DISCUSSION PAPER SERIES

DP14500

GENDER AND CREDIT RISK: A VIEW FROM THE LOAN OFFICER'S DESK

Marta Reynal-Querol and Jose G Montalvo

FINANCIAL ECONOMICS

ISSN 0265-8003

GENDER AND CREDIT RISK: A VIEW FROM THE LOAN OFFICER'S DESK

Marta Reynal-Querol and Jose G Montalvo

Discussion Paper DP14500
Published 18 March 2020
Submitted 15 March 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

GENDER AND CREDIT RISK: A VIEW FROM THE LOAN OFFICER'S DESK

Abstract

In this paper we analyze the effect of loan officers' gender on the approval of loans and, in particular, on their subsequent performance. Using detailed bank information on a sample of close to half a million loans, we show that female loan officers have, conditional on the risk score, around a 15\% lower delinquency rate than that of male officers. In addition to the original scoring of the loans, we also have the recommendation of the expert system. We find that the risk profile of applicants screened by male and female loan officers is very similar, but conditional on risk score, women follow the recommendations more often than men. Moreover, we find evidence of gender bias in terms of a mistake-punishment trade-off, which could explain, at least in part, women's higher compliance with the recommendations. Indeed, there is a double standard in terms of the consequences for breaking the rules: errors, in the form of delinquent loans as a result of not following the recommendation of the system, are forgiven more often for male than for female loan officers.

JEL Classification: G21, G32, J16

Keywords: N/A

Marta Reynal-Querol - marta.reynal@upf.edu
Universitat Pompeu Fabra and CEPR

Jose G Montalvo - jose.garcia-montalvo@upf.edu
Universitat Pompeu Fabra

Gender Differences in Compliance: Evidence from Loan Officers

Jose G. Montalvo and Marta Reynal-Querol*

This version: February 2020

*Montalvo: UPF - BarcelonaGSE - IPEG - ICREA-A, Ramon Trias Fargas 25-27,08005 Barcelona, jose.garcia-montalvo@upf.edu. Reynal-Querol: ICREA-UPF-BarcelonaGSE-IPEG, Ramon Trias Fargas 25-27,08005 Barcelona, marta.reynal@upf.edu. We appreciate the comments of Jose Luis Peydro, Fernando Broner, Ruben Durante, and participants at many seminar and conferences. We acknowledge the excellent research assistance provided by Ece Yagman. We are especially grateful to the financial companies that shared with us the data for this study, and to its members who kindly answered our questions about the data and provided qualitative insights. Financial support from the European Research Council under the European Community ERC Grant Agreement n.647514, the Spanish National Science Foundation AEI/FEDER, UE ECO2017-82696-P, and the Government of Catalonia (ICREA) is gratefully acknowledged. We are also thankful for the financial support provided by the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563).

Abstract

We analyze the effect of loan officers' gender on the approval of loans and their subsequent performance. Using information on close to half a million loans we show that female loan officers have around a 15% lower delinquency rate than male officers despite the fact that the risk profile of applicants screened by male and female loan officers is very similar. In addition to the original scoring of the loans, we also have the recommendation based on specific policies of the bank. Using this information we show that, conditional on risk score, women show a higher degree of compliance with the recommendations than males, which explains their lower delinquency rates. Moreover, we find evidence of a double standard in terms of the consequences for breaking the rules: errors have more negative consequences for females than for males. (JEL: G21, G32, J16)

1 Introduction

There are many theories on the determinants of the financial crisis of 2008, but all share the common idea of a deficient risk management system. The credit and housing bubble were the result of excessive risk-taking, which was ultimately the cause of the crisis. In this paper we study the management of credit risk during the period that led up to the financial crisis. We approach this issue from the perspective of the gender of loan officers. Indeed, the effect of gender on the management of credit risk remains understudied. We examine the lending decisions of more than one thousand loan officers on close to half a million applications for mortgages and consumer loans to several Spanish financial institutions during the period 2000-2013.

Our data and setting provide four basic advantages. First, we have access to information on the individual characteristics of each loan and the officer who screened it. Second, and in contrast to most of the literature on bank loans, our data also include the original risk score of each loan. The models to generate the risk score included many variables, available at the time of screening the loan, related with the characteristics of the applicants and the type of product. These models were critical for the financial institutions since they used IRB (Internal rating-based approach) to estimate their capital requirements. The statistical models had to be approved by the banking supervisor.

Third, we also have the original recommendation based on the credit policies of the bank. This is different from the risk score. For instance a credit policy could be to avoid originating mortgages with a very high loan-to-value ratio. This means that two loans with the same risk score can have different recommendations depending on the specific value taken by the policy variables. For example, a mortgage with a loan-to-value ratio over 90% could get a negative recommendation while, depending on the value of the other variables used in the scoring model, the risk score could be good.

Finally, the data contains a rich set of indicators on the performance of each loan, ranging from their objective payment situation at each point in time (without incident, in arrears for more than 30, 60, or 90 days, etc.) to their legal/accounting status (without incident, in legal litigation, proposed for writing down, wrote down, condoned, etc.).

We reach three basic conclusions. The first set of results concerns the differential delinquency rates of male versus female loan officers. Conditional on risk scores, female loan officers have a 15% lower delinquency rate than male officers. This result is robust to the use of alternative measures of delinquency, various scoring models, different types of loans, or adding other characteristics of the loan officers.

The second set of findings show that, conditional on the risk scoring, women follow the recommendations generated by the credit policies of the bank more often than men. They also less frequently apply exceptional circumstances to overrule the recommendation of the system compared to men. Our data is particularly well-suited to analyzing this issue as we have the recommendation after the original screening process. This finding is consistent with research in other fields (drivers' compliance with traffic regulations, pedestrian behavior, etc.) but, as far as we know, there are no previous examples in the economic literature. Third, and finally, we show that one potential explanation for the higher degree of compliance of women versus men is related to gender bias in the "mistake-punishment trade-off": women's errors, and hence their careers, are more severely penalized conditional on their record of loan performance.

Our results contribute to four strands of the literature. First, there is a long tradition of study on gender discrimination across different contexts and, particularly, in the labor market.

There is also an extended literature on behavioral gender differences with respect to risk. Previous research has shown that women are more risk averse than men (Byrnes et al., 1999; Croson & Gneezy, 2009; C. C. Eckel & Grossman, 2008).[1] The differential responses of men and women can also affect the management of credit risk. We consider an alternative channel for the observed differences in the management of credit risk: women show a higher level of compliance with regulations than do men. Research on the differences between women and men in terms of their respective compliance with rules is scarce, and mostly concentrated around compliance with traffic regulations. We contribute to the literature by showing the differential degree of compliance of men and women in an economic environment.

Third, our results contribute to a very recent economic literature on a different source of discrimination: the possibility that gender influences the way information about others is interpreted. For instance, Sarsons (2019) shows that there is an asymmetric response to mistakes made by surgeons depending on their gender. This implies that the drop in referrals after a bad outcome is much larger for women than men, in turn reducing the possibilities of promotion and higher salaries for female versus male surgeons.

[1] Niederle (2016) argues that the experimental evidence on gender differences in risk aversion is less clear than that on competition, and that there is substantial heterogeneity in results across experimental set-ups and elicitation methods. In the same spirit, Filippin & Crosetto (2016) conclude that gender differences in risk are less frequently found in the literature than usually depicted, and depend largely on the elicitation method.

This theory of differential punishment based on gender has also been analyzed in the context of the financial industry. Egan et al. (2017) show that, after an incident of misconduct, female financial advisors are more likely to lose their job, and spend more time searching for a new one, than are men. Using our data, we find that women who accumulate a high proportion of non-performing loans, which is more likely if they do not follow the recommendations, have a greater probability of being punished than men, conditional on the same level of performance. This double standard helps to explain, from a rational perspective, the higher level of compliance of women with the recommendations of the system.

Finally, our results are related with the findings of Beck et al. (2013), which report a lower likelihood of arrears for loans screened by female loan officers than for those screened by male loan officers. However, the context of the decision is very different. In Beck et al. (2013) the screening of borrowers is performed in an ad-hoc fashion by each officer, while in our case their is a explicit scoring and recommendation for each loan application, that is available for the officer before making a decision. Therefore, we can control explicitly by the risk and recommendation of the system for each application. This set up allow us to test a new explanation for this finding.

The remainder of the article is organized as follows. Section 2 presents an overview of the literature on the relationship between gender and risk. Section 3 discusses the data. We then describe our basic results in Section 4. Section 5 provides a large set of robustness exercises. In Section 6, we explore explanations for the basic findings, including an examination of the gendered double standard relative to punishment-mistakes. Section 7 concludes.

2 Gender and risk

There is a broad literature on gender differences in risk attitudes and the evaluation of risk, where a variety of explanations have been proposed for divergences in risk-taking.[2]. Most of these empirical papers analyze gender differences in the context of market risk. In the banking industry, the bulk of the risk is, however, concentrated around credit risk. In addition, much of the literature on gender differences relative to credit risk focuses on borrowers' gender. Our paper analyzes the influence of gender on the understudied overlap between credit risk and lender behavior.

2.1 Gender and market risk

Much of the empirical literature on the differential risk attitudes of men and women has centered around financial markets and, therefore, market risk. There are somewhat mixed evidence on gender differences in bubble formation suggest that results should be interpreted as context dependent (C. C. Eckel & Füllbrunn (2015) (C. C. Eckel & Füllbrunn, 2017) and Cueva & Rustichini (2015)). Gender differences also produce varying investment styles. Barber & Odean (2001) find that men trade 45% more than women

[2] Croson & Gneezy (2009) summarize this literature.

and, therefore, produce net returns below those of women. Portfolios managed by women have less risky assets and less propensity to engage in extreme investment strategies. Several studies show that women have less tolerance for financial risk than men (Barsky et al., 1997; Olsen & Cox, 2001; Hallahan et al., 2004; Neelakantan, 2010). This difference in risk preferences can lead to men and women adopting different financial strategies, where the latter might be less willing to employ a wider range of strategies with greater variance (Powell & Ansic, 1997).[3]

2.2 Gender and credit risk

Empirical research on gender and risk has concentrated on the management of market risk, as described above. However, the financial crisis of 2008 was mostly associated with a credit bubble that fed a housing bubble through excessive mortgage lending. In the context of the banking industry, it is therefore interesting to characterize gender differences, if any, in the management of credit risk.[4] Such an analysis is particularly important as the recent financial crisis was associated with a banking crisis, which tend to be deeper and more prolonged than other types of crises (Reinhart & Rogoff, 2009).

[3]A caveat is the type of decision frame used in the experiment, as highlighted by Schubert et al. (1999). Results suggest that heterogeneity in risk preferences between males and females arise only in abstract gambles but not in contextual decisions.

[4]There are four basic risk categories that affect banks' profitability and solvency: rate risk, market risk, credit risk, and operational risk. The most significant source of risk in the banking industry as a whole is credit risk.

2.2.1 Gender from the borrowers' perspective

In this paper we study gender differences in credit risk management. There is a long literature documenting the effect of borrowers' gender on approval probability, interest rates and delinquency rates. The microcredit literature argues, for example, that the fact that such credit targets mostly women explains, at least partly, the success of these programs in developing economies (Pitt & Khandker (1998), Khandker (2005)).

Using European Central Bank survey data, Stefani & Vacca (2013) find that the expectation of rejection leads women to apply less frequently for bank loans. Ongega & Popov (2016) and Treichel & Scott (2006) find similar results. Andres et al. (2019), Robb & Wolken (2002) and Moro et al. (2017) find that female entrepreneurs are less likely to ask for a loan than their male counterparts. However, many papers (Ongega & Popov (2016), Moro et al. (2017), Asiedu et al. (2012) and Treichel & Scott (2006)) do not find gender differences in loan denial. In particular, Blanchflower et al. (2003) use data from the Survey of Small Business Finances (SSBF) in the United States and find no difference in loan denial rates by gender. Structural differences (e.g., size of business, age, sector, etc.) between firms owned by males and females explain the difference in rejection rates.

Several papers show that female entrepreneurs are charged higher interest rate than males (Alesina et al. (2013), Muravyev et al. (2009), Mascia & Rossi (2017)). By contrast Ongega & Popov (2016) do not find any difference in interest rate or loan characteristics.

