loan's disbursement and the incidence of arrears slowly increases the longer the time since a loan's disbursement. Third, an alternative explanation for the improved accuracy of low numeracy loan officers (compared to high numeracy loan officers) could be that they became more rigorous in their assessment of loan applicants once the crisis started. An analysis of the processing time of loan applications by numeracy level over our sample period shows that, on average, the processing time increases for all loan officers after the start of the crisis (see Figure A2). However, mean processing times increase the least for low numeracy loan officers. Thus, the relative improvement in the

27. Our main results in Tables 4 and 5 are robust to several alternative specifications. First, in Table A1 we estimate our main regressions using the full sample (instead of sample splits) and interact all explanatory variables with *High numeracy* and *Medium numeracy*. Second, to account for the arguably arbitrary allocation of loan officers to the low, medium and high numeracy groups, we use the linear numeracy score in Table A2. We again estimate our main regressions using the full sample, but this time interacting all explanatory variables with the linear numeracy score. These analyses show that the results do not depend on the definition of the numeracy variable nor the use of sample splits or interaction terms. We also confirm our main results in a shorter sample that only includes loans made before the test date in February 2010 to exclude any potential effect from the test itself on loan officers' lending decisions (see Table A3). In addition, we estimate a nonlinear logit model and the results (available upon request) qualitatively confirm our main findings.

TABLE 4
Numeracy and Accuracy: Full Sample Results

	Total sample: 2007 July–2010 December								
OLS Regression	**Basic Controls**			**Basic Controls with Quarter FE**			**Extended Controls with Quarter FE**		
Numeracy Level	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Dependent variable: Arrears	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Risky	0.107***	0.130***	0.219***	0.121***	0.139***	0.234***	0.060**	0.106***	0.180***
	(0.029)	(0.026)	(0.025)	(0.030)	(0.024)	(0.029)	(0.024)	(0.023)	(0.031)
Mean arrears	0.06	0.07	0.10	0.06	0.07	0.10	0.06	0.07	0.10
Observations	2,119	2,098	2,281	2,119	2,098	2,281	2,040	1,997	2,165
Adj. R-squared	0.044	0.052	0.093	0.051	0.065	0.104	0.063	0.065	0.090
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	No	No	No	No	No	No	Yes	Yes	Yes
Loan officer controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Difference in coefficients of risky: p values of Chow test in parentheses									
Compared to high numeracy	−0.112***	−0.089**		−0.113***	−0.095**		−0.120***	−0.074*	
	(0.003)	(0.012)		(0.006)	(0.010)		(0.002)	(0.051)	
Compared to medium numeracy	−0.023		0.089**	−0.018		0.095**	−0.046		0.076*
	(0.546)		(0.012)	(0.633)		(0.010)	(0.155)		(0.051)

Notes: The dependent variable Arrears is a binary variable equal to 1 if a firm went into 30 day payment arrear within the first 24 months of the loan. Basic controls include Ln(Requested amount), Request Euro, Time relationship. Extended controls include Leverage, ln(Sales), Young firm, Agriculture, Total assets/requested amount. Loan officer controls include Female, Experienced and Age. FE, fixed effects. Standard errors in parentheses; standard errors are clustered on loan officer level; *$p < .1$, **$p < .05$, ***$p < .01$. We compare the coefficients of Risky by numeracy level using a Chow test.

TABLE 5
Numeracy and Accuracy: Subperiod Analysis

OLS Regression	Precrisis: 2007 July–2008 September			Crisis: 2008 October–2010 December		
Numeracy Level	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Dependent Variable: Arrears	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Risky	−0.108*	−0.004	0.131	0.075***	0.117***	0.183***
	(0.059)	(0.036)	(0.078)	(0.023)	(0.025)	(0.034)
Mean arrears	0.10	0.05	0.09	0.06	0.08	0.11
Observations	151	389	293	1,889	1,608	1,872
Adj. R-squared	0.033	0.049	0.010	0.068	0.073	0.100
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer controls	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Difference in coefficients of risky: p values of Chow test in parentheses						
Compared to high numeracy	−0.239***	−0.135*		−0.108***	−0.066	
	(0.008)	(0.087)		(0.009)	(0.105)	
Compared to medium numeracy	−0.104		0.135*	−0.042		0.066
	(0.101)		(0.087)	(0.236)		(0.105)

Notes: The dependent variable Arrears is a binary variable equal to 1 if a firm went into 30 day payment arrear within the first 24 months of the loan. Basic controls include Ln(Requested amount), Request Euro, Time relationship. Extended controls include Leverage, ln(Sales), Young firm, Agriculture, Total assets/requested amount. Loan officer controls include Female, Experienced and Age. Standard errors in parentheses; standard errors are clustered on loan officer level; *$p<.1$, **$p<.05$, ***$p<.01$. We compare the coefficients of Risky by numeracy level using a Chow test.

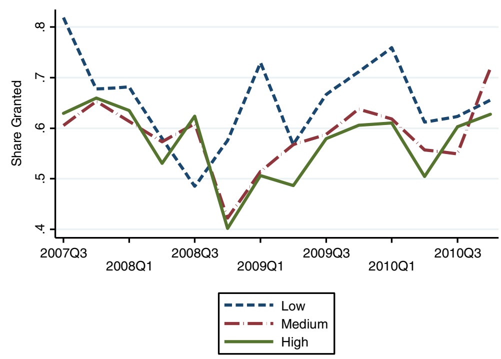
accuracy of low numeracy loan officers does not seem to be driven by a more diligent assessment. Rather our results support the conjecture that it is most difficult to sort borrowers according to their riskiness during boom periods.

