Gender, Credit, and Firm Outcomes

Abstract

Small and micro enterprises are usually majority-owned by entrepreneurs. Using a unique

sample of loan applications from such firms, we study the role of owners' gender in bank credit

decisions and post-credit-decision firm outcomes. We find that, ceteris paribus, female

entrepreneurs are more prudent loan applicants than are males, since they are less likely to apply

for credit or to default after loan origination. The relatively more aggressive behavior of male

applicants pays off, however, in terms of higher average firm performance after loan origination.

Keywords: Gender; loan applications; bank credit decisions; firm performance
JEL Classification: G21; G32; J16

I. Introduction

Are male entrepreneurs more aggressive credit applicants compared to female entrepreneurs? Do banks treat male and female entrepreneurs differently? Are there any differences in the performance of female- and male-owned firms following bank credit decisions (loan origination or rejection)? These questions are important for our understanding of the role of gender in corporate finance and banking. If gender indeed plays a role in entrepreneurial decisions to apply for credit, this phenomenon can trigger a sequence of events at the firm level, ultimately affecting firm performance and real economic outcomes. These effects are especially important for small firms that do not usually have recourse to alternative sources of conventional finance outside of private equity (internal or external) and bank credit (Berg, 2018) and, as highlighted by the European Commission (EC) (2014) and the Organization for Economic Cooperation and Development/European Union (OECD/EU) (2017), female-owned firms are on average smaller and more likely to have single ownership.

We identify entrepreneurs as majority owners of small and micro firms, following the relevant definition of the European Commission (total assets less than €10 million). Our first two questions concern whether male entrepreneurs are more aggressive credit applicants than are female entrepreneurs, and whether such aggression is the result of differential loan demand or supply between genders. Research backing the loan-demand premise suggests that debt aversion mostly characterizes female entrepreneurs because of higher risk aversion than among males (Carter, Shaw, Lam, and Wilson (2007), Dawson and Henley (2015)). Additional reasons include different motivations for female entrepreneurs, especially a higher demand for autonomy and a lower inclination toward firm growth. In turn, supply-side premises note that, in their credit decisions, banks value elements such as female-owned firm underperformance, differential levels

of education between male and female entrepreneurs, and the fact that female-owned firms are typically smaller and younger. Other studies suggest that loan officers discriminate against female entrepreneurs, especially when the former are male (Qi, Ongena, and Cheng (2019)). Recent evidence does not fully back up these propositions (for a recent review, see Dean, Larsen, Ford, and Akram (2019)), however, while the literature identifies such effects to be more pronounced in less developed economies (Asiedu, Kalonda-Kanyama, Ndikumana, and Nti-Addae (2013), Ongena and Popov (2016)).

Future performance indicators such as future profitability and the probability of default might be affected if female-owned businesses were less likely to apply for credit or if access to credit supply were restricted. Concerning the probability of firm default (after a bank's credit decision), the effects can go either way. On the one hand, obtaining more credit implies higher leverage. If this results from less prudent behavior by male entrepreneurs, or if the additional credit is not directed toward purely productive ends, however, the probability of default for male-owned firms will be higher. On the other hand, if additional credit to male-owned firms implies profitable investments, then future profitability ratios will be higher and the corresponding probability of default lower. Our study is the first to pinpoint the effects of the differential likelihood of loan application and bank credit decisions (loan origination or rejection) on the default probability and performance of male- vs. female-owned firms.

We empirically answer our research questions using data on loan applications to a large (systemic) European bank with nationwide coverage. For each loan application, we have full information on several applicant characteristics (gender, income, wealth, family, age, education, etc.); firm characteristics for which the applicant is majority owner (including the firm's financial characteristics and region); loan characteristics (e.g., loan amount, maturity, collateral,

purpose); and the bank's loan decision (approved or rejected). Importantly, we have access to the applicant's credit score on which the bank based its credit decision. We also know whether the applicant has an exclusive relationship with the bank: Applicants with such a relationship are credit constrained (even from other conventional banks) if a specific bank rejects their application. Using these data and repeated loan applications by the same applicants, we construct a panel dataset of loan applicants over the period 2002-2017 (we observe the same applicant over time).

To determine whether male applicants are more likely than females to apply for a loan, we use a model that includes the applicant's credit score. We show that this is an important variable, encompassing information on several characteristics observed by bank loan officers (soft information) that cannot be captured by the relevant hard information disclosed in the loan application. We find that the probability that male entrepreneurs apply for credit is 0.8 percent higher than that for female entrepreneurs. The results are quite similar (the estimate is 1.2 percent) when using an instrumental-variable (IV) method to overcome potential endogeneity concerns (mainly due to the omitted-variable bias). Our IV represents the average, by industry, region, and year, for female entrepreneurs 15 years before a loan application. Overall, the effect of gender on the probability of loan application, even though statistically significant, is economically moderate.

We next examine whether there are differences in the probability of loan origination (the supply-side premise). When controlling for the probability of loan application (estimated under our first research question), we find the effect of gender to be statistically insignificant. Similarly, we find that gender does not significantly affect the loan amount and spread of originated loans.

4

Given these results, we next examine whether rejected male applicants reapply for loans sooner (within one or two years post a bank's credit decision) compared to female applicants. In an interesting finding for the whole lending process, we find that this is indeed the case: Our IV estimations show that female applicants are 3.5 percent (4.2 percent) less likely to reapply within one (two) year(s) post loan rejection.