2.2.2 Gender from the lenders' perspective

This paper considers the role of women as lenders, not as borrowers. In contrast to the wealth of studies on the effect of the gender of the borrower, relatively little work has been carried out on the effect of the gender of the loan officer on the performance of loans. We analyze the determinants of loans' delinquency rates, focusing particularly on differences due to the gender the loan officer screening the application.

The novelty of our data is the fact that loan officers make their decision to grant or deny a loan knowing both the outcome of the scoring process and the recommendation based on the specific policies of the bank. The conditions under which the decision is made thus reduce the complexity of the choice, and generate a clear set-up for the analysis. All the officers have the same information delivered by a common internal risk scoring model and a common set of criteria to determine the recommendation. Conditional on the characteristics of the client and the product, the scoring model produces the same score for any loan officer working at the branches of the bank. The decision faced is therefore similar: the loan officers have the same hard information, synthesized in a risk score rate. They also receive a recommendation, and they must choose whether to follow the recommendation or claim an exception. As we noticed before, two loans with the same risk score can have different recommendations depending on the specific credit policies of the bank.

Our objective is to determine whether there is empirical evidence to support any difference in the performance of loans as a function of the gender of the loan officer making the decision and explain any such variance. Surprisingly, only a handful of studies have examined the impact of loan officers' gender on the screening and outcomes of loan granting. Most of these papers highlight qualitative differences in the screening criteria and processes used by male and female loan officers (Carter et al., 2007; Agier & Szafarz, 2013; Bellucci et al., 2010), especially in the context of lending to businesses.[5] Beck et al. (2013) analyze a data set from a commercial bank in Tirana (Albania) over the period 1996-2006. In most of the exercises, they consider a sample of 6,775 small loans mostly for small and medium size firms (SMEs). The authors conclude that female loan officers have a lower likelihood of granting a problematic loan than male officers. Doering (2018) includes the gender of the loan officer as a control variable to account for the possibility that clients may be less compliant with female officers, based on sociological research. In the context of microcredits Doering (2018) finds that female loan officers' have more missed payments on their loans than males.

In general, loan officers use hard and soft information (Liberti & Mian, 2009; Rajan et al., 2015), garnered from personal interaction. Soft information is particularly important in the context of loans to SMEs where there is no formal scoring process, or the firm's reliability is low. Such cases have

[5]Our study does not consider lending to business or entrepreneurs because it is well known that scoring models for these categories are not very reliable. Moreover, in the data, there was no scoring model for these types of loans, such that all applications from businesses were handled by a loan specialist, or committees of several officers, at central services.

been explored in the literature discussed above. In determining the appropriate choice, women are both more sensitive to social cues and more responsive than men to the specific conditions of the experimental setting (Kahn et al., 1971; Croson & Gneezy, 2009).

In our case, the risk scoring and the recommendation of the system were very salient, leaving much less room for a relevant role of soft information. After introducing all the data of the applicant[6], the loan officer's screen showed information on the risk score of the application, and the recommendation derived from the policies of the bank. The decision of the loan officer is thus mostly based on the risk scoring and the recommendation.[7]

It is thus important to assess the relevance of compliance with the rules relative to the differential results in the lending decisions of male versus female loan officers. Psychological and traffic research shows that women follow the rules more often than men. In our study, this is similarly true, conditional on credit scores. We consequently analyze the incentives of men and

[6]Recent research (Berg et al., 2019) argue that IRB systems, based only on hard information, may generate an incentive on loan officer to alter the information until a positive recommendation is obtained. In their case the decisions of the system could not be override. In our case, as we will show, loan officers can override the recommendations of the system which reduce their incentives to alter the information. As in the case of Berg et al. (2019) we have all the information on scoring trials (any change in the original typing of the information in the application). In fact, in our case the system saved all the keys typed by the loan officer. Also differently from Berg et al. (2019), the loan officer in our sample were aware of the fact that their actions on the keyboard were saved for further inspection. For these reasons we do not observe a significant amount of scoring trial.

[7]Managers and loan officers reported, in personal interviews, that the risk score and the recommendation produced by the application of the policies of the bank were considered fundamental information in their screening process.

women in an effort to explain this phenomenon. Conversations with managers and loan officers indicated that women perceived a gender bias in terms of a mistake-punishment trade-off: women's errors prompted harsher consequences than those of men. Versions of this potential explanation have recently found some academic support. Several papers have shown that women who break the rules are punished more often than men.[8] This creates an incentive that could generate a gender difference in the decision to grant or deny a loan given a recommendation. We also show that, given a specific risk score, overturning recommendations generated a higher delinquency rate than following the suggested course of action.

We revisit the relationship between gender and credit risk management using a large data set of individual loans from several Spanish financial institutions. Spain suffered a large credit and housing bubble in the years leading up to the financial crisis of 2008, providing an appropriate context for the analysis of the management of credit risk. The problems of Spanish banks were not very different from the problems of banks in many other countries. The rate of non-performing loans increased quite substantially in most of the countries, and many banks were under IRB, which implied that they had to have formal models to generate the score of the loans and the corresponding expected rate of default.

We had access to loan level administrative data from more than 400,000

[8] See Egan et al. (2017) and Sarsons (2019). We discuss this possibility in the last section of the paper.

loans applications to several financial institutions that merged to create a large bank.[9] The data includes mortgages and consumption loans, ranging from low to high amounts. A distinct advantage of this data set, as already mentioned, is the fact that it contains the internal scoring used to screen each loan application as well as the recommendation produced by the application of the credit policies of the bank. The data set also includes numerous financial variables on the applicants and the loans, many of which were used in the scoring model. Finally, we also had access to demographic information on the officer who approved each loan, as well as all the Internal Circulars issued by central services to the various branches. The memos contain all the policies regarding risk management, pricing, etc. as well as changes made to these policies over time.

3 Data

3.1 Characteristics of the data set

The global economy suffered a large shock as a consequence of the financial crisis of 2008. The effects were felt especially in the banking industry, where the default rate of loans increased rapidly. The Spanish banking sector was not an exception. Figure 1 shows the fast increase in the rate of non-performing loans in the Spanish banking sector after 2008. The peak of the rate was much higher than the maximum observed in the previous

[9] We hereon refer to ”the bank” or to financial institutions indifferently.

banking crisis of 1992-95. This study is based on a unique and very detailed database that contains more than 400.000 applications for loans to a large Spanish bank during the period 2000-2013. We had access to several data sets. The first one, the validation database, contained all the variables needed for the construction of the internal scoring model. It includes many financial variables[10] considered at the time of the original screening of the operation, and all the characteristics of the applicants that were recognized as potentially relevant, or predictive, in the scoring model. The scoring model was quite sophisticated and most probably[11] included not only demographic, financial, and personal characteristics of the applicants (age, marital status, occupation, type of contract, indebtedness, etc.) but also variables related to the relationship between the client and the bank, transactionality (number of accounts, length of the commercial relationship, average amount held in the account during the last year, etc.) and the type of product (mortgage/consumer loan, loan to value if mortgage, etc.).

The second data set include a variety of performance measures used by the bank to validate the scoring model. Obviously, the validation model must be confronted with the performance of the loans. The performance database is, by its structure, quite different from the validation data set. The latter captures a still picture at the time of approval of each loan, while

[10]This information came from the main data set of the bank, which included all financial information on the accounts and products of the bank used to produce financial statements, regulatory reports, etc.

[11]We do not know the exact model that was used to calculate the risk scores.

the performance database includes the accumulated performance since the approval. For instance, it contains, among many other variables, indicators describing whether there were any late payments[12] of more than 30, 60, or 90 days since the origination of the loan.[13]

A third data set provided information on the loan officers. It included not only some demographic characteristics of the officer (i.e., gender, age) but also the duration of their tenure in the position, and the branch at which they were working. A fourth data set covered all the characteristics of the loans that were approved: maturity, amount, type, purpose, etc. This information also came from the bank's main financial data set.

The database resulting from merging these four data sets is not only very detailed but includes information that makes it unique. First, it contains data not only on loans granted but also applications denied. While the information comes from several financial institutions, the high number of loans and the length of time over which the data are available provide confidence on the external validity of the results.[14] Second, the database includes the risk scoring as well as the recommendation generated by the application of the credit policies of the bank, one of the main novelties of the analysis presented in this paper. Generally, researchers working with administrative

[12] Including the number of late payments during the life of the loan.

[13] We had also access to data sets with the temporal evolution of these performance indicators. In the last section of the paper we use this information to construct the known evolution of the performance of each loan officer at each point in time.

[14] The use of detailed banking information from one, or a few, loan providers instead of the whole sector is not uncommon in the recent literature. See for instance Campbell & Cocco (2015) and Rajan et al. (2015).

data on individual loans rarely have access to the internal scoring of the loans. Consequently, some papers use the interest rate as a proxy of the quality of the loan. In the Spanish case, the interest rate would be a questionable indicator of the quality of a loan since, in general, the interest rate is set independently of mortgage characteristics (scoring, LTV, etc.) (Mayordomo et al. (2019)) and banks ration credit through quantities instead of prices (Bentolila et al. (2017)).[15] The bank analyzed here provides a clear example: the interest rate was only a function of buying other products of the bank together with the loan. The Internal Circulars of the bank state that the standard common rate could be reduced by 0.1 points for subscribing to life insurance; 0.05 additional points for buying home insurance; 0.1 points for getting a credit card; and 0.1 points for direct payment of paychecks to the account of the bank. No reference is made to any influence of the scoring on the interest rate. This is, moreover, a general feature of Spanish banks: interest rates are insensitive to mortgage characteristics (risk scores, LTV, etc.) in the segment of retail banking clients.[16]

Depending on the size of the requested loan, the decision was either made at that branch or was elevated to a specialized committee in the bank's central services. During the analyzed period, most households' applications for loans were initiated at a branch of the bank. One basic operating principle was the

[15]In any case, this is not very relevant in our case since we do not need to proxy the quality of the loan using the interest rate given that we know the original score.

[16]This fact simplifies the calculation of the risk-adjusted return on capital (RAROC) of each individual loan.

delegation of the ability to authorize different types of loans. The Internal Circulars issued by central services to the branches[17] confirm that during the period of study the loan officers at the branches could approve mortgages up to 350,000 euros.[18] We do not consider applications that requested amounts above the limits of concession at the branches, which were sent to the bank-wide committee.[19]

A second novelty of the study is the analysis of the final decision made on a given application. The loan officers knew the recommendation before making their decision, although they could "exceptionally" overrule this recommendation. The Internal Circulars state that when the recommendation system provided a favorable recommendation (positive or very positive), the loan could automatically be granted. If the recommendation was unfavorable (negative or very negative), then the operation should be denied. However, the Internal Circulars add that "in exceptional cases the officer can ultimately approve the loan, explaining why she disagrees with the recommendation provided by the system." This option was frequently used by loan officers who, during the period 2002-2008, granted around 80% of the loans that the system recommended rejecting.[20]

[17] We had access to all internal communications between central services and the branches.

[18] Note that the average price of a typical house in Spain was around 155,000 euros, meaning that loan officers at the branches could authorize most mortgages. In the case of personal loans, the limit before delegation to central services was 110,000 euros.

[19] The role of gender in decisions made by committees composed of many individuals is complex and reflects many different influences.

[20] It is in this sense that the Spanish banking crisis was a classical banking crisis derived from excessive risk taking, as it was the case in many other countries.

3.2 Characteristics of the scoring

An important aspect of our paper is the use of the scoring as conditioning variable to control for the risk and quality of the loans when analyzing the relationship between delinquency and gender of the loan officer. In fact, credit score models are generally not publicly available since they are a very sensitive element of the credit risk management of financial institutions. Rajan et al. (2015) argue that the scoring models used in the US during the period 1997-2006 were unstable because securitization changed the incentives of lenders. The securitization process that took place in the US during the period of 2000-2006 did not also happen in the financial system of many other countries. For instance, in the Spanish case, banking regulation did not allow to deconsolidate SPVs created with securitized mortgages and, therefore, banks could not improve their capital ratios by securitizing mortgages US-style. In addition, as we show later in this section, the internal risk models of the banks were validated every year and updated if there was any significant loss of predictive power. Internal documents of the bank show that the AUC[21] of the scoring model was systematically over 80% during the period under study.