B. *Loan Approval and Sample Selection*

The analysis so far has focused on the sample of granted loans and studied the accuracy of loan officers' credit assessments by comparing ex ante risk ratings with ex post loan performance. However, if numeracy is related to the ability to pick out risky borrowers, it might also lead to systematic differences between the samples of loans which are approved in the first place. The observed differences in the screening performance of loan officers of different numeracy levels would then be influenced by their preceding approval vs. rejection decisions.

Our dataset covers all loan applications processed by our sample of loan officers during the sample period. Figure 5 displays the quarterly approval rate for first time applicants by the level of the loan officers' numeracy. Over the entire sample period 62% of all loan applications are granted (see also Table 1, Panel B). Low numeracy loan officers display higher approval rates (67%) compared to loan officers with medium numeracy (60%) and high numeracy (58%).

FIGURE 5
Quarterly Approval Rate by Numeracy over the Sample Period



Notes: Share of granted first time loans by quarter and level of numeracy.

In Table 6, we estimate a linear probability model of the approval decision. The dependent variable is *Granted*, which is a dummy variable that is 1 if the loan application is granted and 0 if it is rejected. All regressions include as explanatory variables the loan application characteristics *(Ln)Requested amount*, *Request Euro* and *Time relationship*. All regressions further include controls for loan officer characteristics (gender, experience, age) as well as for branch and quarter fixed effects.

TABLE 6
Numeracy and Loan Rejections

Dependent Variable: Granted	Subsample by Numeracy Level Low (1)	Medium (2)	High (3)	Total Sample: 2007 July–2010 December (4)	Precrisis: 2007 July–2008 September (5)	Crisis: 2008 October–2010 December (6)
High Numeracy				−0.036*	−0.001	−0.043*
				(0.021)	(0.063)	(0.022)
Medium Numeracy				−0.046*	0.035	−0.046
				(0.025)	(0.061)	(0.029)
Ln(Requested amount)	−0.063***	−0.057***	−0.062***	−0.064***	−0.071***	−0.064***
	(0.010)	(0.012)	(0.010)	(0.007)	(0.010)	(0.008)
Request Euro	−0.052*	0.014	0.012	−0.000	−0.042	0.003
	(0.030)	(0.025)	(0.028)	(0.017)	(0.052)	(0.019)
Time relationship	0.135***	0.143***	0.154***	0.145***	0.196***	0.142***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.011)	(0.005)
Mean Granted	0.67	0.60	0.58	0.62	0.62	0.61
Observations	3,143	3,482	3,937	10,562	1,475	9,087
Adj. R-squared	0.313	0.312	0.311	0.305	0.222	0.323
Loan officer controls	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable Granted is a binary variable equal to 1 if a loan application was granted and 0 otherwise. Loan officer controls include Female, Experienced and Age. Standard errors are clustered on loan officer level; $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$. We compare the coefficients of available application controls in the subsample analysis (1)–(3) using a Chow test. Results suggest that the difference in the coefficients of application controls in the subsamples of low and high numeracy loan officers is only slightly significant for Request Euro. Comparing coefficients of the medium numeracy subsample to high/low subsamples, the only significant difference (10%-level) exists for Request Euro between medium and high numeracy. FE, fixed effects.

In columns 1–3 of Table 6 we estimate the model separately for low, medium and high numeracy loan officers. The results suggest that—at all levels of numeracy—loan officers are more likely to approve applications for smaller loans as well as applications from clients with a longer relationship with the bank. We then compare the column 1–3 coefficients across numeracy levels applying Chow tests. We only find a slightly significant difference for *Request Euro* between coefficients of low and high numeracy loan officers. Thus the approval behavior of loan officers seems to be similarly related to observable borrower characteristics, independent of the loan officer's numeracy level.

The observed differences in average approval rates between the low versus medium / high numeracy loan officers could be caused by differences in the assigned application pool. Comparing the characteristics of loan applications (see Table 1, Panel B) highlights that medium and high numeracy loan officers are indeed more likely to handle loan applications with a larger requested loan size as well as applications from clients with shorter bank relationships. In columns 4–6 of Table 6, we examine whether loan officer numeracy influences approval rates conditional on loan application characteristics. We pool the samples of applications across loan officers and add our indicators of *High numeracy* and *Medium numeracy* to the regression model. Column 4 reports results for the full sample period, while columns 5 and 6 report results for the precrisis and crisis period separately. The column 4–6 estimates show that, controlling for loan application characteristics, high numeracy loan officers are significantly less likely to approve loans than low numeracy loan officers. Over the entire observation period the estimated difference in approval rates is 3.6 percentage points. This amounts to almost 6 % of the average approval rate in the sample (62%) and accounts for more than one-third of the observed difference in approval rates between low and high numeracy loan officers. The sub-period analysis shows that there is no significant difference in the approval rate before the crisis (column 5) but that the significantly lower approval rate of high compared to low numeracy loan officers observed in the full sample stems from the crisis period (column 6).