Having established that male entrepreneurs are somewhat more active in seeking credit from their bank, especially when they have been rejected, we seek to answer the question as to whether this behavior pays off in terms of future firm performance. We examine the role of credit demand (given that the role of credit supply is insignificant) in the nexus between gender and firm performance. To this end, the bank's credit score creates a sharp discontinuity between applicants whose loans are approved and those rejected (or those subject to additional review at a later stage). Thus, our identification approach relies on a regression discontinuity design (RDD), with the credit score being the assignment variable (Berg (2018)). This implies identification from comparing changes in measures of future firm performance of approved and denied applicants with credit scores around the cutoff prior to the bank's credit decision. We estimate the RDD separately for male and female applicants and our preferred specification relies on standard nonparametric techniques (Cattaneo, Titiunik, and Vazquez-Bare (2016); Delis, Fringuellotti, and Ongena (2019)). All relevant tests suggest that loan applicants cannot manipulate their credit scores, and graphical evidence clearly points to a sharp discontinuity.

Our results suggest that male loan applicants' firms have a higher default probability three years after bank credit decisions (by approximately 2.1 percent) compared to those of female loan applicants. Furthermore, male applicants are more aggressive in reapplying for a loan once rejected. Specifically, rejected male applicants are three percent (4.2 percent) more

likely to apply for a loan in the one (two)-year window post the original rejection (compared to rejected female applicants). Given this finding, we also show that firms of male applicants have higher leverage levels three years after bank credit decisions. These findings are consistent with the premise that female entrepreneurs are more prudent loan applicants.

An important question arising from these results is whether and how the tradeoff between the additional credit and higher probability of firm default for male entrepreneurs affects firm performance. We show that, for female-owned firms, positive bank credit decisions increase firms' return on assets (ROA) three years onward by approximately 29 percent (compared to the average ROA of female-owned firms). For male-owned firms, however, the equivalent increase is approximately 38 percent. We show that most of this nine-percent difference in forward ROA is indeed due to male entrepreneurs' aggressiveness in applying for credit. Specifically, when controlling for the probabilities of loan application and firm default (as estimated within our previous analysis), the effect of gender on forward ROA is minimal.

The key implications drawn from our results are as follows: Female entrepreneurs are somewhat more prudent loan applicants; the lower probability of their applying for credit pays off in terms of lower firm default after loan origination. However, the relatively more aggressive behavior of male applicants pays off in terms of higher average firm performance after loan origination, a mechanism that explains a large part of the difference in performance between male- and female-owned firms. This difference is negligible when controlling for the probability of applying for credit and of firm default.

Our research relates to two broad and interrelated strands of literature on the role of gender in financial access and firm performance. The first strand aims to identify the reasons behind gender-related differential access to finance. Muravyev, Talavera, and Schäfer (2009), Alesina,

6

Lotti, and Mistrulli (2013), Bellucci, Borisov, and Zazzaro (2010), Beck, Behr, and Guettler (2012), and Ongena and Popov (2016) provide evidence for gender discrimination in bank credit decisions, while other studies suggest that discrimination is very limited if present at all (Asiedu et al. (2013)), especially in more developed economies (Delis and Papadopoulos (2018)). Other studies suggest that the nature of female-owned firms (smaller firms with lower startup capital) causes these differences, which might be attributed to differences in risk-taking incentives, financial motivations (especially shying away from competition), education, connections, and autonomy (Alsos, Isaksen, and Ljunggren (2006), Niederle and Vesterlund (2007), Roper and Scott (2009), Alesina, Lotti, and Mistrulli (2013), Piacentini (2013), and Fang and Huang (2017)). In contrast, several studies suggest that such effects are due to external unobserved factors, and that similarities between the sexes are more important than are their differences (Ahl (2006); Dean, Larsen, Ford, and Akram (2019)).

The second and even more general strand of literature considers how gender affects firm performance (for a recent review, see Dean et al., 2019). Earlier studies (e.g., Baker, Aldrich, and Liou (1997)) allege that female attributes are key to the understanding of the so-called "underperformance hypothesis" of female-owned firms, which emanates from stylized facts showing a gender gap in entrepreneurship (EU (2014), OECD/EU (2017)). However, recent studies using better data and identification techniques challenge this view (Coleman (2007), Conroy and Weiler (2016)). In the finance literature, a natural laboratory for the gender-firm performance relationship is composition of firms' or banks' boards (e.g., Huang and Kisgen (2013), Schwartz-Ziv (2017)). Several studies show a positive effect of gender diversity on the board on firm performance (e.g., Chen, Leung, and Goergen (2016), and references therein). Other studies, however, suggest that gender diversity does not necessarily imply better

performance (Adams and Ferreira, 2009; Sila, Gonzalez, and Hagendorff (2016), Schwartz-Ziv (2017)). Recently, Karavitis, Kokas, and Tsoukas (2019) show that board gender diversity lowers corporate loan spreads. Our study relates to these general strands of literature and draws on them in many theoretical and empirical respects, but is also the first to pinpoint and analyze all steps in the role of bank credit in the nexus between gender and firm performance.

The rest of the paper proceeds as follows. Section 2 describes our data set. Section 3 discusses the empirical analysis, stressing our identification approaches for the different questions, and analyzes the empirical results. Section 4 concludes the paper.

II. Data

We source our data from a major European bank, for which we have access to its full loan portfolio between 2002 and 2017. We use data for the bank's loans to domestic small and micro firms (total assets of up to €10,000, per the EU definition).[1] Using data from a single bank is common practice when detailed data are required (e.g., Adams, Einav, and Levin (2009), Iyer and Puri (2012), Berg (2018)). The bank from which we obtain the data is a systemically important financial institution, according to the definition of the European Banking Authority (EBA).[2] The bank operates on a global scale and provides credit to all business types.