The bank provided two scores: a behavioral scoring and a concessional

[21] The AUC, or Area Under the ROC Curve, is the usual machine learning device to check the discrimination ability of a binary classifier. It compares the sensitivity of the procedure (true positive rate) with the false positive rate (one minus the specificity). The integral of that area, normalized, is the AUC. It basically measures the probability of correctly identifying a good loan if faced with one random good and one random bad loan.

scoring. The former was used to offer small amount, pre-approved loans while the latter was used when the client did not have enough data to construct the behavioral score, or when the amount applied for was over the limit of the pre-approved loan. All the loans in our database were screened using the concessional, or standard, score. We use the behavioral score as an additional measure of the quality of the applicants, and for robustness purposes.

The bank did not share with us the full specification of their risk scoring models. In order to check the accuracy of their claims relative to the quality of their scoring model, we constructed our own model using most of the variables included in the validation data set[22]. In particular, we considered a variety of characteristics of the borrower and the loan: age, marital status, job contract type, destination of the loan, leverage ratio, debt over wealth, loan to value ratio, monthly mortgage payment over 6-month average bank account balance, nationality of the client, number of years at the current job, average bank balance over 6 months, 6/12 month bank balance ratio, an indicator for whether the individual is a bank client or not, number of years of continuous relationship with the bank, and number of years as a bank client.[23] Using this specification we derived the AUC for consumer loans (Figure 2) and mortgages (Figure 3). Our specification covers the whole period and, therefore, it is not strictly comparable with the results of the

[22] We only excluded variables that were mostly redundant or, in a few cases, had missing values for most of the loans.

[23] The bank did not use the gender of the client as a determinant of the scoring. Rather, the scoring was applied to all the adult members of the family, very frequently a couple, and the final screen was performed on the member with the highest score.

internal documents of the bank. The area under the ROC curve was 77.7 in the case of mortgages and 74.5 in the case of consumer loans. These results confirm the good quality of the data supporting the risk scoring model[24].

Financial institutions use diverse scoring models for different clients and products, and this is true of our data on concessional scoring. It is, for example, common to have one model for clients and another for non-clients, given that the respective availability of data is very different. It is also common to use diverse models to score applications for distinct products (mortgages, consumer loans, etc.). In fact, tables may also change over time when the models are updated. Furthermore, these scores generate different tables by product and/or client that evolve over time, with diverse ranges of variation. For this reason, the risk scores of the different models are frequently aligned into one adjusted score that synthesizes all the tables and allows to check the goodness of fit of the risk management system as a whole. While the bank provided the aligned behavioral score, it did not provided an adjusted concessional score.

We consequently generated a standardized concessional score, that we name "adjusted score" so as to distinguish it from the aligned score produced by the standarization of the behavioral scoring constructed directly by the bank. We use the following procedure. Denote $F(.)$ as the distribution function of the scores of each table. The reference score function is table

[24] An AUC of 70% or greater is the goal in information-rich environment as the one we discuss in this paper.

0 corresponding to product 0. Therefore, for any table i we can calculate the aligned score using the following algorithm. In step one we run the probability of default (PD) of the table of reference (0). Then, we run the probability model for all the scoring models (i). Using the predicted probabilities derived from that model and the parameters estimated in the reference model, we can obtain the adjusted scores. The empirical findings check the robustness of the results using alternatively the aligned behavioral score of the bank and our adjusted concessional score.

$$\begin{aligned} PD_0 &= F(\beta_0 * Score_0) \\ PD_i &= F(\beta_i * Score_i) \\ AdjScore &= F^{-1}(\hat{\beta}_0 * \hat{PD}_i) \end{aligned} \tag{1}$$

3.3 Characteristics of the sample

The advantage of including only loans to households is the fact that the internal risk assessment produces risk scores for all cases. Therefore, each of the loan officers had the same summary information about the quality of the loan based on the observable quantitative indicators used by the scoring system. By contrast, loans to SMEs and micro-companies are much more difficult to score appropriately and, consequently, no risk scoring is usually available.[25] This is also the case for the bank that provided the data. We eliminate from the population of household loans those that were authorized by a risk committee in central services due to the size of the request exceeding the authorization of the loan officer at the branch. Summarizing, we start with 422,302 applications for mortgages and personal loans. This is the whole population of those two types of loans handled by the bank during the period of analysis. In 40,648 cases the decision was taken by a committee at the central services of the bank because the loan overcame the delegation limits. This leaves 381,654 loans to households that were screened by the branches, our basic loan level administrative data. We do not consider the application available in the dataset after 2012 for reasons that we explain in the next section. Using these conditions we work with 380,237 observations. Finally, when we analyze the determinants of delinquency we obviously only consider the approved applications. This sample add up to 362,898 observations. The dataset include all the 1,507 loan officers of the bank. Female loan officers represent 22% of all the officers.

Table 1 presents the basic statistics of the data. The average punctuation of the adjusted concessional score[26] is almost identical for applications managed by male versus female loan officers. We can also examine differences in the distribution of the score of applicants depending on the gender of the loan

[25] Loans for large corporations are mostly scored using the ratings produced by rating agencies.

[26] Described in the previous section.

officer who managed the application. Figure 4 shows that the distributions of the standarized score of the applications are virtually identical for male and female loan officers. In addition, the distributions of the recommendations on the applications submitted to male and female loan officers are almost identical as shown by the second panel of Table 1. Moreover, all the loan officers, independently of their gender, have a very similar composition of applications in terms of the recommendations. Therefore, whether we look at the risk scores or the recommendations, the distributions of the applications received by male and female loan officers are very much alike.

However, the approval rate of loans is higher among male than female loan officers. The difference is four percentage points, although Table 1 shows that this is mostly concentrated among the loan applications with a rejection recommendation. This implies that the overruling rate, or the approval of loans notwithstanding a negative recommendation, is much higher for male loan officers than that for females loan officers. More specifically, we observe in Table 1 that the overruling rate for men is 13 percentage points greater than that for women. Table 1 also shows that the approval rate of men and women are almost identical for applications with an acceptance recommendation. The approval rate for application with a rejection recommendation are substantially higher for men than for women.

4 Basic results

The basic regression analyzes the relationship between the gender of the loan officer and the delinquency rate of the loans conditional on the quality of the applicant, as determined by the internal scoring rate. The basic specification is a logit model[27]

$$logit(Delinq_{ijt}) = \alpha\, male_{ijt} + \beta\, Score_{ijt} + \sum \gamma_k X_{ijkt} + \mu_t + \mu_j \qquad (2)$$

where *Delinq* is a dummy variables that takes value 1 if the loan has missed any payment for more than 90 days, which is the standard definition of delinquency; *male* is a dummy variable that takes value 1 if the loan officer was a man; *score* corresponds to the different versions of the score; X includes other explanatory variables; μ_t is a time dummy while μ_j is a geographical dummy.

Column 1 of Table 2 shows that loans approved by male loan officers have a delinquency rate that is 1.7 points higher than that of female loan officers.[28] This difference increases to 2.5 percentage points if we consider the cohort of the loan (column 2). This figure is statistically very significant but also economically important since the average delinquency rate of the loans in the sample is 12%. Conditional on the aligned behavioral score provided

[27]A linear probability model delivers almost identical results. The estimations report robust standard errors.

[28]To facilitate the interpretation of the parameters, they are expressed as average marginal effects in all of the tables. In addition, the variables score, age and tenure enter in the estimation divided by 100.

by the bank (Column 3), the loans approved by male loan officers have a delinquency rate that is 2.4 pp (percentage points) higher than that of female loan officers. The score is statistically very significant in the explanation of the delinquency rate. In particular, an increase of 100 points in the score decreases the probability of delinquency 1.5 pp. The result remains basically unaffected when we add experience or demographic characteristics of the loan officer (age). Older loan officers have a higher probability of granting loans that will be delinquent.[29] However, it is not very relevant in economic terms: 10 more years of age implied an increase in the delinquency rate of 0.2 pp. Experience as a loan officer reduces the probability of delinquency of the loans. These results are unaffected by the inclusion of geographical dummies[30].

As discussed above, the banks use different scoring models for different products, types of clients, and periods. Each of these models defines a particular scoring table. For example, Table 3 was used to obtain the scoring for mortgages for non-clients during the period 2003-09. The banks worked with a concessional scoring divided into 13 tables, with different models for clients and non-clients[31], and for personal loans and mortgages. The updating of

[29]This result is in line with the career concern model of Agarwal & Ben-David (2018), but contrasts the results of Beck et al. (2013).

[30]As a robustness check we have ran all the estimations using clustered standard errors at the level of each loan officer. In those cases the only variables that are always statistically significant are gender and score. We believe that our data do not fulfill the requirements for this clustering strategy (Abadie et al. (2017)). Results are available upon request.

[31]A client who opened an account less than 6 months before the calculation of the score is considered, from a scoring perspective, as a non-client given that some of the relevant variables used for the scoring of clients (e.g.,average account balance over the last

the different models over time also generated new scoring tables since the specification of the models changed.

Table 3 analyzes differences in delinquency rates by gender of the loan officer, considering the concessional score before adjustment. This approach avoids the need to adjust the scores to make them comparable across tables and periods. Table 3 reports the baseline probability and the increase in the probability of delinquency for males (interaction effect). The basic results of Table 2 are supported by the use of the concessional score by each scoring table. In general, female loan officers have a lower delinquency rate for loans they approved than do male loan officers. Scoring tables 8 to 13, which correspond to the scoring models used after 2009, represent an exception to this general finding. As argued above, and based on the analysis of the Internal Circulars, after 2009 there is a clear change in the management of credit risk, once the financial crisis was clearly impacting the Spanish economy. The ability of loan officers at the branches to grant loans was reduced and the exceptional conditions used to override the recommendation in the case of a rejection recommendation were eliminated. This shift corresponded to a general contraction in new loan origination and more restrictive practices by all Spanish financial institutions. The fact that after 2008 there is not a significant effect of gender on the delinquency rate implies that when scoring controls are tightened, for instance by eliminating the possibility of overriding a negative recommendation, male and female loan officers perform

6 months) cannot be calculated.

similarly.

That said, most of the loans of the sample belong to Scoring Tables 2 to 7, corresponding to the period prior to 2009, which show a statistically higher delinquency rate for loans granted by male loan officers. In particular, Scoring Tables 3, 5, and 6, which correspond to non-client applicants, show the largest difference in the delinquency rate between loans granted by male and female loan officers.

Table 4 replicates the estimation of the basic specification of Table 2 using our adjusted score, calculated as described in Section IV. The results of Tables 2 and 3 are confirmed. Female loan officer have a lower delinquency rate than male loan officers, ranging from 2.0 to 2.5 percentage points when there is a control for the cohort of the loan. As argued above, from the analysis of the Internal Circulars, and the results of Table 3, we know that the period before 2009 (prior to the banking crisis) was quite different from that after the beginning of the crisis. We also include a final column (7), which considers only those loans produced before 2009. The results show a difference of 2.1 percentage points, very similar to the findings using the full sample. This outcome is reasonable given that after 2008 the number of loans originated is very low compared to the pre-2009 period. Interestingly, the explanatory power of the specification using our version of the adjusted score is almost double the pseudo R^2 obtained using the aligned behavioral score provided by the bank. In the following sections, we consequently check the robustness of the results to the pre-2009 sample, and include our adjusted score as the indicator of the risk quality of the loan.[32]

5 Robustness

In the previous section, we showed, using alternative measures of the quality of the loans, that male loan officers have a higher delinquency rate than female officers. This section investigates the robustness of this finding to the inclusion of additional explanatory variables and specifications.

5.1 Adding characteristics of the loan officers

Table 5 includes some robustness checks. The basic results shared above are robust to these changes. Adding as an explanatory variable the interaction between tenure and male officer shows that improvement in the ability to screen applicants increases much faster for women officers than for males officers. In other words, an enhanced ability to screen bad loans, understood as a reduction in the delinquency rate of the loans approved as function of the years spent as a loan officer, occurs more quickly among female officers compared to male officers.

Of interest as well is the fact that having experienced the previous crisis as a banking employee does not immunize the loan officer from granting bad loans. Measuring exposure to a previous crisis represents a challenge. In

[32]The results remain basically unchanged if we use the aligned score provided by the bank.

order to gauge the possible influence of this experience, we employ specific time periods as reference points. The first and main threshold is defined as being hired initially before 1995, which corresponds to the previous banking crisis of the Spanish economy. The crisis, starting in 1992, involved the failure of Banesto, a major Spanish bank, and a rapid increase in the proportion of non-performing loans, peaking in 1995.[33] We find that the experience of a previous financial crisis did not prevent loan officers from approving bad loans. In fact, quite the opposite: loan officers who were already working in the banking sector during the previous financial crisis present a statistically significant higher delinquency rate than other loan officers, although the effect is economically small. The average number of loans approved by loan officers also increases the delinquency rate.