We implement a Heckman two-step model to account for sample selection in our main analysis of granted loans. Specifically, we estimate a predicted probability of approval for each loan in our main sample of granted loans. We augment the approval regression with *New client*, a binary variable equal to 1 if a client has no account history with the bank, to satisfy the exclusion restriction. We then add the corresponding inverse Mill's ratio as a further control to our main regression analysis. The results are presented in Table 7, where Panel A displays our main analysis and Panel B our first-stage results. The results confirm our main findings: Panel A of Table 7 shows that the credit assessments of high numeracy loan officers are more accurate than those of medium and low numeracy loan officers and that the effect is more pronounced in the precrisis period. The coefficient of the inverse Mill's ratio in Panel A of Table 7 does not suggest a robust selection effect in our main analysis. The negative coefficient in eight out of nine columns is in line with the conjecture that loans selected for approval are less likely to fall into arrears. However, the coefficient is significant in only three out of the nine columns.

C. *Hard Versus Soft Information*

Our baseline analysis shows that higher numerical skills are associated with more accurate credit assessments. There are two potential drivers of this superior accuracy. First, high-numeracy loan officers may be better able to draw meaning from existing "hard" quantitative information on the borrower. Second, they may be better able to assess and verify "soft" qualitative information. Our definition of hard and soft information is based on Liberti and Petersen (2017). They define three main characteristics that relate to hard information: Numbers (vs. text), the unimportance of context and the possible separation of information collection and decision-making. In our context, we accordingly define all financial information as hard information. In contrast, the main components of soft information in our setting—the entrepreneur's character and managerial ability as well as the market outlook for the business—involve qualitative assessments, are context-dependent and become less useful when separated from the environment in which they are collected, which make it hard to separate information collection and decision-making.

In Table A5 we examine—separately for low, medium and high numeracy loan officers—to what extent the risk classification of a borrower is related to observable characteristics of the borrower and his application. We find that there is no significant difference in the influence of observed application or borrower characteristics on the risk classification except for *Leverage*. This suggests that the higher accuracy of high numeracy loan officers is not primarily driven by a different interpretation of well observable, "hard" financial information.

In Table A6 we examine to what extent the risk classification of the loan officer helps predict loan arrears beyond the available hard financial information on the borrower. The degree to which this is the case provides us with an indicator of the value of the loan officer's assessment of soft, qualitative information about the borrower. Columns 1, 3 and 5 of Appendix I show that the explanatory power (R^2) of the simple regressions containing only the basic controls vary very little between the three numeracy groups. However, when adding the *Risky* indicator in columns 2, 4, and 6, the explanatory power is much higher in the regression for the high numeracy loan officers than for the medium and low numeracy loan officers. Results including the extended controls are qualitatively the same. This suggests that high numeracy loan officers are more accurate because they are better able to collect and assess the soft information that enters the rating decision.

Our estimates in Table 4 account for differences in average borrower characteristics between the pools of loans processed by high, medium and low numeracy loan officers. However, the loan portfolios may also differ with respect to the variation in observable "hard" characteristics across borrowers. The higher accuracy of high numeracy loan officers might therefore be partially explained by the fact that it is just easier for them to classify risky versus safe borrowers, because there is more variation in the pool of loans they process. Table A7 compares the distribution of observable borrower characteristics for the pool of loans processed by low, medium and high numeracy loan officers. We find that the standard deviation of some variables (*Time relationship*, *Leverage*, *Total assets/requested amount*) is indeed somewhat higher in the pool of loans processed by high numeracy loan officers. That said, the range of the distributions of all variables largely overlaps. Thus, our main results can hardly be explained by the fact that high numeracy loan officers have more variation to exploit in their loan portfolios.

TABLE 7

Heckman Selection Model

OLS Regression	**Total sample: 2007 July–2010 February**				**Precrisis: 2007 July–2008 September**			**Crisis: 2008 October–2010 February**		
Numeracy Level	**Low**	**Medium**		**High**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Dependent Variable: Arrears	**(1)**	**(2)**		**(3)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A: Second stage results										
Risky	0.059***	0.105***		0.180***	−0.107	−0.006	0.133**	0.075***	0.115***	0.182***
	(0.017)	(0.016)		(0.018)	(0.090)	(0.045)	(0.054)	(0.017)	(0.017)	(0.019)
Inverse Mill's ratio	−0.005	−0.036*		−0.047**	0.014	−0.043	−0.076	−0.014	−0.038	−0.045*
	(0.018)	(0.020)		0(.022)	(0.072)	(0.036)	(0.055)	(0.019)	(0.024)	(0.024)
Mean arrears	0.06	0.07		0.10	0.10	0.05	0.09	0.06	0.08	0.11
Observations	3,064	3,381		3,821	250	655	490	2,814	2,726	3,331
Wald test (chi2)	188.9***	189.7***		273.8***	34.3*	54.1***	43.6	185.3***	172.8***	262.2***
Basic controls	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer controls	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: First stage results										
New client	−0.605***	−0.668***	−0.643***	−0.909***	−0.730***	−0.658***	−0.575***		−0.651***	−0.643***
	(0.041)	(0.048)	(0.043)	(0.192)	(0.101)	(0.107)	(0.042)		(0.058)	(0.048)
Mean granted	0.67	0.60	0.58	0.62	0.62	0.62	0.68		0.60	0.57
Observations	3,064	3,381	3,821	250	655	490	2,814		2,726	3,331
Pseudo R-squared	0.395	0.373	0.354	0.486	0.320	0.318	0.395		0.394	0.366
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes
Extended controls	No	No	No	No	No	No	No		No	No
Loan officer controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes

Notes: This table shows results from the first stage of a two-step Heckman selection model. Effects are displayed as marginal effects at the mean. The dependent variable Granted is a binary variable equal to 1 if a loan application was granted and 0 otherwise. New client is the selection variable and is a dummy equal to 1 if the borrower has a bank account since <0.1 year and 0 otherwise. Basic controls include Ln(Requested amount), Request Euro, Time relationship. Loan officer controls include female, experienced, and age. FE, fixed effects. Standard errors in parentheses; *$p < .1$, **$p < .05$, ***$p < .01$.