We restrict our sample to small and micro enterprises as defined above because we require that loan applicants, whether male or female, are majority owners (own more than 50 percent) of the firm. We consider all loan types, including working capital loans, real estate

---

[1] A similar dataset is used by Delis et al. (2019).
[2] This includes the EBA definitions for the Global Systemically Important Institutions (G-SIIs) and the Other Systemically Important Institutions (O-SIIs). See, https://eba.europa.eu/risk-analysis-and-data/global-systemically-important-institutions and https://eba.europa.eu/risk-analysis-and-data/other-systemically-important-institutions-o-siis-.

loans, venture loans for startups, lines of credit, and others.[3] For our end sample of loan applications, we obtain information on several applicant characteristics, including gender, income, wealth, education, age, marital status, credit score by the bank, and number of dependents. We also have a full array of firm characteristics such as size, leverage, ROA, liquidity, and the firm's region and industry. At the loan level, besides being aware of the bank's credit decision (approved or denied loans), we know the loan amount and maturity applied for, as well as loan characteristics when originated (i.e., whether the loan has performance pricing provisions and/or is secured with collateral). Using this information, our sample consists of 95,340 loan applications between 2002 and 2017. We define all the variables in Table 1.

[Please insert Table 1 about here]

Two important sources of available information are the credit score assigned by a bank to each applicant, and the exclusivity of the relationship between the applicant and the bank. The credit score contains all information (both hard and soft) for the bank's credit decision. Hard information refers here to all the information on paper, i.e., both firm and applicant characteristics on the application files (i.e., the relevant variables listed in Tables 1 and 2). Soft information refers to the residual, i.e., an explanation for the credit score that is not on paper (e.g., the bank's and loan officer's perception of the firm, the applicant, and the investment idea; the strength of the bank-firm relationship, and other such ties). For a credit score above a given cutoff, the bank approves (originates) a loan; for a credit score below this cutoff, the bank denies the loan or suggests reexamination at a later date. We are not allowed to expose the precise cutoff and thus we normalize it to 0. Thus, in our empirical analysis, the bank originates the loan if the credit score is equal to or larger than 0, and the bank denies the loan otherwise.

---

[3] Distinguishing between loan types has no effect on any of our empirical inferences.

Our bank knows exactly which firms apply for loans to other regulated and supervised banks (by the EBA or the country's regulatory and supervisory authority), irrespective of whether these firms have had a loan approved or rejected by our bank. Our bank also knows the timing of these applications and their outcome and obtains this information from both the firms and the country's credit register. This represents the essence of the exclusive relationship between the firms and our bank.[4] For small firms, having an exclusive relationship with a bank is fairly common and is the case for 65 percent of the firms in our full sample. Using summary statistics from previous studies on multiple or exclusive lending relationships, we note that Berger et al. (2011) document a 71-percent exclusive relationship between banks and SMEs in three European countries (Germany, Italy, and the UK), while this is less often the case in the U.S. (Berger, Goulding, and Rice (2014) document a 57-percent rate).[5] For most applicants, we observe more than one loan application during our sample period.

From the 95,340 loan applications and the bank's continuing information for the applicant and firm characteristics, we construct a balanced panel of applicant (firm)-year observations. We need to maintain a balanced panel to observe important firm and applicant characteristics over the full sample period. We thus discard loans from applicants establishing a relationship with the bank in the middle of our period, as well as applicants (both accepted and

---

[4] The bank is naturally unaware of any loan applications to the shadow banking system that was largely unregulated during the sample period. In the most comprehensive available report, Kaya and the Deutsche Bank (2014) suggest that, in Europe, competition that banks face from shadow banks in SME lending is fairly small (European shadow banks usually engage in other activities). Furthermore, even if SMEs access the shadow banking sector, the lending terms would be completely different (more unfavorable). Thus, we should exclude shadow bank SME lending from our analysis even if we had this information.

[5] It is hard to find much more evidence precisely on whether (small) firms have one or more banking relationship(s) in north European countries. Farinha and Santos (2002) report similar statistics for Portugal (70% of firms with fewer than 10 employees have one bank relationship). More recently, Bonfim at al. (2018) report a mean value of 2 banks for small Portuguese firms, but the Portuguese banking sector is much less concentrated than that in our bank's country. Essentially, the available evidence suggests that the percentage of exclusive relationships in our sample is comparable to that in previous papers on relationship banking.

rejected) that never reapply for loans. Essentially, all individuals reapply for loans within a four-year period. In other words, all observed firms have a relationship with the bank from 2004 onward (the bank has information for applicants from 2002 onward).[6]

This panel includes 357,056 observations. The reason the panel has more observations than the number of loans is simply that firm owners do not apply for a loan every year. However, the bank continues to holds information on applicant characteristics after the loan application because a new application takes place in the future and the bank requests information on applicants' income and wealth in the past, which is what allows us to generate the panel of loan applicants. For education and marital status, we observe some changes from year to year, and when we do not know the precise year of the change (i.e., there is no loan application in two consecutive years), we assume that a change occurs in the middle of the time interval between the two loan applications. This assumption does not affect our main results. We also fill in the observations with the last credit score calculated by the bank. Thus, if there is a loan application in year t but not one in year $t+1$, we assign in year $t+1$ the credit score in year t.[7]

We report summary statistics in Panel A of Table 2. For the applicant and firm characteristics, we have a balanced panel of 357,056 applicant (or firm)-year observations and 95,340 loan applications, from which there are 79,470 originated loans. This number represents approximately 83 percent of loan applications and is somewhat lower than that reported in the Survey of Access to Finance of Enterprises (SAFE), which also includes the relatively safer medium-sized firms. Applications by female entrepreneurs amount to approximately 20 percent

---

[6] This comes at the expense of dropping observations from respective applications that do not fulfill our criteria. Running our analysis on an unbalanced sample does not affect our inferences and in fact strengthens the results only in the cases where these are statistically significant. However, using an unbalanced panel implies that we do not have important dynamic information on certain applicant characteristics (especially income, wealth, and changes in family status) and an observed exclusive bank-firm relationship.