5.2 Decomposing the score

In the previous tables, we used two alternative scores (the aligned behavioral score and the adjusted concessional score) as indicators of the quality of the loans. As argued in Section 3, the banks' validation reports of the scoring show AUCs above 80%, implying a good level of accuracy. Nonetheless, we investigate the robustness of the results to the use of variables that are known to be determinants of the quality of loans. More specifically, rather than using the scores provided by the bank, which were calculated using a confidential model that the bank did not share with us, we generated our

[33] See Figure 1.

own scoring model to assess whether the results are robust to the direct use as explanatory variables of those factors commonly included in the calculation of scoring models.[34] The estimates are obtained using variables that, with high likelihood, were part of the banks' scoring, such as age, marital status, type of job contract, loan type, destination of the loan, leverage ratio, loan to value ratio (in case of mortgages), total debt over wealth, average balance over six months, 6/12 months bank balance ratio, monthly mortgage payment over 6-months, average bank account balance, nationality, number of years in the current job, indicator of client and number of years as client of the bank, as well as the years of continuous relationship with the bank[35]. Table 6 include the descriptive statistics of those variables.

The results of Table 7, where we substitute for score with the abovementioned variables, are consistent with the previous results: loans granted by female officers present a delinquency rate between 1.6 and 1.8 percentage points lower their male counterparts. The effect of age and tenure have the same sign as before but are not statistically significant. Although if we also include the adjusted score (Columns 5 and 6), we still find some additional explanatory power, although the coefficient is largely reduced compared with previous results. Table A.1, in the Appendix, shows that the basic results using the components of the score are robust to using additional controls for

[34]Figures 2 and 3 already show the Area Under the Curve (AUC) of these scoring models for consumer loans and mortgages respectively.

[35]These are the same variables that we used to obtain the accuracy of the scoring model in Section III

the demographic characteristics of the loan officer.

5.3 Alternative definitions of delinquency

The performance database contains several indicators of delinquency depending on how many times, or for how long, the client missed a payment. These are codified for the whole life of the loan. The performance is measured as the number of missing payments over 30 days, 60 days, and 90 days. In previous sections, we defined a delinquent loan as having missed at least one payment for a period of 90 days. The 90-day threshold is the standard used in many countries to define a non-performing loan (NPL).

It is hence of interest to check the robustness of our findings to changes in the measurement of the performance of the loan. Table 8 considers 60 days as the threshold to classify a loan as non-performing, while Table 9 considers as a NPL those missing at least one payment over 30 days. The basic results are robust to these new definitions of performance of the loans. In fact, for periods below 90 days, the performance of male loan officers relative to female loan officers worsens with respect to the 90-day threshold. Tables A.2 and A.3, in the Appendix, show that the results of Tables 8 and 9 are robust to including additional controls for the demographic characteristics of the loan officers.

5.4 Panel data analysis using branches

Table 10 shows the estimation of the basic specification using a two-way panel data model with branches as the reference of the analysis.[36] The results show that male loan officers have between a 1.1 and 1.5 point higher delinquency rate than female loan officers, consistent with the results of previous exercises. If we consider the period prior to 2009, the estimate is 1.4 percentage points. The estimator is statistically significant in all of the columns as well as economically important: women loan officers show an 11.8% lower delinquency rate than men. As in previous tables, the delinquency rate is higher for older loan officers although, in this case, it is not statistically significant. Here, however, there is no effect of tenure on the proportion of non-performing loans.

6 Explaining the findings

The previous sections have shown that there is a significant difference between the delinquency rate of men and women loan officers conditional on the risk score of each loan.

Is this difference relevant for the bank? Could it be that, even having a higher delinquency rate, men generate a higher return adjusted by risk than women? This is unlike since, as we explained before, the interest rate charged

[36]We resume using as the definition of non-performing loan those missing at least one payment over 90 days.

for loans in Spain during the period of analysis was quite insensitive to the risk score of the client. A precise calculation of the Risk Adjusted Return on Capital (RAROC) confirms the previous hypothesis[37]. The interest rate of mortgages approved by male and females loan officers are almost identical. The same is true for consumption loans. Financial cost is also similar for both. We assume the same loss given default (LGD) for males and females' loan officers since this parameter is not available, although we assume different LGD for mortgages (15%) and consumption loans (45%). Obviously, given the higher default rate of loans approved by males, the largest difference between males and females in the calculation is the expected loss. The cost of the structure is supposed to be the same independent of the gender of the officer who approved the loan. Finally, economic capital is calculated as the sum of the required capital associated to credit risk and operational risk. It depends on the size of the loans but it is obviously independent of the gender of the loan officer. This calculation leads, in fact, to a higher RAROC for loans approved by female than male loan officers.

Having discussed the financial implications of the differences in the delinquency rate, in this section, we introduce a new piece of information - the recommendation of the system - and offer an explanation for the gender difference in delinquency rates.

[37] In this section we present a discussion of the main hypotheses and results of the calculation. The precise calculations are available upon request

6.1 Gender differences in rule compliance

In previous sections, we analyzed the effect of the credit score. Here we instead consider the influence of the recommendations of the system on the decisions of the loan officers, conditional on the risk score. The recommendation system reflected specific credit policies of the bank, and took the form of a categorical variable with five levels[38]. The Spearman's correlation between risk score and recommendation is -0.63, implying that the correlation between the risk measured by the scoring model and the recommendation is far from perfect[39]. The five recommendation categories are: very positive (A1), positive (A2), neutral (A3), negative (D1) and very negative (D2). Categories D1 and D2 implied a recommendation to reject the application.

Loan officers could, however, override the recommendation in "exceptional cases." The reasons for doing so, and forcing an approval, were explained using several standard sentences, codified into 34 categories. Most were very subjective justifications, such as "the applicant has good prospects of generating future business with the bank" (36.6%) and "the client has a positive credit history with the bank" (21.3%). Other comments downplayed important components of the scoring model. For instance, some operations with a negative recommendation were approved arguing that "the applicant has a temporary contract but has been working continuously in recent years." There were also cases where the loan officer chose to overrule the recom-

[38]The bank did not share with us the algorithms that generated the recommendations.
[39]Later in this section we show that the recommendation has explanatory power on the delinquency rates ever after controlling by the risk scoring.

mendation even when the client was included on a list of known delinquent debtors[40] or had experienced issues in the payment of previous loans at the bank. In such cases, the reason given for overriding the recommendation is that "the incidence has been regularized."

After 2008, many restrictions were placed on the ability of branch loan officers to approve loans. Credit contraction, a consequence of the financial crisis, meant a tightening of the rules. To this regard, the Internal Circular A2-088/08 states that branches could not approve new loans to any applicant, either holder or guarantor of a previous loan, who had had any loan delinquent for more than 30 days, a refinancing operation, or any incidence in the risk information service. It also prohibited, with no exception, the approval of applications for which the system had recommended rejection (negative and very negative).

As a first approximation to understand the observed difference in the delinquency rates of loans handled by men and women, we can decompose, from a purely accounting perspective, the default rate for each gender into the delinquency rate of positive and negative recommendation loans.[41]

$$\begin{aligned} P(D|G) = {} & P(PR|G)P(A|PR,G)P(D|A,PR,G) \\ & + P(NR|G)P(A|NR,G)P(D|A,NR,G) \end{aligned} \tag{3}$$

[40]For example, the applicant was listed in the ASNEF registry, the "black list" of defaulters managed by EQUIFAX.

[41]To simplify the decomposition, we also include the neutral level in the positive recommendation category.

where D equals 1 if the loan is delinquent, PR is a positive recommendation, NR is a negative recommendation, A is approval, and G is either a man or a woman. We observe no gender difference in the proportion of loans with a negative recommendation handled by men and women during the relevant period (before 2009). However, female loan officers rejected 35% of loans classified by the recommendation system as negative or very negative, whereas male officers rejected only 19% of such loans. In addition, the delinquency rate among negative and positive recommendation loans, conditional on approval, was higher for men than for women. The difference between the genders in the case of positive recommendation loans is 1 percentage point while for negative recommendation loans it is 4.1 percentage points.

These distinctions in the delinquency rates of men and women reflect three components: differences in the likelihood of handling positively and negatively labeled applications; differences in the approval rates of positive and negative recommendation applications; and differences in the delinquency rates conditional on the recommendation. For positive recommendation loans, the difference in the delinquency rate component explains 98.5% of the gender difference. This is reasonable since the difference in the approval rate of positively recommended loans is almost identical (98.4% for men and 98% for women). Women seemingly have a greater ability to read soft information compared to men, given that all the hard information was already included in the risk score and the recommendation, and the approval rates are very similar in both cases. For loans with a negative recommendation, the gender difference in the delinquency rate component explains 25% of the difference in the delinquency rate, the discrepancy in the approval rate explains around 70%, and the difference in the proportion of negatively labeled loans explains the remaining 5%.

To analyze men's and women's compliance with the rules we run several empirical exercises. Table 11 presents some logit specifications to explain the differences in the approval rates of men versus women by type of recommendation. The set of explanatory variables is the same as in previous regressions. Table 11 shows, as expected, that the approval rate decreases with the worsening of the recommendation. More interestingly, the difference in the approval rates of men and women loan officers increases monotonically with the worsening of the classification of the loan. For very positive and positive recommendations, there is basically no difference in the rejection rates. However, for the neutral recommendation loans, the difference is 1.1 percentage points, which increases to 3.3 for negative recommendation and 5.8 for very negative recommendation loans.[42]

An analysis of the overruling behavior of men and women offers another perspective of the results presented in the previous paragraph. Considering only the loans that received a negative recommendation label, Table 12 shows that men overrule the recommendation of the system significantly more often

[42]An Oaxaca-Blinder decomposition shows that the difference in the approval rates of men versus women, generally as well as for positive recommendation and negative recommendation loans, is almost completely due to the structure (coefficients), with a very small contribution of the composition effect.

than women. Consistent with the results of Table 11, the difference in the overruling proportion increases with the worsening of the category assigned by the recommendation system. The coefficient on the combination of negative recommendation loans by men implies that their overruling rate is 5.5 percentage points higher than that of women, signifying that female loan officers comply with the rules more often than men.

In previous sections, we showed that the risk scoring, in different versions, is a statistically significant determinant of the delinquency rate. Does following the rules also provide an advantage in terms of lower delinquency rates? Table 13 analyzes the determinants of the delinquency rate by recommendation category controlling for risk scores. As expected, the delinquency rate increases monotonically with the worsening of the loan recommendation. But, as in the previous table, conditional on the recommendation, male loan officers are associated with a higher rate of delinquency, especially for loans in the negative or very negative categories.

These results indicate that female loan officers reject more loans with negative recommendations than do men. They also select, among the negative recommendation loans, a pool with a lower default probability than that chosen by men. This explains why the loans produced by female loan officers show a lower delinquency rate than those approved by male loan officers.

Therefore, even after conditioning by the risk score, the recommendation should contain relevant information. This can be seen in Tables 11 and 13 where both risk score and recommendation are strongly statistically significant. We can thus interpret this result as meaning that conditional on the risk score, females loan officers follow the rules more often than male loan officers, and that this provides an advantage in terms of avoiding loan delinquency.

In the first section of this paper, we discussed several theories that could explain why gender may have an effect on the lower delinquency rate of loans monitored by women. Our study is unique in that, differently from previous research, the officer observes the risk score and the recommendation of the system before making a decision on a loan. Therefore, their initial decision is based on following or overruling the recommendation of the system.

Social psychology research finds that women are more compliant than men.[43] There is also extensive evidence that dangerous behavior and involvement in car accidents among adults are more often due to rule-breaking among males than females. Women abide by road signs more often than men; they also less frequently violate pedestrian rules.[44] This section has similarly shown that female loan officers tend to follow the rules more often than their male counterparts, which consequently means that they generate lower levels of delinquency.