Overall, our results point to high numeracy loan officers being better able to assess and verify "soft" qualitative information.

VI. CONCLUSION

We provide novel evidence documenting that employees with high numerical skills perform better on the job. In the context of small business lending we relate the numeracy of loan officers to the accuracy of their credit assessments. In line with findings from experimental studies, we document significant differences in accuracy between loan officers with low versus high numeracy. Initial ratings assigned by high numeracy loan officers are better able to predict which borrowers will default and which will not.

The difference in accuracy between high and low numeracy loan officers is most pronounced in the precrisis credit boom phase. This finding is in line with Becker, Bos, and Roszbach (2018) who show that it is most difficult to accurately sort borrowers according to their riskiness during boom periods in which informational frictions are highest. Our results thus provide evidence that hiring skilled loan officers is most important during boom times when separating borrowers by quality is most difficult. Our findings further show that higher numerical skills are a complement to other characteristics (gender, experience) that have been connected to improved loan performance in the literature.

APPENDIX A: EXAMPLE QUESTIONS FROM THE NUMERACY TEST

1. Calculate the value of the following expressions. [3.3]

(3 pts. for each correct answer)

$$\frac{\left(\frac{3}{4}+2\right)}{\left[\frac{2\bullet 3-2\bullet(-6)}{3}-7\right]}=$$

2. Calculate the original price if the current price of 88 EUR was obtained after the original price was first increased by 10% and then decreased by 4%. [4.15]

(4 pts. for the correct answer)

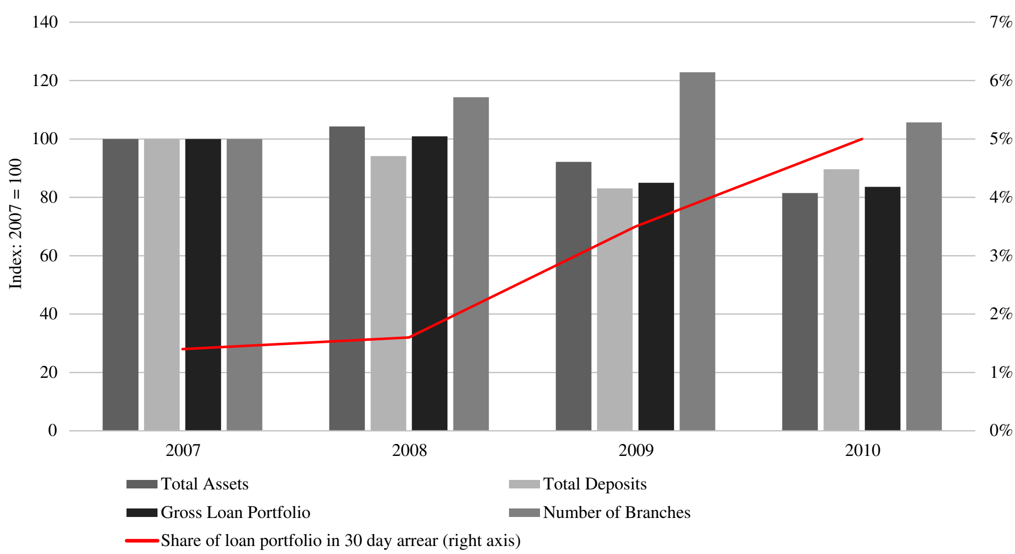
3. Six friends want to buy a piece of land, each paying an equal share. The day before the contract is signed two of the friends decide to withdraw their offer. The remaining four friends must therefore each increase their share by 4500 EUR in order to be able to pay the asking price. Calculate the price of the land. [6.4]

(5 pts. for the correct answer)

Notes: The three questions are taken from the bank's numeracy test. They are representative for the overall level of difficulty of the test.

FIGURE A1

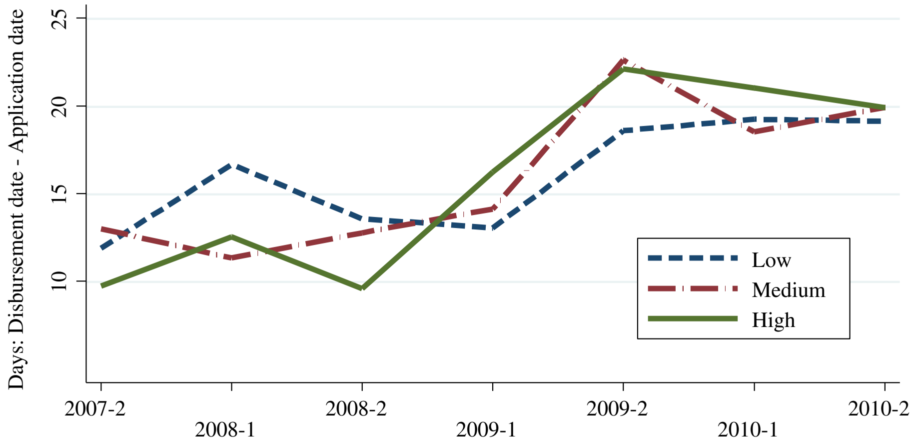
Development of the Bank.