[7] Again, sliding this forward to $t+2$ or $t+3$, no new loan application in this time interval does not affect our results.

of total loan applications. Furthermore, the mean applicant is close to having tertiary education, is approximately 44 years old, and has slightly fewer than two dependents.

[Please insert Table 2 about here]

Panels B and C of Table 2 report the equivalent summary statistics separately for female and male applicants. The statistical difference is significant for income, wealth, firm size, firm leverage, firm cash, and the number of previous loan applications. There is also a significant difference in the credit score in favor of male applicants (pointing to a significant gender gap). In contrast, the two groups are not statistically different in terms of education, age, number of dependents, and (interestingly) firm ROA.

We run four checks to establish that the bank and firms in our sample have similar characteristics with other systemic European banks and other small European firms, respectively. First, we collect Compustat data for the financial statements of the 32 systemically important banks (according to the EBA's definition) for the period 1985-2018. Using these data, we compare the annual means of important bank characteristics of these banks with the respective characteristics of our bank. Specifically, we use the ratio of liquid assets (cash plus short-term securities) to total assets, the ratio of market value to book value, and profitability (before tax return on assets). In Figures 1a to 1c, we report scatterplots and regression lines for our bank's values (y-axes) against the systemic banks' averages (x-axes). The correlations are all positive and highly statistically significant (at the one-percent level). Thus, the bank used in our analysis has similar characteristics over time with other systemic European banks.

We next move to an analysis of credit standards, using data from the SAFE. It is possible to compare the Eurozone average rejection rate with that of our bank because we obtain the definition of small and micro enterprises and the definition of rejection from the European

12

Union. Figure 2 graphs the eurozone average rejection rate (available for half-years since 2010) against the rejection rate from our bank. The correlation for these 19 observations is quite high (equal to 0.86) with our bank being somewhat more stringent than the Eurozone average during the European crisis (2010-2014) and slightly less stringent from 2015 onward.

Third, the characteristics of small and micro enterprises in our sample are very similar to others in Europe, given that we strictly follow the European definition of small and micro firms. Figures 3a and 3b plot the annual average leverage and profitability ratios of small and micro firms in several core European countries (Austria, Belgium, Denmark, France, Germany, and the Netherlands) against the averages in our sample. The data for these countries are from Orbis (information is only available since 2008). Again, the comparisons reflect important similarities: On average, the firms in our sample have a 1.1 percent lower leverage ratio and a 0.76 percent higher ROA. These small differences most likely exist because our bank is based in one of the highest-income European countries and was not significantly affected by the economic downturn of 2010-2014.

Fourth, approximately 20 percent of the loan applicants in our sample are female entrepreneurs. Females represent 25 percent of the total number of entrepreneurs in our bank, which is close to the 29 percent reported for Europe in 2012 (European Commission (2014)). The discrepancy probably occurs because our sample begins in 2002, where we report lower entrepreneurship rates for females (in 2012, the average in our sample is 28.3 percent). Piacentini (2013) suggests that the share of female employers was stable at around 25 percent throughout the 2000s. Finding equivalent data for loan applications by female entrepreneurs across Europe is a big challenge. According to Stefani and Vacca (2013), SAFE data in 2010 shows that 25 percent (23 percent) of male (female) entrepreneurs applied for bank loans in the past six months. For the

same period, the equivalent statistics in our sample are 27 percent and 24 percent. According to

Piacentini (2013), about 30 percent of women-led startups use bank credit; in our sample, this

figure is 34 percent. We also document similar differences in the earnings of female and male

entrepreneurs with those reported by OECD (2017). All in all, our sample has characteristics

similar to aggregate data obtained for European countries (especially those with high income

levels).

[Please insert Figures 1 - 3 about here]

III. Empirical Analysis

A. Gender and the Probability of Loan Application

We first study gender differences in loan-application probability using the full sample of 357,056

individual-year observations. With this analysis, we aim to examine gender differences related to

the probability of seeking bank credit, conditional on observed individual and firm

characteristics. We estimate the following model:

(1) $\quad \text{APPLY}_{it} = a_0 + a_1 \text{GENDER}_i + a_2 x_{i(f)t} + u_{it},$

where APPLY is a binary variable, taking the value 1 if an individual *i* in our sample applies for

a loan at year *t* (and 0 otherwise),[8] GENDER is a binary variable equal to 1 for male applicants

---

[8] Thus, we assign a value 1 to the 95,340 observations for which individuals applied for a loan in a specific year during our sample period and a value 0 to the rest of the observations in the sample of 357,056 (individuals in our sample who did not apply for a loan in a specific year).The best alternative for a control group is all individuals who do not apply (i.e., firms with valuable investment opportunities). Identifying these firms is of course impossible unless there are specific demand-side survey data (which have other limitations such as the non-inclusion of credit scores). A more viable alternative would be to use all small and micro firms in the country (data from Amadeus or Orbis). However, this perplexes identification because bank-firm relationships are not easy to break and firms might choose banks on the basis of such relationships, irrespective of other economic incentives. Nevertheless, we run a search in Orbis but find only 64 firms in the bank's country using the EU definition of a maximum turnover of 10 million euros. Using larger firms is not an option because most of these are not majority-owned by specific individuals.

and equal to 0 for female applicants, and x is a vector of control variables reflecting individual (i) or firm (f) characteristics. All specifications include regional and year fixed effects.[9]