Could an alternative reason explain the findings? Can differences in risk attitudes between males and females explain the results? If females are more risk averse than males then, as risk increases, female loan officers should

[43] See the classical reference of Tittle (1980).
[44] Rosenbloom (2009).

reject more loans than their male counterparts. Unfortunately, since the recommendation is not just a function of the score, it is not possible to run a regression discontinuity exercise to identify gender differences in the rejection rate around the negative recommendation threshold. However, looking at very positive, positive and neutral recommendations there is not much of a difference in the rejection rate of females and males, even though average risk clearly increases as you move toward the neutral recommendation. The difference between the rejection rate of female and male loan officers jumps to 15 pp when the recommendation is negative.

An alternative approach is to check if the typical difference in risk aversion between males and female can explain the results of the previous sections. There is a large literature on gender differences in risk aversion although, there is less agreement on the extent of that difference[45] We consider a type of lottery in which with probability 0.87 the loan officer gets an application with a positive recommendation, and with a probability 0.13 the officer gets an application with negative recommendation. These are the probabilities observed in our data. Using the probabilities of rejection, default conditional on rejection, etc. of male and female officers, discussed above in this section, we can calculate the difference in risk aversion needed for female officers to choose the rejection rates observed in the data instead of the rates showed by

[45] We consider the gender differences in risk calculated in Filippin & Crosetto (2016) using the Holt & Laury (2002) procedure and the risk-elicitation procedure of C. Eckel & Grossman (2002).

male loan officers[46]. Using a constant relative risk aversion function we find that the rate of risk aversion required to justify the rejection rate of negatively recommended loans by female loan officers is 10 times the highest value obtained by Filippin & Crosetto (2016). Obviously, this is just a back of the envelop calculation since it is not clear that a constant relative risk aversion is a good representation of the utility function, and there is controversy over the actual value of the gender difference in risk aversion. Nevertheless, these calculations indicate that it is difficult to claim that the gender differences observed in the data are simply derived from differences in risk aversion.

6.2 Gender bias in the mistake-punishment trade-off

Why do women follow the recommendation of the system more often than men? High level managers in the risk department of the bank described a phenomenon that they define as gender bias in the "mistake-punishment trade-off." More specifically, one reason why female loan officers were afraid to deviate from the recommendation of the system was that if they approved a negatively recommended loan that then became a non-performing loan,

[46]The bonus scheme provided a strong incentive to approve loans. The variable pay was function of several indicators, and could reach up to 20% of the fixed pay in case of reaching 100% of the objective in all the indicators. The indicators related with the number of loans approved by an officer amounted to 60% of the variable pay. There was no variable pay if the indicator did not reach at least 50% of the objective. We have also estimated, using our data, the probability of loosing the conditions of loan officer as a function of the accumulated default rate. For reasons that will become clear next section, we use the same rate for males and females.

their careers would be damaged more so than those of men[47]. They therefore had a strong incentive to follow the recommendation of the system.

The literature has recently discussed the possibility of double standards in terms of punishment for breaking the rules. Egan et al. (2017), for example, analyze gender discrimination in the financial services industry. Using a panel data on misconduct reported to FINRA for 1.2 million financial advisors in the US between 2005 and 2015, they find that women face harsher punishment for misconduct. They are, in fact, 20% more likely than men to lose their job after a misconduct incident. Sarsons (2019) analyzes primary care physicians' (PCPs) referrals to surgeons and describes an asymmetric updating in terms of gender. PCPs drop referrals to female surgeons more sharply than to male surgeons after a patient death. She concludes that women have fewer chances to make mistakes, which in turn could mean lower promotion rates compared to men.

As we have the career histories of the loan officers and can thus link the performance of the loans they approved with the latter, we are able to investigate potential gender bias in the mistake-punishment trade-off. In particular, we are interested in whether gender is a determinant of the duration of the position as loan officer conditional on the accumulation of non-performing loans, which increases the probability of being demoted or, eventually, dismissed.

[47] Conversations with managers and many loan officers indicated that female loan officers were aware of this gender bias in terms of a mistake-punishment trade-off.

To check the performance of each loan officer, we constructed the variable BadL, which corresponds to the accumulation of bad loans by an officer. More specifically, this variable represents the proportion of loans generated by each loan officer that are delinquent for more than 90 days, as per the definition of delinquency used in the previous sections of the paper. Figure 5 shows the Kaplan-Meier non-parametric estimator of the survival function for loan officers who have accumulated at least 4% of bad loans versus officers who have less than 4% of delinquent loans[48]. Figure 5 shows that the probability of being demoted as a loan officer, or dismissed, increases drastically when bad loans accumulate in the portfolio of a particular loan officer. This result confirms the usefulness of this indicator as a measure for the evaluation of loan officers. Figure 6 shows the estimation of the survival functions for male and female loan officers. We observe that, unconditionally, males remain longer in the position of loan officer than females, although the difference in the survival functions is less striking than that depicted in Figure 5. That said, as these unconditional figures are not evidence of the hypothesis of gender bias relative to the mistake-punishment trade-off, we run several duration models.

Using the bad loans variable, we can investigate the effect of the accumulation of bad loans on the careers of men and women, and determine if,

[48]The results of the figures and the analysis in this section are unaffected if we used the proportion of delinquent loans conditional on a negative recommendation instead of the unconditional delinquency rate. This is not surprising given the high correlation between delinquency and recommendations.

conditional on the accumulation of such loans, there is a differential tenure in the position of loan officer by gender. Survival analysis modelling usually assumes that the values of all the covariates were determined a time 0. However, in this case, it is necessary to consider at least one time-varying covariate, as the proportion of bad loans accumulated by each loan officer could change over time. The basic variables in our specification are gender, age, and the proportion of bad loans. Table 14 shows the results using different specifications for the hazard rate of the tenure as loan officer. The first column includes the estimation of a proportional hazard model.

$$\begin{aligned}\lambda(t, x, \beta) = \lambda_0(t) \exp(&\beta_1 * male_i + \beta_2 * age_i \\ &+ \beta_3 * BadL_{it})\end{aligned} \tag{4}$$

The results show that, conditional on a loan officer's age and the proportion of bad loans accumulated in the past, men have a 34% lower rate of being demoted from their position as loan officer than do women. The hazard rate increases 3% for each point of increase in the accumulation of bad loans. As expected from Figure 5, reaching a high proportion of bad loans implies a high hazard of being demoted from the position. In particular, the hazard rate increases 12.3%, for a rise of four points in the proportion of bad loans.

Columns 2 to 4 include the estimation of parametric models of increasing flexibility in terms of the shape of the hazard function. The Weibull and Gompertz models in the following two columns generate similar results, and their respective ancillary parameters, p and γ, signal that the hazard rate is monotonically increasing. Unlike the previous parametric models, the Gamma regression coefficients in Column 5 can only be reported in accelerated failure-time metric. The results here support the previous findings that men have a lower probability of being demoted than women.

Table 15 includes as an additional explanatory variable the cross-product of gender by bad loans. The term captures gender bias relative to the mistake-punishment trade-off. In this specification, the dummy for male is not statistically significant, as observed for the results in Table 14. However, the interaction effect of male and the accumulated proportion of bad loans is statistically significant. Taking the results of the proportional hazard model[49] we observe that while the estimated log hazard function with respect to the proportion of bad loans has the same origin (since the male dummy is not statistically significant), the slope of females' log hazard function is higher than that of males. For instance, at 2% of accumulated bad loans, men have a 9% lower probability of being demoted or dismissed than women. At 4% of bad loans, males have a 16% lower probability of being demoted or dismissed than females. The results remain basically unchanged using the same basic specification for alternative parametric models.

Tables 16 adds to the specification the average number of loans produced by each loan officer every month. This variable considers the revenue side of the production of loans. The basic results of Table 15 are fundamentally

[49] Note that in this case, describing the model as proportional hazard is not, strictly speaking, appropriate. We use this terminology since it is generally adopted in the literature.

unchanged. The male dummy is not statistically significant. The effect of bad loans and the interaction of bad loans by gender are similar to those discussed in Table 16. Finally, the number of loans produced by loan officers is not statistically significant in most of the specifications. In fact, it appears that the more credits that are approved, the higher the hazard of dismissal, which may be due to a higher proportion of bad loans associated with a fast rate of loan production. This conclusion would be consistent with the results of Table 5.

7 Conclusions

In this paper we analyze the effect of the gender of loan officers on the approval of loans and, in particular, on posterior delinquency rates. Using information from a large Spanish bank, we show that female loan officers produce loans that have a smaller non-performing rate than those screened by male loan officers. In fact, in our sample of close to half a million loans, female loan officers have a delinquency rate that is 1.5 to 2.5 points lower than their male counterparts. This result is economically very significant since it amounts to between 12.5% and 20% of the average delinquency rate. The fact that after 2008 there is no longer a significant effect of gender on the delinquency rate implies that when controls are tightened, for instance eliminating the possibility of overriding the recommendation of rejecting a loan, male and female loan officers perform similarly. One reason for the better performance of the loans screened by women is that, conditional on risk measured by the score, female loan officers followed the recommendation of the system more often than men. This higher compliance with the rules is potentially explained by a differential punishment of women versus men in the case of bad outcomes/performance.

Using our data, we find that women who accumulate a high proportion of non-performing loans have a greater probability of being punished than do men conditional on the same level of performance.

References

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2017). When should you adjust standard errors for clustering? *NBER Working Paper 24003*.

Agarwal, S., & Ben-David, I. (2018). Loan prospecting and the loss of soft information. *Journal of Financial Economics*, *129*, 608-628.

Agier, I., & Szafarz, A. (2013). Microfinance and gender: Is there a glass ceiling on loan size? *World Development*, *42*, 165-181.

Alesina, A. F., Lotti, F., & Mistrulli, P. E. (2013). Do women pay more for credit? Evidence from Italy. *Journal of the European Economic Association*, *11*, 45–66.

Andres, P., Gimeno, R., & Mateos, R. (2019). The gender gap in credit access. *WP Bank of Spain*, *1945*.

Asiedu, E., Freeman, J., & A., N.-A. (2012). Access to credit by small business: how relevant are race, ethnicity and gender? *American Economic Review*, *102 (5)*, 532-537.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence and common stock investment. *Quarterly Journal of Economics*, *116*, 261-292.

Barsky, R. B., Juster, T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement study. *Quarterly Journal of Economics*, *112*, 537-579.

Beck, T., Behr, P., & Guettler, A. (2013). Gender and banking: Are women better loan officers? *Review of Finance*, *17*, 1279-1321.

Bellucci, A., Borisov, A., & Zazzaro, A. (2010). Does gender matter in bank—firm relationships? Evidence from small business lending. *Journal of Banking & Finance*, *34*, 2968-2984.

Bentolila, S., Jansen, M., & Jimenez, G. (2017). When credit dries up: job losses in the great recession. *Journal of the European Economic Association*, *16*, 650-695.

Berg, T., Puri, M., & Rochol, J. (2019). Loan officer incentives, internal rating models and default rates. *Review of Finance*, *Forthcoming*.

Blanchflower, D. G., Levine, P. B., & Zimmerman, D. J. (2003). Discrimination in the small-business credit market. *Review of Economics and Statistics*, *85*(4), 930–943.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*, 367-383.

Campbell, J. Y., & Cocco, J. F. (2015). A model of mortgage default. *The Journal of Finance*, *70*, 1495-1554.

Carter, S., Shaw, E., Lam, W., & Wilson, F. (2007). Gender, entrepreneurship, and bank lending: The criteria and processes used by bank loan officers in assessing applications. *Entrepreneurship Theory and Practice*, *31*, 427-444.

Croson, R., & Gneezy, U. (2009). Gender difference in preferences. *Journal of Economic Literature*, *47*(2), 448-74.

Cueva, C., & Rustichini, A. (2015). Is financial instability male-driven? Gender and cognitive skills in experimental asset markets. *Journal of Economic Behavior & Organization*, *119*, 330-344.

Doering, L. (2018). Risks, returns and relational lending: Personal ties in microfinance. *American Journal of Sociology*, *123 (5)*, 1341-1381.

Eckel, C., & Grossman, P. (2002). Sex differences and statistical stereotyping in attitudes towards financial risk. *Evolution and Human Behavior*, *23 (4)*, 281-295.

Eckel, C. C., & Füllbrunn, S. C. (2015). Thar she blows? Gender, competition, and bubbles in experimental asset markets. *American Economic Review*, *105*, 906-920.