Notes: This graph shows the development of the bank's total assets, total deposits, gross loan portfolio, branches and loan performance based on annual reports. The total assets, total deposits, gross loan portfolio and the number of branches are indexed at December 2007 = 100.

FIGURE A2

Processing Time of Loan Applications over Time.



Notes: The figure displays the average processing time of loan applications by half year and numeracy level. The processing time is defined as Disbursement date—Application date.

TABLE A1

Accuracy on Loan Level: Total Sample with Interaction Terms

OLS Regression	**Total Sample**		**Precrisis**		**Crisis**	
Dependent Variable: Arrears	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
High numeracy × Risky	0.112***	0.121***	0.187**	0.240**	0.111***	0.108***
	(0.038)	(0.039)	(0.077)	(0.094)	(0.040)	(0.041)
Medium numeracy × Risky	0.023	0.047	0.100**	0.104	0.021	0.041
	(0.038)	(0.033)	(0.049)	(0.066)	(0.040)	(0.034)
High numeracy	0.454***	0.469***	−1.425***	−0.497	0.427***	0.575***
	(0.138)	(0.158)	(0.444)	(2.023)	(0.141)	(0.167)
Medium numeracy	−0.006	0.210	−1.607***	−0.491	−0.070	0.178
	(0.141)	(0.140)	(0.362)	(1.991)	(0.148)	(0.185)
Risky	0.107***	0.060**	−0.075**	−0.108*	0.117***	0.075***
	(0.029)	(0.024)	(0.032)	(0.055)	(0.030)	(0.023)
Mean arrears	0.08	0.08	0.07	0.07	0.08	0.08
Observations	6,498	6,202	913	833	5,585	5,369
Adj. R-squared	0.073	0.079	0.029	0.042	0.086	0.088
Basic controls × Numeracy level	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls × Numeracy level	No	Yes	No	Yes	No	Yes
Loan officer controls × Numeracy level	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE × Numeracy level	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE × Numeracy level	No	Yes	No	Yes	No	Yes

Notes: The dependent variable Arrears is a binary variable equal to 1 if a firm went into 30 day payment arrear within the first 24 months of the loan. Basic controls include Ln(Requested amount), Request Euro, Time relationship, New client. Extended controls include Leverage, Ln(Sales), Young firm, Agriculture, Total assets/requested amount. Loan officer controls include Female, Experienced and Age, FE, fixed effects. Standard errors in parentheses; standard errors are clustered on loan officer level; $^{*}p<.1$, $^{**}p<.05$, $^{***}p<.01$.

TABLE A2

Accuracy on Loan Level: Linear Numeracy Score

OLS Regression	**Total Sample**		**Precrisis**		**Crisis**	
Dependent Variable: Arrears	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Transformed numeracy score × Risky	0.140**	0.152**	0.234**	0.362***	0.142**	0.136**
	(0.061)	(0.062)	(0.117)	(0.130)	(0.065)	(0.065)
Transformed numeracy score	0.484**	0.435**	−1.245	−1.608	0.444*	0.611**
	(0.225)	(0.211)	(1.450)	(1.717)	(0.236)	(0.254)
Risky	0.076**	0.038	−0.092	−0.183**	0.084**	0.055
	(0.038)	(0.037)	(0.059)	(0.070)	(0.040)	(0.037)
Mean arrears	0.08	0.08	0.07	0.07	0.08	0.08
Observations	6,498	6,202	913	833	5,585	5,369
Adj. *R*-squared	0.068	0.072	0.023	0.029	0.080	0.081
Basic controls × Transformed numeracy score	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls × Transformed numeracy score	No	Yes	No	Yes	No	Yes
Loan officer controls × Transformed numeracy score	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE × Transformed numeracy score	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE × Transformed numeracy score	No	Yes	No	Yes	No	Yes

Notes: The table displays results of the linear influence of the numeracy score. The numeracy score (values 0.65 – 1) is transformed so that the lowest value is 0 and the highest is 1: (numeracy score – 0.65)/0.35. Hence, the coefficients reflect the effect of moving from the lowest to the highest numeracy score. The dependent variable Arrears is a binary variable equal to 1 if a firm went into 30 day payment arrear within the first 24 months of the loan. Basic controls include Ln(Requested amount), Request Euro, Time relationship, New client. Extended controls include Leverage, Ln(Sales), Young firm, Agriculture, Total assets/requested amount. Loan officer controls include Female, Experienced and Age. FE, fixed effects. Standard errors in parentheses; standard errors are clustered on loan officer level; $*p < .1$, $**p < .05$, $***p < .01$.