An important element of our identification approach is that the credit score not only contains information on observed but also on several unobserved characteristics (e.g., soft information), which are nonetheless known to the individual and to the bank. This information is important because the inclusion of these characteristics within the credit score affects the probability of a loan application and should considerably limit the effect of related omitted-variable bias on GENDER. Thus, the coefficient on GENDER should better capture the relevant gender-related reasons behind the probability of applying for a business loan.[10]

We report marginal effects from a probit model in column 1 of Table 3. We cluster the standard errors in all our analyses by individual applicants.[11] In addition to our control variables, we also use year and regional dummy variables. We find that female entrepreneurs have a 0.8 percent lower probability of applying for a loan. While the effect is statistically significant, its economic significance is clearly not particularly large. In column 2, we report the equivalent results excluding CREDIT_SCORE. The coefficient on GENDER jumps to 2.7 percent, showing either that the credit score contains valuable information that is erroneously being attributed to gender differences or that the bank discriminates between male and female applicants when constructing the credit score. In the next section, we show that the latter is not the case.

---

[9] We also experiment with industry fixed effects. These do not affect the OLS results but introduce convergence problems in the probit models. Another level of information that might be important for the formation of the credit score and thus the credit decision is the loan officer. Unfortunately, we do not have this information. However, the bank notes that, especially for marginal cases, such as those we mostly rely on in our empirical analysis, the ultimate credit decision is made at the headquarters by the credit department. For this reason, we also consider relevant corporate governance characteristics in robustness tests.

[10] If GENDER were not included in equation (1), CREDIT_SCORE would also capture its effect as part of a soft-information component. That is, the inclusion of GENDER extracts the relevant information from the more general credit-score variable. The same holds for the rest of the controls used in equation (1).

[11] Clustering at a more aggregate level (by region) or double clustering (by individual and year or by region and year) does not affect our inferences.

[Please insert Table 3 about here]

Despite considering the information contained in the credit score, the simple probit model may still suffer from the omitted-variable bias, mainly due to population differences in female and male entrepreneurs.[12] To this end, we use an instrumental-variable (IV) probit model. Our instrument is the average (by industry, region, and year) ratio of female to male entrepreneurs 15 years before the loan application (named AVERAGE_FEMALE_ENTREPRENEURS). For example, the value of the ratio in 1987 enters for loans that originated in 2002. The premise for the use of this IV is that the local shares of female to male entrepreneurs 15 years ago are predetermined and do not directly affect the probability to apply for a loan, conditional on our control variables (including the credit score). By using AVERAGE_FEMALE_ENTREPRENEURS directly in equation (1), we show that this is the case (results in Appendix Table A1), especially when controlling for the applicants' credit score. In contrast, as we show, in the lower part of the tables, that AVERAGE_FEMALE_ENTREPRENEURS strongly correlates with GENDER.

The first-stage results of the two-stage IV probit model in Table 3 indicate a strong correlation between AVERAGE_FEMALE_ENTREPRENEURS and APPLY.[13] Specifically, a one-standard-deviation higher AVERAGE_FEMALE_ENTREPRENEURS is associated with a 2.5 percent higher probability that the loan applicant is a female (statistically significant at the one-percent level). This is intuitive, given that the pre-existence of more female entrepreneurs in a given region, industry, and year yields a higher probability that the loan applicant is a female

---

[12] We are less concerned here with other sources of omitted-variable bias stemming from individual characteristics of female vs. male loan applicants (e.g., risk-taking incentives). We precisely leave these differences to be captured by GENDER.

[13] Using probit in both stages of the model improves econometric efficiency because the outcome variables in both stages are binary.

entrepreneur. Importantly, the effect of GENDER on APPLY shows that female entrepreneurs have a 1.2 percent lower probability of applying for a loan compared to male entrepreneurs. This is our preferred estimate for the rest of our empirical analysis.

Moreover, in column 4 we use a linear probability model, estimated via OLS. This model is preferred in some specifications, especially when using several fixed effects. The results are almost identical to those of the first specification.

An additional robustness test considers the effect of corporate governance characteristics. We first add fixed effects based on the head of the bank's credit department. For borderline credit decisions (rejections or approvals), which are the most important for identification purposes, loan officers consult the bank's credit department to make their final decision. Thus, a change in the head of the credit department could signal a change in the role of gender. Second, we add fixed effects reflecting the addition of a female director to the board. Adding a female director might also signal changes in the attitude toward female entrepreneurs. We add these fixed effects into our specifications in Table 4 and report the results in specification 1 of Appendix Table A2. The results are consistent with our baseline, suggesting that the inclusion of these fixed effects do not affect our estimates on gender.

In a nutshell, our results show that male entrepreneurs display a higher probability of applying for credit, implying that male entrepreneurs are either more willing to take credit risks or the bank is more willing to supply credit to them. However, the economic significance of this finding is not particularly high, showing that the implied probability is between 0.8 percent and 1.2 percent in our preferred specifications. We next aim to investigate whether the effect is supply (bank) driven.

B. Probability of Loan Origination and Reapplication

We examine potential differences in the probability of loan origination for female- and male-owned firms using the following model:

(2) $\text{GRANTED}_{it} = a_0 + a_1\text{GENDER}_i + a_2 x'_{i(f)t} + u_{it}.$

In equation (2), GRANTED is a binary variable equal to 1 if the loan is approved by the bank (i.e., the credit score is positive) and 0 if the loan application is rejected (i.e., the credit score is negative). The rest of the control variables are equivalent to those in equation 1, with the obvious exclusion of CREDIT_SCORE that perfectly identifies GRANTED.

We estimate equation (2) using the 95,340 observations, where individuals in our sample apply for a loan (APPLY = 1), and for which the bank makes a credit decision. Our identification approach follows the previous section. However, given that the credit score perfectly defines the bank's credit decision and cannot be included, any omitted-variable bias seems more important in equation (2). Thus, we place weight on the results from the IV model.