Eckel, C. C., & Füllbrunn, S. C. (2017). Hidden vs. known gender effects in experimental asset markets. *Economics Letters*, *156*, 7-9.

Eckel, C. C., & Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, *68*, 1-17.

Egan, M. L., Matvos, G., & Seru, A. (2017). When Harry fired Sally: the double standard in punishing misconduct [Working Paper]. *National Bureau of Economic Research*(23242).

Filippin, A., & Crosetto, P. (2016). A reconsideration on gender differences in risk attitudes. *Management Science*, *62 (11)*, 3085-3105.

Hallahan, T. A., Faff, R. W., & McKenzie, M. D. (2004). An empirical investigation of personal financial risk tolerance. *Financial Services Review*, *13*, 57.

Holt, C., & Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review*, *92 (5)*, 1644-1655.

Kahn, A., Hottes, J., & Davis, W. L. (1971). Cooperation and optimal responding in the prisoner's dilemma game: Effects of sex and physical attractiveness. *Journal of Personality and Social Psychology*, *17*(3), 267-279.

Khandker, S. (2005). Microfinance and poverty: Evidence using panel data from bangladesh. *The World Bank Economic Review*, *19*, 263-286.

Liberti, J. M., & Mian, A. R. (2009). Estimating the effect of hierarchies on information use. *Review of Financial Studies*, *22*, 4057-4090.

Mascia, D. V., & Rossi, S. P. S. (2017). Is there a gender effect on the cost of bank financing? *Journal of Financial Stability*, *31*, 136-153.

Mayordomo, S., Rachedi, O., & Rodriguez, M. (2019). Bank regulatory capital arbitrage: evidence from housing overappraisals. *WP Bank of Spain*.

Moro, A., Wisniewski, T., & Mantovani, G. (2017). Does a manager's gender matter when accessing credit? evidence from european data. *Journal of Banking and Finance*, *80*, 119-134.

Muravyev, A., Talavera, O., & Schäfer, D. (2009). Entrepreneurs' gender and financial constraints: Evidence from international data. *Journal of Comparative Economics*, *37*, 270-286.

Neelakantan, U. (2010). Estimation and impact of gender differences in risk tolerance. *Economic Inquiry*, *48*, 228-233.

Niederle, M. (2016). Gender. In *Handbook of experimental economics* (p. 481-553). Princeton University Press: John Kagel and Alvin E. Roth.

Olsen, R. A., & Cox, C. M. (2001). The influence of gender on the perception and response to investment risk: The case of professional investors. *The Journal of Psychology and Financial Markets*, *2*, 29-36.

Ongega, S., & Popov, A. (2016). Gender bias and credit access. *Journal of Money, Credit and Banking*, *48 (8)*, 1691-1724.

Pitt, M. M., & Khandker, S. R. (1998). The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter. *Journal of Political Economy*, *106*, 958–996.

Powell, M., & Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of economic psychology*, *18*, 605—-628.

Rajan, U., Seru, A., & Vig, V. (2015). The failure of models to predict failure: Distance, incentives and defaults. *Journal of Financial Economics*, *115*, 237–260.

Reinhart, C., & Rogoff, K. S. (2009). *This time is different: Eight centuries of financial folly.* Princeton: Princeton University Press.

Robb, A., & Wolken, J. (2002). Firm owner, and financial characteristics: differences between female and male-owned small business. *Federal Reserve Finance and Economic Discussion series*, *18*.

Rosenbloom, T. (2009). Crossing at a red light: Behaviour of individuals and groups. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*, 383–394.

Sarsons, H. (2019). Interpreting signals in the labor market: Evidence from medical referrals. *mimeo*.

Schubert, R., Brown, M., Gysler, M., & Brachinger, H. W. (1999). Financial decision-making: Are women really more risk-averse? *The American Economic Review*, *89*, 381–385.

Stefani, M. L., & Vacca, V. P. (2013). Credit access for female firms: Evidence from a survey on european SMEs. *Questioni di Economia e Finanza (Occasional Papers)*, *176*.

Tittle, C. R. (1980). *Sanctions and social deviance: The question of deterrence*. New York: Praeger.

Treichel, M., & Scott, J. (2006). Women owned business and access to bank credit: evidence from three surveys since 1987. *Venture Capital*, *8 (1)*, 51-67.