TABLE A3

Accuracy on Loan Level: Sample until Test Date

OLS Regression	**Total Sample: 2007 July–2010 February**			**Precrisis: 2007 July–2008 September**			**Crisis: 2008 October–2010 February**		
Numeracy Level	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Dependent Variable: Arrears	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Risky	0.097*	0.169***	0.256***	−0.108*	−0.004	0.131	0.154***	0.208***	0.279***
	(0.050)	(0.044)	(0.044)	(0.059)	(0.036)	(0.078)	(0.052)	(0.051)	(0.050)
Mean arrears	0.06	0.08	0.11	0.09	0.05	0.09	0.05	0.09	0.11
Observations	1,067	1,218	1,303	151	389	293	916	829	1,010
Adj. *R*-squared	0.055	0.086	0.129	0.033	0.049	0.010	0.072	0.103	0.162
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Difference in coefficients of risky: *p* values of Chow test in parentheses									
Compared to high numeracy	−0.159**	−0.087		−0.239***	−0.135*		−0.125*	−0.071	
	(0.014)	(0.151)		(0.008)	(0.091)		(0.074)	(0.307)	
Compared to medium numeracy	−0.072		0.087	−0.104*		0.135*	−0.054		0.071
	(0.266)		(0.151)	(0.099)		(0.091)	(0.446)		(0.307)

Notes: The dependent variable Arrears is a binary variable equal to 1 if a firm went into 30 day payment arrear within the first 24 months of the loan. Basic controls include Ln(Requested amount), Request Euro, Time relationship, New client. Extended controls include Leverage, Ln(Sales), Young firm, Agriculture, Total assets/requested amount. Loan officer controls include Female, Experienced and Age. Standard errors in parentheses; standard errors are clustered on loan officer level; $*p < .1$, $**p < .05$, $***p < .01$. The sample includes only loans granted before the math test date in February 2010. We compare the coefficients of Risky by numeracy level using a Chow test.

TABLE A4

Accuracy on Loan Level: Only Loan Officers Who Were in Precrisis Sample

OLS regression	**Total Sample: 2007 July–2010 December**			**Precrisis: 2007 July–2008 September**			**Crisis: 2008 October–2010 December**		
Numeracy level	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Dependent Variable: Arrears	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Risky	0.056	0.117***	0.172***	−0.108*	−0.004	0.131	0.092*	0.128***	0.171***
	(0.044)	(0.026)	(0.040)	(0.059)	(0.036)	(0.078)	(0.043)	(0.028)	(0.046)
Mean arrears	0.06	0.07	0.10	0.10	0.05	0.09	0.06	0.08	0.11
Observations	652	1,584	1,622	151	389	293	501	1,195	1,329
Adj. R-squared	0.057	0.082	0.077	0.033	0.049	0.010	0.073	0.101	0.092
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan officer controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Difference in coefficients of Risky: p values of Chow test in parentheses									
Compared to high numeracy	−0.116**	−0.055		−0.239***	−0.135*		−0.079	−0.043	
	(0.043)	(0.242)		(0.008)	(0.091)		(0.194)	(0.420)	
Compared to medium numeracy	−0.061		0.055	−0.104*		0.135*	−0.036		0.043
	(0.213)		(0.242)	(0.099)		(0.091)	(0.461)		(0.420)

Notes : This table contains results for the subsample of loan officers who were already working at the bank in the precrisis period. The dependent variable Arrears is a binary variable equal to 1 if a firm went into 30 day payment arrear within the first 24 months of the loan. Basic controls include Ln(Requested amount), Request Euro, Time relationship, New client. Extended controls include Leverage, ln(Sales), Young firm, Agriculture, Total assets/requested amount. Loan officer controls include Female, Experienced and Age. Standard errors in parentheses; standard errors are clustered on loan officer level; ; $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$. We compare the coefficients of Risky by numeracy level using a Chow test.

TABLE A5

Influence of Loan Characteristics on Risky

OLS Regression	**Total Sample: 2007 Jul–2010 Dec**			**Difference in coefficients p values of Chow test in parentheses**		
Dependent Variable: Risky	**Low (1)**	**Medium (2)**	**High (3)**	**Low vs medium**	**Low vs high**	**Medium vs high**
Ln(Requested amount)	−0.010	−0.001	−0.022	−0.009	0.012	0.021
	(0.014)	(0.017)	(0.015)	(0.665)	(0.539)	(0.334)
Request Euro	0.316***	0.309***	0.262***	0.007	0.054	0.047
	(0.075)	(0.047)	(0.043)	(0.939)	(0.527)	(0.458)
Time relationship	0.003	0.005	0.002	−0.002	0.001	0.003
	(0.004)	(0.007)	(0.006)	(0.780)	(0.957)	(0.774)
Leverage	0.022**	0.047***	0.046***	−0.025***	−0.024***	0.001**
	(0.010)	(0.013)	(0.011)	(0.000)	(0.002)	(0.013)
Ln(Sales)	0.029	0.012	0.047	0.017	−0.018	−0.035
	(0.032)	(0.027)	(0.030)	(0.142)	(0.110)	(0.958)
Young firm	−0.131***	−0.125***	−0.083***	−0.006	−0.048	−0.042
	(0.031)	(0.042)	(0.030)	(0.682)	(0.672)	(0.375)
Agriculture	0.001	0.002	0.002***	−0.001	−0.001	0.000
	(0.001)	(0.002)	(0.001)	(0.907)	(0.258)	(0.403)
Total assets/requested amount	0.001	0.002	0.002***	−0.001	−0.001	0.000
	(0.001)	(0.002)	(0.001)	(0.645)	(0.593)	(0.929)
Mean risky	0.19	0.25	0.26			
Observations	2,040	1,997	2,165			
Adj. R-squared	0.472	0.330	0.396			
Loan officer controls	Yes	Yes	Yes			
Branch FE	Yes	Yes	Yes			
Quarter FE	Yes	Yes	Yes			