Any significant results in this section would reveal that the bank favors loan applicants based on gender. Identifying a positive (negative) and significant coefficient on GENDER implies that the bank favors male (female) applicants. Taken together with the findings in the previous section and understanding that outright gender discrimination is unlikely for a large European bank, a negative coefficient on GENDER is, if anything, the more likely outcome. This would imply that the bank observes the higher probability of loan applications by male entrepreneurs and is more cautious to originate loans to them given credit-risk considerations. In other words, the bank would "price" the higher credit risk of male applicants as soft information in the credit score and "discriminate" against male applicants for that reason.

18

The first specification of Table 4 reports results from the simple probit model. The coefficient on GENDER shows that male applicants are 0.5 percent less likely to be granted a loan compared to female applicants. Even though this effect is statistically significant at the 10-percent level, its economic significance is moderate if not small. In column 2, we use the IV probit model. Again, in the first stage, the coefficient on AVERAGE_FEMALE_ENTREPRENEURS strongly correlates with GENDER. In the second stage, we observe that the coefficient on GENDER remains practically unchanged compared to specification 1, but the standard error rises to levels that renders the estimate statistically insignificant at conventional levels.

[Please insert Table 4 about here]

In principle, we should draw similar inferences from models using credit score as the outcome variable in equation (2). This implies that we revert to the full sample of 357,055 observations. The importance of using this sample is that we can also control for the probability that an individual actually applies for a loan (APPLICATION_PROBABILITY), which is of obvious importance based on our findings in Table 3 that females are less likely to apply. In other words, the inability to include APPLICATION_PROBABILITY might be the reason for the marginally significant coefficient in column 1 of Table 4. APPLICATION_PROBABILITY equals the predicted values of specification 3 in Table 3. We report the results in the last two specifications of Table 4 (OLS and two-stage least squares (2SLS), respectively). The results from both the OLS and the 2SLS specifications show that gender does not explain the credit score. Thus, APPLICATION_PROBABILITY is indeed an omitted variable in the estimation of the probability of loan origination and accounting for it implies that GENDER's effect is insignificant.

As an additional test, we also examine differences in the loan amount and spread. The results in Table 5, especially those from 2SLS, suggest that the effect of GENDER is statistically insignificant. Furthermore, using fixed effects reflecting changes in the head of credit department or the addition of female directors to the bank's board does not affect the results in this section (see specifications 2 to 4 in Appendix Table A2).

Our results are in line with recent evidence by Dobbie et al. (2018), who study consumer lending using administrative data from a high-cost lender in the United Kingdom and with Delis and Papadopoulos (2018), who study mortgage loans. Our findings are contrary to the majority of studies on small-business lending (Muravyev et al., 2009; Bellucci et al., 2010; Ongena and Popov, 2016), who typically use data from several countries. Thus, we find that it is important to have detailed loan-level data with important individual characteristics to examine gender discrimination in small business lending. We conclude that the effect of gender is statistically insignificant either in the probability of loan origination or the main loan terms (loan amount and spreads).

[Please insert Table 5 about here]

Consistent with our results in Table 3, we next look into gender differences concerning the probability that rejected applicants reapply for a loan within a specific time period (one or two years).[14] We expect that within the sample of rejected applicants, and given that male applicants have a higher probability of applying for a loan in the general sample, the effect of GENDER would be more potent. That is, rejected male applicants will seek to reapply for a loan sooner than will rejected female applicants. To examine this premise, we construct the dummy variable REAPPLY, which takes the value 1 for rejected applicants who reapplied for a loan

---

[14] To be included in the balanced panel, all these individuals reapply for loans within a period of four years.

within one or two years after the bank's credit decision (takes the value 0 for those who did not reapply).[15]

For this exercise, we use the sample of rejected applicants (15,826 observations) and report the results in Table 6. The first two columns report probit estimations and specifications 3 and 4 report IV probit estimations.[16] We use a one-year window in columns 1 and 3, and a two-year window in columns 2 and 4. The results show a considerably higher probability that male applicants reapply within one or two years after a rejected application. Specifically, based on the IV estimates, we find that rejected male applicants are three percent more likely to apply for a loan in the one-year window after the original rejection. This likelihood increases to 4.2 percent in the two-year window. These probabilities are higher than those identified in Table 3 and reflect the comparative readiness (cautiousness) of male (female) applicants to reapply for credit. We view this line of results as particularly important for the subsequent empirical analysis.[17]

[Please insert Table 6 about here]

C. Gender and Firm Outcomes

Noting that male entrepreneurs are somewhat more likely to apply for a loan and that the bank does not discriminate based on gender, the most interesting question becomes whether gender affects firm outcomes via the credit channel. We first look into the probability that a firm defaults in the period following the loan origination, using the following model:

---

[15] We also know that those applicants did not reapply for credit to another bank (at least at banks being actively regulated and supervised by national or European authorities).

[16] An alternative would be to estimate duration models (e.g., Cox hazard models). We do not favor this approach here because, by construction of our panel to observe important applicant characteristics, individuals reapply for loans within four years. Thus, we document gender differences in the readiness to apply for credit within the first two years post-rejection.

[17] Again, our results are robust to the inclusion of corporate governance fixed effects (specification 5 of Table A2).

(3) $\text{DEFAULT}_{i,t+n} = a_0 + a_1 \text{GENDER}_i + a_2 x'_{i(f)t} + u_{it},$

where the outcome variable is a dummy equal to 1 when a firm defaults within n years after the current year (and 0 otherwise).