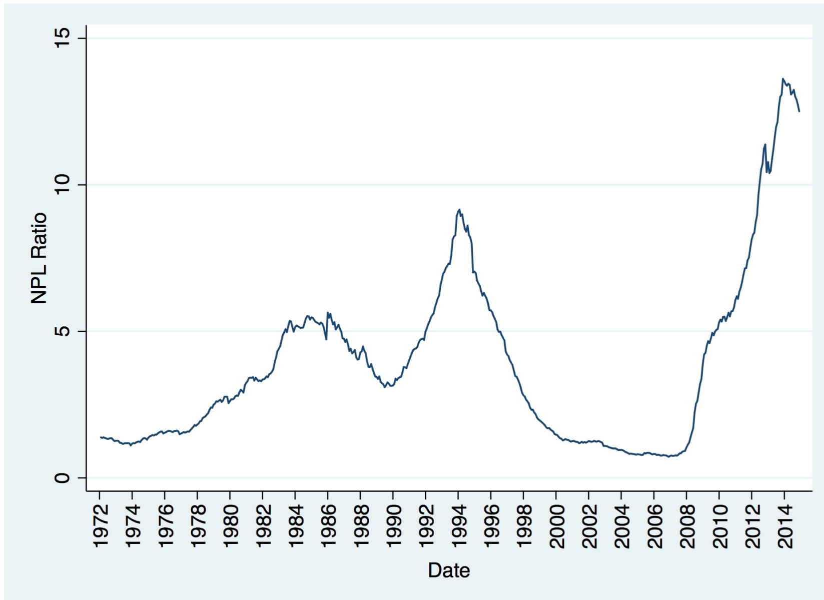

Figure 1: Non-performing loans rate

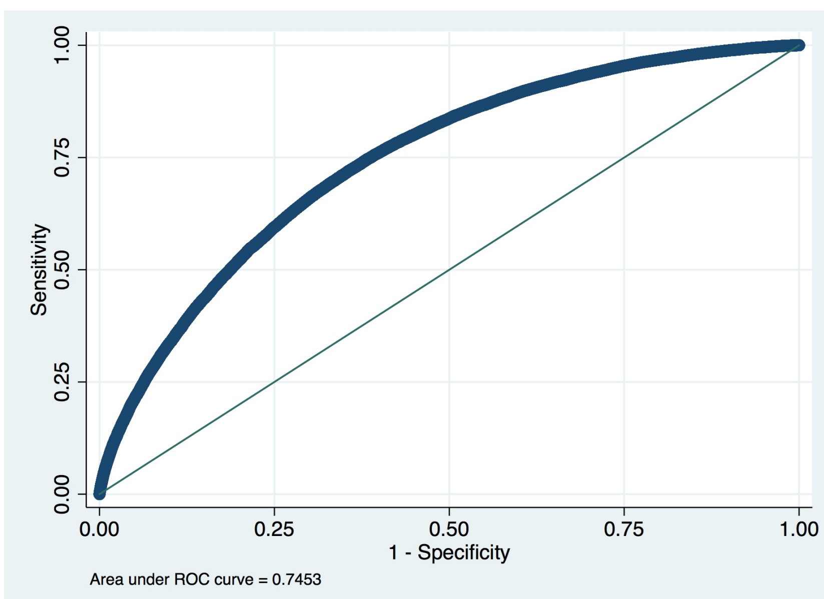

Figure 2: AUC-consumer loans

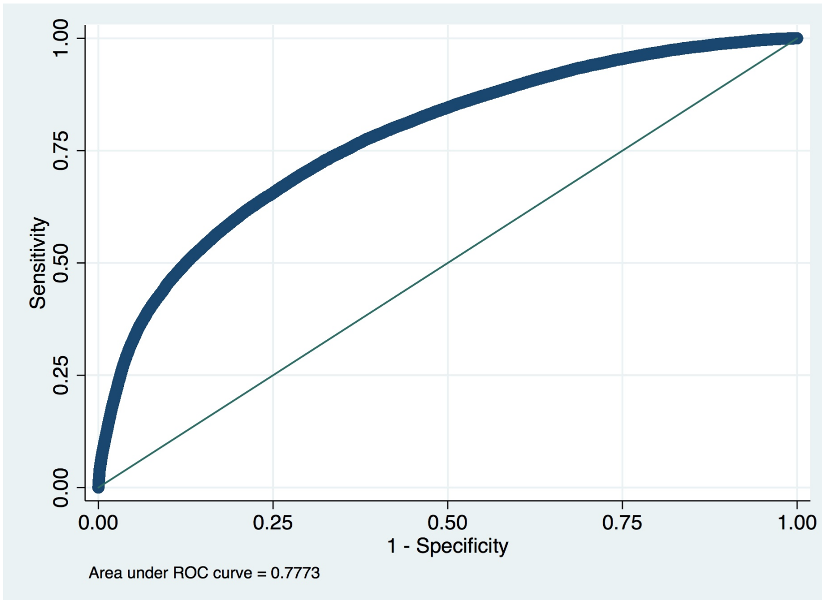

Figure 3: AUC-mortgage loans

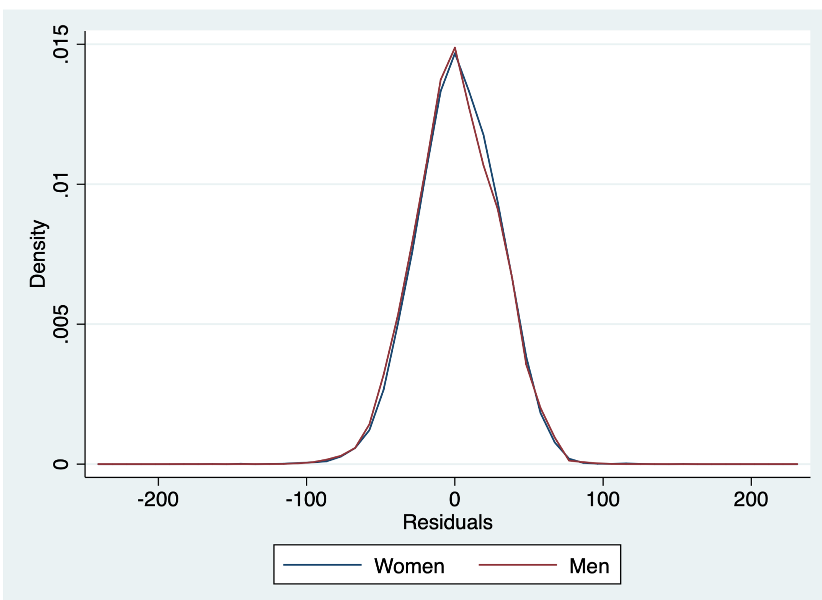

Figure 4: Standardized score of applications by gender

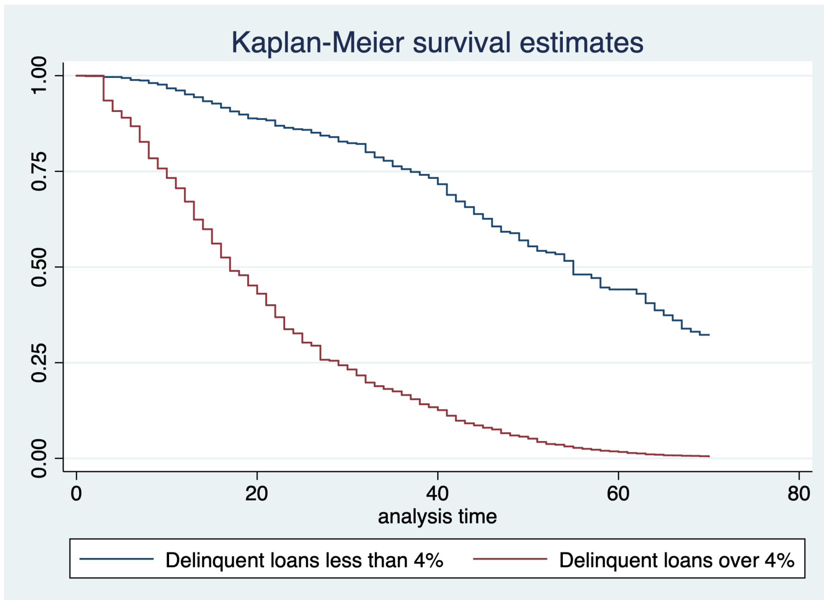

Figure 5: K-M SURVIVAL ESTIMATES BY PROPORTION OF DELINQUENT LOANS

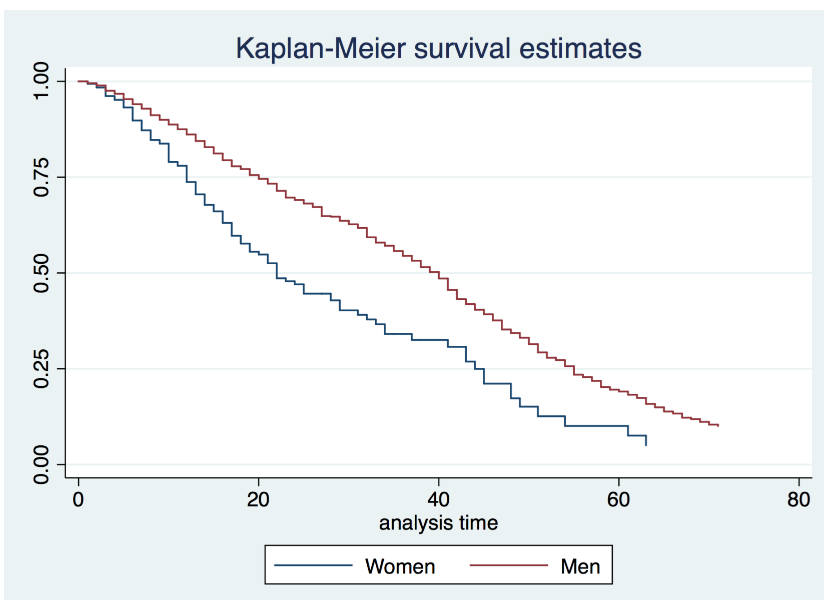

Figure 6: K-M SURVIVAL ESTIMATES BY GENDER

Table 1: Descriptive Statistics

	Total	Women	Men
Adjusted score	475	475	476
Applications by recommendation			
Very positive (A1)	0.30	0.29	0.30
Positive (A2)	0.43	0.43	0.43
Neutral (A3)	0.14	0.14	0.14
Negative (D1)	0.09	0.09	0.09
Very negative (D2)	0.05	0.05	0.04
Overall approval rate	0.95	0.92	0.96
Overruling share	0.77	0.66	0.79
Delinquency rate	0.12	0.10	0.12
Approval rate by recommendation			
Very positive (A1)	0.99	0.99	0.99
Positive (A2)	0.98	0.98	0.98
Neutral (A3)	0.96	0.94	0.96
Negative (D1)	0.83	0.70	0.85
Very negative (D2)	0.66	0.59	0.67

Table 2: Basic regression analysis with aligned score

	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.017***	0.025***	0.024***	0.022***	0.023***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Aligned score			−0.015***	−0.015***	−0.015***	−0.015***
			(0.000)	(0.000)	(0.000)	(0.000)
Age				0.026***	0.048***	0.035***
				(0.008)	(0.009)	(0.009)
Tenure					−0.087***	−0.047***
					(0.014)	(0.014)
Time effect	No	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	No	No	Yes
Observations	362898	362898	362898	362898	362898	362871
Pseudo R^2	0.00	0.03	0.05	0.05	0.05	0.06

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Basic regression analysis with score tables

	(1)	(2)	(3)	(4)	(5)
Table 1: Personal loans	0.148***	0.217***	0.232***	0.233***	0.230***
Non-client (2000-04)	(0.013)	(0.018)	(0.019)	(0.019)	(0.019)
Table 1 × Male	0.008	0.013	0.008	0.007	0.010
	(0.014)	(0.019)	(0.019)	(0.019)	(0.019)
Table 2: Personal loans	0.051***	0.077***	0.086***	0.086***	0.086***
Client (2000-04)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
Table 2 × Male	0.011***	0.018***	0.016***	0.016***	0.015***
	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
Table 3: Mortgage loans	0.145***	0.116***	0.128***	0.128***	0.133***
Non-client (2003-09)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Table 3 × Male	0.067***	0.070***	0.067***	0.068***	0.066***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Table 4: Mortgage loans	0.093***	0.074***	0.083***	0.083***	0.084***
Client (2003-09)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Table 4 × Male	0.017***	0.020***	0.018***	0.018***	0.017***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Table 5: Personal loans	0.181***	0.123***	0.135***	0.134***	0.135***
Non-client (2005-2006)	(0.015)	(0.011)	(0.012)	(0.011)	(0.012)
Table 5 × Male	0.104***	0.080***	0.077***	0.078***	0.077***
	(0.016)	(0.012)	(0.012)	(0.012)	(0.012)
Table 6: Personal loans	0.255***	0.171***	0.186***	0.185***	0.185***
Non-client (2006-09)	(0.016)	(0.012)	(0.013)	(0.013)	(0.013)
Table 6 × Male	0.066***	0.044***	0.040***	0.041***	0.042***
	(0.017)	(0.013)	(0.013)	(0.013)	(0.013)

Continued on next page

Table 3: (Continued)

	(1)	(2)	(3)	(4)	(5)
Table 7: Personal loans	0.104***	0.073***	0.082***	0.082***	0.082***
Client (2005-2011)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Table 7 × Male	0.023***	0.013***	0.011***	0.012***	0.011***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Table 8: Mortgage loans	0.372***	0.383***	0.393***	0.394***	0.394***
Foreigners (2009-13)	(0.044)	(0.043)	(0.044)	(0.044)	(0.043)
Table 8 × Male	0.050	0.054	0.049	0.049	0.048
	(0.049)	(0.050)	(0.050)	(0.050)	(0.050)
Table 9: Personal loans	0.130***	0.146***	0.158***	0.159***	0.159***
Foreigners (2009-13)	(0.021)	(0.023)	(0.024)	(0.024)	(0.024)
Table 9 × Male	0.047	0.058	0.056	0.055	0.054
	(0.024)	(0.026)	(0.026)	(0.026)	(0.027)
Table 10: Personal loans	0.117***	0.135***	0.148***	0.148***	0.147***
Non-client (2010-13)	(0.028)	(0.032)	(0.034)	(0.034)	(0.034)
Table 10 × Male	−0.002	−0.002	−0.004	−0.005	0.002
	(0.032)	(0.036)	(0.037)	(0.037)	(0.037)
Table 11: Mortgage loans	0.122***	0.122***	0.134***	0.134***	0.135***
Client (2010-13)	(0.010)	(0.011)	(0.012)	(0.012)	(0.012)
Table 11 × Male	0.012	0.013	0.011	0.011	0.010
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Table 12: Mortgage loans	0.038***	0.038**	0.043**	0.043**	0.045**
Non-client (2010-13)	(0.019)	(0.019)	(0.021)	(0.021)	(0.022)
Table 12 × Male	0.004	0.004	0.004	0.004	0.002
	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)

Continued on next page

Table 3: (Continued)

	(1)	(2)	(3)	(4)	(5)
Table 13: Personal loans	0.034***	0.033***	0.037***	0.037***	0.037***
Client (2011-13)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Table 13 × Male	0.005	0.005	0.004	0.004	0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age			0.029***	0.053***	0.038***
			(0.008)	(0.009)	(0.009)
Tenure				−0.092***	−0.042***
				(0.014)	(0.014)
Time effect	No	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	No	Yes
Observations	362861	362861	362861	362861	362834
Pseudo R^2	0.04	0.05	0.05	0.05	0.06

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Basic regression analysis with adjusted score

	(1) All years	(2) All years	(3) All years	(4) All years	(5) All years	(6) All years	(7) Before 2009
Male	0.017***	0.025***	0.023***	0.020***	0.021***	0.020***	0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Adjusted score			−0.061***	−0.061***	−0.061***	−0.061***	−0.062***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age				0.027***	0.048***	0.032***	0.031***
				(0.008)	(0.009)	(0.009)	(0.009)
Tenure					−0.080***	−0.041***	−0.043***
					(0.013)	(0.014)	(0.014)
Time effect	No	Yes	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	No	No	Yes	Yes
Observations	362898	362898	362898	362898	362898	362871	338978
Pseudo R^2	0.00	0.03	0.10	0.10	0.10	0.10	0.10

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Robustness checks

	(1) All years	(2) All years	(3) All years	(4) All years	(5) Before 2009
Male	0.017***	0.017***	0.015***	0.014***	0.015***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Adjusted Score	−0.067***	−0.067***	−0.066***	−0.067***	−0.068***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.051***	0.048***	0.026***	0.007	0.006
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
Tenure	−0.277***	−0.278***	−0.338***	−0.292***	−0.322***
	(0.088)	(0.088)	(0.090)	(0.092)	(0.099)
Tenure × Male	0.200**	0.192**	0.219**	0.207**	0.231**
	(0.089)	(0.089)	(0.090)	(0.092)	(0.099)
Hired before 1995		0.003*	0.005***	0.005***	0.005***
		(0.002)	(0.002)	(0.002)	(0.002)
Number of loans			0.018***	0.021***	0.023***
			(0.002)	(0.002)	(0.002)
Time effect	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	Yes	Yes
Observations	362898	362898	362898	362871	338978
Pseudo R^2	0.10	0.10	0.10	0.10	0.10

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: VARIABLES USED IN THE SCORING MODEL

	Mean		Mean
Mortgage	0.34	*Marital status:*	
Personal guarantee	0.66	Married	0.51
Average loan amount	€39,747	Single	0.32
Average maturity	9y 10m	Other	0.12
Loan to value	64	Common law	0.05
Monthly mortgage payment over 6-month balance	131	*Work contract type:*	
		Temporary	0.12
Credit over wealth	46	Permanent	0.60
Leverage	30	Self-employed	0.14
Interest rate	7.04	Retired	0.10
Bank client	0.83	Unemployed	0.04
Foreign client	0.12	*Destination of the loan:*	
Age	41	Family home	0.51
Average years as a bank client	12	Personal	0.15
Average number of years at the current job	5	Others	0.34
Average bank balance over 6-months	€3924.21		
Ratio of 6m/12m bank balance	81.18		
Average number of years of continuous relationship with the bank	7.70		

Table 7: Regression analysis with decomposed score

	(1) All years	(2) All years	(3) All years	(4) All years	(5) All years	(6) All years	(7) Before 2009
Male	0.013***	0.018***	0.017***	0.017***	0.016***	0.015***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age			0.014	0.029***	0.018*	0.019**	0.022**
			(0.008)	(0.009)	(0.009)	(0.009)	(0.010)
Tenure				−0.061***	−0.026*	−0.026*	−0.028*
				(0.014)	(0.014)	(0.014)	(0.015)
Adjusted Score						−0.031***	−0.031***
						(0.001)	(0.001)
Time effect	No	Yes	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	No	No	Yes	Yes
Observations	329996	329996	329996	329996	329979	329979	316889
Pseudo R^2	0.12	0.14	0.14	0.14	0.15	0.15	0.15

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors in parentheses. Controls include characteristics of the borrower and the loan: age, marital status, job contract type, original interest rate, destination of the loan, leverage ratio, debt over wealth, loan to value ratio, monthly mortgage payment over 6-month average bank account balance, nationality, number of years at the current job, average bank balance over 6 months, 6/12 month bank balance ratio, an indicator for if the individual is a bank client or not, number of years of continuous relationship with the bank, and number of years as a bank client.