Notes: This table displays results of a linear probability model estimation. The dependent variable Risky is a binary variable equal to 1 if a loan was classified as risky at loan disbursement. Basic controls include Ln(Requested amount), Request Euro, Time relationship, New client. Extended controls include Leverage, ln(Sales), Young firm, Agriculture, Total assets/requested amount. Loan officer controls include Female, Experienced and Age. Standard errors in parentheses; standard errors are clustered on loan officer level; ; $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

TABLE A6

Predictive Power of Hard Information

	Basic controls						Basic and Extended controls					
Numeracy level	**Low**		**Medium**		**High**		**Low**		**Medium**		**High**	
Dependent Variable: Arrear	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Risky		0.111***		0.136***		0.220***		0.042*		0.101***		0.170***
		(0.027)		(0.026)		(0.025)		(0.023)		(0.025)		(0.028)
Observations	2,119	2,119	2,098	2,098	2,281	2,281	2,040	2,040	1,997	1,997	2,165	2,165
Adj. R-squared	0.003	0.026	0.000	0.039	0.000	0.077	0.033	0.036	0.021	0.041	0.024	0.068
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table displays the predictive power of application and firm variables for the outcome variable Arrears. The dependent variable Arrears is a binary variable equal to 1 if a firm went into 30 day payment arrear within the first 24 months of the loan. Basic controls include Ln(Requested amount), Request Euro, Time relationship, New client. Extended controls include Leverage, Ln(Sales), Young firm, Agriculture, Total assets/requested amount. Standard errors in parentheses; standard errors are clustered on loan officer level; $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

TABLE A7

Distribution of Firm Characteristics

Variable	**Numeracy**	**Mean**	***SD***	**p10**	**p25**	**p50**	**p75**	**p90**
Ln(Requested amount)	Low	8.14	0.95	6.88	7.54	8.17	8.85	9.36
	Medium	8.45	0.89	7.28	7.80	8.46	9.15	9.55
	High	8.42	0.95	7.14	7.78	8.45	9.14	9.60
Time relationship	Low	1.88	1.78	0.00	0.00	1.74	3.15	4.44
	Medium	1.85	1.92	0.00	0.00	1.53	3.13	4.77
	High	1.85	1.82	0.00	0.00	1.62	3.14	4.51
Leverage	Low	0.99	2.92	0.15	0.24	0.48	0.96	1.99
	Medium	1.72	18.55	0.17	0.31	0.57	1.16	2.33
	High	1.36	6.79	0.16	0.28	0.58	1.25	2.24
Ln(Sales)	Low	7.02	1.44	5.36	6.01	6.77	7.94	8.96
	Medium	7.60	1.37	6.00	6.59	7.37	8.64	9.49
	High	7.43	1.44	5.78	6.40	7.13	8.49	9.50
Total assets/requested amount	Low	5.13	6.49	0.93	1.69	3.20	5.86	11.25
	Medium	5.23	11.64	0.83	1.58	3.06	5.79	10.43
	High	5.85	12.45	0.76	1.55	3.24	6.40	12.22

REFERENCES

Agarwal, S., and I. Ben-David. "Loan Prospecting and the Loss of Soft Information." *Journal of Financial Economics*, 129(3), 2018, 608–28.

Agarwal, S., and B. Mazumder. "Cognitive Skills and Household Financial Decision Making." *American Economic Journal: Applied Economics*, 5(1), 2013, 193–207.

Andersson, P. "Does Experience Matter in Lending? A Process-Tracing Study on Experienced Loan Officers' and Novices' Decision Behavior." *Journal of Economic Psychology*, 25(4), 2004, 471–92.

Beck, T., P. Behr, and A. Guettler. "Gender and Banking: Are Women Better Loan Officers?" *Review of Finance*, 17(4), 2013, 1279–321.

Beck, T., H. Degryse, R. deHaas, and N. vanHoren. "When Arm's Length is Too Far. Relationship Banking over the Credit Cycle." *Journal of Financial Economics*, 127, 2018, 174–96.

Becker, B., M. Bos, and K. Roszbach. "Bad Times, Good Credit." Swedish House of Finance Research Paper, Issue 15-05, 2018.

Berg, T. "Playing the Devil's Advocate: The Causal Effect of Risk Management on Loan Quality." *Review of Financial Studies*, 28(12), 2015, 3367–406.

Berger, A. N., and G. F. Udell. "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business*, 68, 1995, 351–81.

———. "The Institutional Memory Hypothesis and the Procyclicality of Bank Lending Behavior." *Journal of Financial Intermediation*, 13(4), 2004, 458–95.

Bank for International Settlements. "Studies on the Validation of Internal Rating Systems." Working Paper No. 14, 2005.

Brown, M., K. Kirschenmann, and S. Ongena. "Bank Funding, Securitization, and Loan Terms: Evidence from Foreign Currency Lending." *Journal of Money, Credit and Banking*, 46(7), 2014, 1501–34.

Brown, M., M. Schaller, S. Westerfeld, and M. Heusler. "Internal Control and Strategic Communication within Firms—Evidence from Bank Lending." University of St. Gallen, School of Finance Research Paper No. 2015/04.

Bruns, V., D. V. Holland, D. A. Shepherd, and J. Wiklund. "The Role of Human Capital in Loan Officers' Decision Policies." *Entrepreneurship Theory and Practice*, 32(3), 2008, 485–506.

Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the*

National Academy of Sciences of the United States of America, 106(19), 2009, 7745–50.
Chevalier, J., and G. Ellison. "Are some Mutual Fund Managers Better than Others? Cross-Sectional Patterns in Behavior and Performance." *Journal of Finance*, 54(3), 1999, 875–99.
Cole, S., M. Kanz, and L. Klapper. "Incentivizing Calculated Risk-Taking: Evidence from an Experiment with Commercial Bank Loan Officers." *Journal of Finance*, 70(2), 2015, 537–75.
Custodio, C., and D. Metzger. "Financial Expert CEOs: CEO's Work Experience and Firm's Financial Policies." *Journal of Financial Economics*, 114(1), 2014, 125–54.
Dell'Ariccia, G., D. Igan, and L. Laeven. "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market." *Journal of Money, Credit and Banking*, 44(2–3), 2012, 367–84.
Dell'Ariccia, G., and R. Marquez. "Lending Booms and Lending Standards." *Journal of Finance*, 61(5), 2006, 2511–46.
Diamond, D. W. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies*, 51(3), 1984, 393–414.
Engelmann, B., E. Hayden, and D. Tasche. "Measuring the Discriminative Power of Rating Systems." Discussion Paper 01/2003, Deutsche Bundesbank Research Center, 2003.
Fang, J., A. Kempf, and M. Trapp. "Fund Manager Allocation." *Journal of Financial Economics*, 111(3), 2014, 661–74.
Frederick, S. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives*, 19(4), 2005, 25–42.
Gerardi, K., L. Goette, and S. Meier. "Numerical Ability Predicts Mortgage Default." *PNAS*, 110(28), 2013, 11267–71.
Ghazal, S., E. T. Cokely, and R. Garcia-Retamero. "Predicting Biases in Very Highly Educated Samples: Numeracy and Metacognition." *Judgment and Decision making*, 9(1), 2014, 15–34.
Ginsburg, L., M. Manly, and M. J. Schmitt. "The Components of Numeracy." NCSALL Occasional Paper. National Center for the Study of Adult Learning and Literacy (NCSALL), 2006.
Grinblatt, M., S. Ikaheimo, M. Keloharju, and S. Knupfer. "IQ and Mutual Fund Choice." *Management Science*, 62(4), 2015, 924–44.
Grinblatt, M., M. Keloharju, and J. Linnainmaa. "IQ, Trading Behavior and Performance." *Journal of Financial Economics*, 104(2), 2012, 339–62.
Hertzberg, A., J. M. Liberti, and D. Paravisini. "Information and Incentives inside the Firm: Evidence from Loan Officer Rotation." *Journal of Finance*, 65(3), 2010, 795–828.
Joensen, J. S., and H. S. Nielsen. "Is There a Causal Effect of High School Math on Labor Market Outcomes?" *Journal of Human Resources*, 44(1), 2009, 171–98.
Kaplan, S. N., M. M. Klebanov, and M. Sorensen. "Which CEO Characteristics and Abilities Matter?" *Journal of Finance*, 67(3), 2012, 973–1007.
Kirschenmann, K. "Credit Rationing in Small Firm-Bank Relationships." *Journal of Financial Intermediation*, 26, 2016, 68–99.
Koedel, C., and E. Tyhurst. "Math Skills and Labor-Market Outcomes: Evidence from a Resume-Based Field Experiment." *Economics of Education Review*, 31, 2012, 131–41.
Li, H., X. Zhang, and R. Zhao. "Investing in Talents: Manager Characteristics and Hedge Fund Performances." *Journal of Financial and Quantitative Analysis*, 46(1), 2011, 59–82.
Liberti, J. M., and A. R. Mian. "Estimating the Effect of Hierarchies on Information Use." *Review of Financial Studies*, 22(10), 2009, 4057–90.
Liberti, J. M., and M. A. Petersen. "Information: Hard and Soft." *Review of Corporate Finance Studies*, 2017, 1–42.
Lusardi, A., and P. Tufano. "Debt Literacy, Financial Experiences, and Overindebtedness." *Journal of Pension Economics and Finance*, 14(04), 2015, 332–68.
Moody's Investor Services. "Measuring the Performance of Corporate Bond Ratings." Special Comment, 2003.
Morales-Acevedo, P., and S. Ongena. "Fear, Anger and Credit. On Bank Robberies and Loan Conditions." *Economic Inquiry*, 2019. Forthcoming.
Pachur, T., and M. Galesic. "Strategy Selection in Risky Choice: The Impact of Numeracy, Affect, and Cross-Cultural Differences." *Journal of Behavioral Decision Making*, 26(3), 2013, 260–71.
Peters, E., D. Västfjäll, P. Slovic, C. K. Mertz, K. Mazzocco, and S. Dickert. "Numeracy and Decision Making." *Psychological Science*, 17(5), 2006, 407–13.
Qian, J., P. Strahan, and Z. Yang. "The Impact of Incentives and Communication Costs on Information Production and Use: Evidence from Bank Lending." *Journal of Finance*, 70(4), 2015, 1457–93.
Reyna, V. F., W. L. Nelson, P. K. Han, and N. F. Dieckmann. "How Numeracy Influences Risk Comprehension and Medical Decision Making." *Psychological Bulletin*, 135(6), 2009, 943–73.