We begin by using the full sample of 357,055 observations and again explicitly control for the APPLICATION_PROBABILITY and the CREDIT_SCORE. Of course, these variables do not exhaustively control for omitted-variable bias in this case (as in equation (1)) because the bank cannot perfectly foresee a nonperforming loan. Thus, simple probit or OLS models might be biased; nevertheless, we report them in columns 1 and 3 of Table 7 for comparability purposes. The probit (OLS) results show that a firm default is 1.3 percent (1.4 percent) more probable among male applicants (results statistically significant at the one-percent level). Moving to the more reliable IV probit results (column 2 of Table 7), we observe an increase in our estimated probability to two percent. Notably, despite the use of instrumentation, the standard error is not significantly higher, making the IV probit model our preferred one. The equivalent 2SLS estimate (in column 4) is even higher, but the standard error also significantly increases, displaying econometric inefficiency.[18]

[Please insert Table 7 about here]

To pinpoint gender differences in the probability of firm default, we next use only those observations for which there is a loan application, as the bank's credit decision might be an important factor in firm default. We report the results in Table 8. We use all control variables of the previous estimations, but do not report the estimates from this point onward (results are available on request). We first show marginal effects from the four combinations of GRANTED and GENDER. The first two specifications include the interaction term GRANTED $\times$ GENDER

---

[18] In specification 6 of Table A2, we show that result are robust to the inclusion of relevant corporate governance fixed effects.

(along with the main terms), while specifications 3 and 4 also include the triple term GRANTED × GENDER × CREDIT_SCORE (along with the relevant main and double terms). The latter approach offers additional information on possible heterogeneity on our results due to the credit score.

[Please insert Table 8 about here]

The results in column 1 indeed show that when the loan is not granted (i.e., when GRANTED = 0), the probability of default is higher for both female and male applicants. The results in column 2 show that this effect is lower when we control for the credit score and also within a triple interaction term. This is intuitive, given the information in the credit score for the applicant's quality. The difference in the effect of GENDER when GRANTED = 0 is generally statistically insignificant, with coefficient estimates essentially being equal when we control for the triple term in columns 2 and 4 (probit and OLS models, respectively). Interestingly, this is not the case when GRANTED = 1. According to column 2, we observe a 1.7 percent difference between male applicants (whose probability of firm default is 2.3 percent) and female applicants (whose probability of firm default is only 0.6 percent).

We essentially confirm these effects by estimating an IV probit model for the average effect of gender on the probability of firm default (column 5 of Table 8). The first stage of the model again shows that the effect of AVERAGE_FEMALE_ENTREPRENEURS is highly statistically and economically significant. The second-stage results show that male applicants are 2.1 percent more likely to experience their firms' default, an effect statistically significant at the five-percent level. This effect is also quite strong economically, considering that apart from income, wealth, and education, no other control variable reflecting individual characteristics has a higher marginal effect.

The second important question concerns measures of firm profitability and leverage three years after the bank's credit decision. This is the most important outcome we are considering in terms of the role of the bank's credit decision in the gender-firm performance nexus. To this end, in the same fashion with equations (1)-(3), we estimate the following model:

(4) $\text{FORWARD ROA (LEVERAGE)}_{i,t+3} = a_0 + a_1 \text{GENDER}_i + a_2 x'_{i(f)t} + u_{it},$

where FORWARD_ROA or FORWARD_LEVERAGE is observed three years after the bank's credit decision (i.e., at $t+3$).

The OLS results are in Table 9. Similar to Table 8, we report results for the four combinations of interaction between GRANTED and GENDER. We observe that when the bank's credit decision is negative, the effect of GENDER on both FORWARD_ROA and FORWARD_LEVERAGE is statistically and economically similar. The respective effects when GRANTED = 1 is statistically different when comparing male and female entrepreneurs. Specifically, based on our preferred estimates in column 2 (probit model including the triple term with CREDIT_SCORE), the firms of male borrowers have a 1.4 points higher FORWARD_ROA compared to firms of female borrowers. This comes at the cost of higher FORWARD_LEVERAGE for the firms of male borrowers (0.209 vs. 0.195 of female borrowers).

[Please insert Table 9 about here]

The results in Tables 7 to 9 might be due to the omitted-variable bias more than to the respective results in previous tables (which also include IV methods), because the credit score cannot contain adequate information about future firm outcomes, i.e., no bank has perfect foresight on future firm outcomes. Furthermore, the IV used in the previous estimations might not be as suitable in the context of forward ROA models because future profitability is a function

of many current and future developments that might be correlated with regional dynamics in the three years from application to forward ROA materialization.

A solution to this important identification problem comes from the dichotomy between the estimation results when GRANTED =1 vs. GRANTED = 0, which implies that the bank's credit decision and the underlying credit score (which is strictly used to reach this decision) create a sharp RDD (see also Berg, 2018). The theoretical reason is straightforward: A positive credit decision helps firms generate liquidity and investment, increase their future profitability, and lower the future probability of default. We validate this theoretical assertion with several empirical tests.