Table 8: Basic regression analysis with delinquency after 60 days

	(1) All years	(2) All years	(3) All years	(4) All years	(5) All years	(6) All years	(7) Before 2009
Male	0.019***	0.029***	0.026***	0.023***	0.024***	0.022***	0.023***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Adjusted Score			−0.080***	−0.081***	−0.081***	−0.081***	−0.081***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age				0.036***	0.057***	0.043***	0.048***
				(0.009)	(0.010)	(0.010)	(0.011)
Tenure					−0.081***	−0.042***	−0.047***
					(0.015)	(0.016)	(0.016)
Time effect	No	Yes	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	No	No	Yes	Yes
Observations	362898	362898	362898	362898	362898	362871	338978
Pseudo R^2	0.00	0.03	0.09	0.09	0.09	0.10	0.09

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Basic regression analysis with delinquency after 30 days

	(1) All years	(2) All years	(3) All years	(4) All years	(5) All years	(6) All years	(7) Before 2009
Male	0.018***	0.031***	0.027***	0.024***	0.024***	0.023***	0.025***
	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Adjusted Score			−0.106***	−0.106***	−0.106***	−0.106***	−0.106***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age				0.040***	0.057***	0.051***	0.060***
				(0.011)	(0.012)	(0.012)	(0.013)
Tenure					−0.064***	−0.034*	−0.049**
					(0.018)	(0.018)	(0.019)
Time effect	No	Yes	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	No	No	Yes	Yes
Observations	362898	362898	362898	362898	362898	362898	338996
Pseudo R^2	0.00	0.02	0.09	0.09	0.09	0.09	0.09

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Two-way panel data estimation

	(1) All years	(2) All years	(3) All years	(4) All years	(5) Before 2009
Male	0.012***	0.011***	0.011***	0.013***	0.014***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Adjusted Score	−0.055***	−0.055***	−0.055***	−0.055***	−0.053***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age		0.017	0.022	0.021	0.025
		(0.012)	(0.014)	(0.014)	(0.015)
Tenure			−0.014	0.089	0.092
			(0.021)	(0.100)	(0.113)
Tenure × Male				−0.106	−0.112
				(0.102)	(0.114)
Constant	−0.002***	−0.002***	−0.002***	−0.002***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	362898	362898	362898	362898	338996
R^2	0.04	0.04	0.04	0.04	0.04
Adjusted R^2	0.04	0.04	0.04	0.04	0.04

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Loan approval rate by recommendation before 2009

	(1)	(2)	(3)	(4)
Very positive (A1)	0.996***	0.996***	0.996***	0.996***
	(0.001)	(0.001)	(0.001)	(0.001)
Very positive × Male	0.001	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Positive (A2)	0.988***	0.988***	0.988***	0.989***
	(0.001)	(0.001)	(0.001)	(0.001)
Positive × Male	0.002**	0.002**	0.002**	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Neutral (A3)	0.962***	0.962***	0.962***	0.966***
	(0.003)	(0.003)	(0.003)	(0.003)
Neutral × Male	0.011***	0.011***	0.011***	0.007***
	(0.003)	(0.003)	(0.003)	(0.003)
Negative (D1)	0.895***	0.895***	0.896***	0.912***
	(0.006)	(0.006)	(0.006)	(0.005)
Negative × Male	0.033***	0.033***	0.033***	0.019***
	(0.006)	(0.006)	(0.006)	(0.006)
Very negative (D2)	0.675***	0.675***	0.675***	0.699***
	(0.015)	(0.015)	(0.015)	(0.016)
Very negative × Male	0.058***	0.058***	0.058***	0.041**
	(0.016)	(0.016)	(0.016)	(0.016)
Observations	350518	350518	350518	349739
Pseudo R^2	0.27	0.27	0.27	0.28

Dependent variable: loan approval rate. Standard errors in parentheses.
Specification (1) controls for score and time, (2) controls for score, time, and age, (3) controls for score, time, age, and tenure, and (4) controls for score, time, age, tenure, and geography.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Loan Override Rate Before 2009

	(1)	(2)	(3)	(4)
Negative recommendation	0.849***	0.848***	0.850***	0.849***
	(0.006)	(0.006)	(0.006)	(0.006)
Negative recom. × Male	0.051***	0.052***	0.051***	0.051***
	(0.007)	(0.007)	(0.007)	(0.007)
Very negative recommendation	0.690***	0.690***	0.691***	0.692***
	(0.012)	(0.012)	(0.012)	(0.012)
Very negative recom. × Male	0.059***	0.059***	0.058***	0.056***
	(0.012)	(0.012)	(0.012)	(0.012)
Reject recommendation	0.801***	0.798***	0.800***	0.799***
	(0.006)	(0.006)	(0.006)	(0.006)
Reject recommendation × Male	0.055***	0.055***	0.054***	0.053***
	(0.006)	(0.006)	(0.006)	(0.006)
Observations	40589	40589	40589	40589
Pseudo R^2	0.31	0.31	0.31	0.32

Dependent variable: loan override rate. Standard errors in parentheses.
Specification (1) controls for score and time, (2) controls for score, time, and age, (3) controls for score, time, age, and tenure, and (4) controls for score, time, age, tenure, and geography.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Delinquency rate by recommendation before 2009

	(1)	(2)	(3)	(4)
Very positive (A1)	0.038***	0.038***	0.038***	0.038***
	(0.002)	(0.002)	(0.002)	(0.002)
Very positive × Male	−0.002	−0.002	−0.002	−0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Positive (A2)	0.104***	0.104***	0.104***	0.104***
	(0.003)	(0.003)	(0.003)	(0.003)
Positive × Male	0.011***	0.011***	0.011***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Neutral (A3)	0.139***	0.139***	0.139***	0.138***
	(0.006)	(0.006)	(0.006)	(0.006)
Neutral × Male	0.009	0.009	0.009	0.010
	(0.006)	(0.006)	(0.006)	(0.006)
Negative (D1)	0.157***	0.157***	0.157***	0.155***
	(0.009)	(0.009)	(0.009)	(0.009)
Negative × Male	0.029***	0.029***	0.029***	0.030***
	(0.009)	(0.009)	(0.009)	(0.009)
Very negative (D1)	0.170***	0.170***	0.169***	0.167***
	(0.015)	(0.015)	(0.015)	(0.015)
Very negative × Male	0.071***	0.071***	0.071***	0.072***
	(0.016)	(0.016)	(0.016)	(0.016)
Observations	338994	338994	338994	338976
Pseudo R^2	0.10	0.10	0.10	0.08

Standard errors in parentheses.

Specification (1) controls for score and time, (2) controls for score, time, and age, (3) controls for score, time, age, and tenure, and (4) controls for score, time, age, tenure, and geography.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Duration Model: Performance and Employment

	(1) Proportional	(2) Weibull	(3) Gompertz	(4) Gamma
Male	-0.420^{***}	-0.430^{***}	-0.424^{***}	0.224^{***}
	(-3.34)	(-3.33)	(-3.45)	(3.33)
Age	-0.022^{***}	-0.021^{***}	-0.022^{***}	0.011^{***}
	(-3.02)	(-2.92)	(-2.97)	(2.60)
Bad loans	0.029^{***}	0.028^{***}	0.029^{***}	-0.014^{***}
	(5.95)	(5.62)	(5.92)	(-4.45)
Constant		-6.046^{***}	-3.476^{***}	3.152^{***}
		(-17.99)	(-11.17)	(15.05)
Observations	17100	17100	17100	17100
p		1.928		
γ			.033	
σ				.514

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Duration Model with Interaction: Performance and Employment

	(1) Proportional	(2) Weibull	(3) Gompertz	(4) Gamma
Male	-0.175	-0.183	-0.157	0.0949
	(-1.09)	(-1.07)	(-1.02)	(1.08)
Age	-0.022^{***}	-0.021^{***}	-0.021^{***}	0.011^{**}
	(-2.96)	(-2.87)	(-2.89)	(2.52)
Bad loans	0.070^{***}	0.070^{***}	0.075^{***}	-0.036^{***}
	(4.61)	(4.21)	(5.27)	(-4.18)
Bad loans $\times$ Male	-0.042^{***}	-0.042^{**}	-0.047^{***}	0.022^{**}
	(-2.64)	(-2.44)	(-3.11)	(2.46)
Constant		-6.294^{***}	-3.764^{***}	3.299^{***}
		(-17.69)	(-11.46)	(14.79)
Observations	17100	17100	17100	17100
p		1.923		
γ			.033	
σ				.513

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16: DURATION MODEL WITH INTERACTION: PERFORMANCE AND EMPLOYMENT

	(1) Proportional	(2) Weibull	(3) Gompertz	(4) Gamma
Male	-0.195	-0.205	-0.185	0.107
	(-1.15)	(-1.14)	(-1.14)	(1.15)
Age	-0.020^{**}	-0.018^{**}	-0.020^{**}	0.0093^{**}
	(-2.54)	(-2.37)	(-2.55)	(2.08)
Bad loans	0.075^{***}	0.075^{***}	0.080^{***}	-0.039^{***}
	(4.91)	(4.44)	(5.62)	(-4.42)
Bad loans $\times$ Male	-0.044^{***}	-0.044^{**}	-0.048^{***}	0.0231^{**}
	(-2.68)	(-2.43)	(-3.12)	(2.46)
Average loans	0.005^{*}	0.004	0.006^{**}	-0.002
	(1.79)	(1.60)	(2.06)	(-1.64)
Constant		-6.426^{***}	-3.922^{***}	3.409^{***}
		(-17.32)	(-11.40)	(14.11)
N	17076	17076	17076	17076
p		1.9		
γ			.033	
σ				.515

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix

Table A.1: ROBUSTNESS CHECKS WITH DECOMPOSED SCORE

	(1) All years	(2) All years	(3) All years	(4) All years	(5) Before 2009
Male	0.016***	0.016***	0.015***	0.015***	0.014***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age	0.029***	0.030***	0.017*	0.019**	0.021**
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Tenure	−0.078	−0.078	−0.059	−0.058	−0.084
	(0.091)	(0.091)	(0.093)	(0.092)	(0.097)
Tenure × Male	0.018	0.018	0.031	0.031	0.054
	(0.092)	(0.092)	(0.094)	(0.093)	(0.098)
Hired before 1995		−0.000	0.001	0.000	0.000
		(0.002)	(0.002)	(0.002)	(0.002)
Adjusted Score				−0.027***	−0.027***
				(0.001)	(0.001)
Time effect	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	Yes	Yes	Yes
Observations	329996	329996	329979	329979	316889
Pseudo R^2	0.14	0.14	0.15	0.15	0.15

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors in parentheses. Controls include characteristics of the borrower and the loan: age, marital status, job contract type, original interest rate, destination of the loan, leverage ratio, debt over wealth, loan to value ratio, monthly mortgage payment over 6-month average bank account balance, nationality, number of years at the current job, average bank balance over 6 months, 6/12 month bank balance ratio, an indicator for if the individual is a bank client or not, number of years of continuous relationship with the bank, and number of years as a bank client.

Table A.2: Robustness check with delinquency after 60 days

	(1) All years	(2) All years	(3) All years	(4) All years	(5) Before 2009
Male	0.017***	0.017***	0.015***	0.014***	0.015***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Adjusted Score	−0.074***	−0.074***	−0.074***	−0.074***	−0.074***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.055***	0.049***	0.021**	0.003	0.005
	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)
Tenure	−0.394***	−0.396***	−0.472***	−0.454***	−0.464***
	(0.098)	(0.097)	(0.099)	(0.101)	(0.109)
Tenure × Male	0.319***	0.303***	0.338***	0.352***	0.350***
	(0.098)	(0.098)	(0.100)	(0.102)	(0.109)
Hired before 1995		0.005***	0.008***	0.008***	0.009***
		(0.002)	(0.002)	(0.002)	(0.002)
Number of loans			0.023***	0.028***	0.031***
			(0.002)	(0.002)	(0.002)
Time effect	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	Yes	Yes
Observations	362896	362896	362896	362896	338976
Pseudo R^2	0.09	0.09	0.10	0.10	0.10

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Robustness check with delinquency after 30 days

	(1) All years	(2) All years	(3) All years	(4) All years	(5) Before 2009
Male	0.014***	0.015***	0.012***	0.011***	0.013***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Adjusted Score	−0.097***	−0.097***	−0.097***	−0.097***	−0.097***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.055***	0.047***	0.018	0.006	0.014
	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)
Tenure	−0.484***	−0.487***	−0.569***	−0.585***	−0.583***
	(0.111)	(0.111)	(0.113)	(0.114)	(0.123)
Tenure × Male	0.429***	0.405***	0.443***	0.483***	0.459***
	(0.112)	(0.112)	(0.113)	(0.115)	(0.123)
Hired before 1995		0.007***	0.010***	0.009***	0.010***
		(0.002)	(0.002)	(0.002)	(0.002)
Number of loans			0.025***	0.033***	0.036***
			(0.003)	(0.003)	(0.003)
Time effect	Yes	Yes	Yes	Yes	Yes
Zip code controls	No	No	No	Yes	Yes
Observations	362896	362896	362896	362896	338994
Pseudo R^2	0.09	0.09	0.09	0.09	0.09

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$