In our context, we need to examine the heterogeneous effect of granting a loan by gender. Using an RDD with interaction terms to infer heterogeneous effects is not common practice in the related literature. Therefore, we mostly rely for identification of the effect of gender on future firm outcomes to the estimation of equation (4) separately for female and male entrepreneurs.[19] The credit score is the assignment (also referred to as "the running" or "the forcing") variable (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). Also following the literature, we use a nonparametric local linear regression, which has the advantage of assigning higher weights to observations closer to the cutoff value of zero.[20] We determine the optimal bandwidth using the

---

[19] In general, the advantage of using two separate regressions is that the slopes of all the right-hand side variables are allowed to differ and this might be preferable when these variables have largely different correlations by gender. In contrast, the advantage of the model with interaction terms is using information from the full sample once. In our context, the two separate regressions have another important advantage. The "rdrobust" Stata tools by Calonico et al. (2014), Cattaneo et al. (2016), Calonico et al. (2018), Cattaneo et al. (2018), and related papers allow identifying the validity of the RDD and produce robust estimates. These imply improved inference and associated transparency. However, these tools come at the expense of some flexibility loss, especially as we cannot introduce the interaction term GRANTED × GENDER. In the technically most relevant recent study, Berg (2018) uses a local linear regression and more standard software allowing the regression function to differ on both sides of the cutoff point (see also Lee and Lemieux, 2010, p. 318). This is different from our specification, where we opt to examine heterogeneous effects by GENDER. When comparing estimates without an interaction term between standard Stata packages and the packages by Calonico et al., we find them to be essentially the same.
[20] In contrast, a parametric OLS regression places an equal weight to all observations.

approach in Calonico, Cattaneo, and Titiunik (2014), and for efficient estimation we base our inference on the local-quadratic bias-correction in Calonico, Cattaneo, and Farrell (2018) and Cattaneo, Jansson, and Ma (2018).

In Figures 4 and 5, we provide a graphical representation of the relation between the credit score and our outcome variables for the full sample of loan applicants (i.e., APPLY = 1), as well as for the separate samples of male and female applicants. The points represent local sample means of the applicant's income for a set of disjointed bins of control and treatment units spanning the full sample. We select evenly spaced bins that mimic the underlying variability of the data using spacings estimators. The continuous line represents a fourth-order polynomial fit used to approximate the conditional mean of applicants' income below and above the cutoff. All the figures show clear upward shifts in both FORWARD_ROA and FORWARD_LEVERAGE around the cutoff. This indeed shows that the treatment (GRANTED = 1) entails a sharp discontinuity in both the outcome variables for the full sample and for the separate samples of male and female applicants. In addition, the relations reflect nonlinearity, which is more pronounced for female applicants and for FORWARD_LEVERAGE. In that sense, the local linear regression helps with identification, as the family of nonparametric models is better suited to account for nonlinearity.

[Please insert Figures 4 & 5 about here]

The key assumption for the validity of the RDD is that applicants cannot *precisely* manipulate their credit score (Lee and Lemieux, 2010). Manipulation of the credit score is difficult to achieve consistently and precisely, assuming that the bank is a value-maximizing entity aiming to minimize nonperforming loans. In Figure 6, we run a manipulation test proposed by Cattaneo et al. (2018). The test uses the local quadratic estimator with cubic bias-correction

26

and a triangular kernel. In line with our theoretical priors, the formal test confirms there is no statistical evidence of manipulation of the assignment variable. In the Appendix, we also provide additional evidence for the validity of our RDD. Specifically, we show that the distribution of the credit score does not jump around the cutoff (Figure A1). A second assumption on RDD validity is that our control variables are not discontinuous at the cutoff and we indeed find this to be the case (figures available on request).

[Please insert Figure 6 about here]

Following the validity tests, we report our baseline RDD results in Table 10. We report the bias-corrected RD estimates with conventional variance estimator. The equivalent results with a robust variance estimator are almost the same. For the estimation, the specific RDD method uses a specific number of observations right and left of the cutoff (reported as effective observations in Table 10); this also implies that the approach is quite less sensitive to difference in the sample size between male and females. The estimate in column 1 suggests that a positive credit decision lowers the probability of firm default for female applicants by a substantial 18.9 percent. The equivalent estimate for male applicants in column 4 is 25.9 percent. This 7 percent difference is statistically significant at the one-percent level and suggests that male applicants rely on a loan origination to avert their firms' default much more than do female applicants. This finding is consistent with the more risk-averse behavior of female entrepreneurs, which is also uncovered in their lower probability of applying for credit, and their generally lower default rates, as shown in Tables 7 and 8.

The corresponding effects on FORWARD_ROA and FORWARD_LEVERAGE are more indicative of our overall findings. The higher leverage assumed by male applicants and the corresponding higher probability of default trade off the higher average future ROA.

27

Specifically, we find that a positive credit decision increases FORWARD_ROA of the firms by male entrepreneurs by 0.011 more than the corresponding increase of female entrepreneurs. This is a large difference given the mean average ROA of 0.068 (0.082) for female-owned firms in our sample. Phrased differently, for female-owned firms, a positive credit decision by the bank increases a firm's return on assets (ROA) three years onward by approximately 29 percent (compared to the average ROA of female-owned firms). For the male-owned firms, the equivalent increase is approximately 38 percent.

Is this profitability difference really the result of the propensity of male entrepreneurs to reapply for loans more frequently and take on higher leverage and probability of default? To answer this question, we repeat our analysis in columns 2 and 4 of Table 10, this time including APPLICATION_PROBABILITY and DEFAULT_PROBABILITY (as estimated from the prediction of specification 2, Table 7). We report the results in Table 11. We note that the effect of the bank's credit decision on the increase in FORWARD_ROA of female applicants is almost the same with that in Table 10. However, the equivalent response for male applicants is significantly lower and converges to the estimate of females.[21] This analysis essentially implies that the higher ROA observed for firms owned by male entrepreneurs is predominantly due to their propensity to apply for credit somewhat more easily and the associated increase in investment. This advantage comes, however, at the expense of a limited number of male-owned firms defaulting on their higher leverage.

[Please insert Table 11 about here]

IV. Conclusions and Directions for Future Research

---

[21] Sequentially introducing the application and the default probability in our model shows that they contribute to this decline with an almost equal weight.

We use a unique sample of loan applications to a single bank, which provides all available information on applicants, their credit score, the resultant loans, and the involved firms. The applications are for corporate loans by majority owners of small and micro enterprises. This dataset allows a deep examination of the differential role of credit, from application to origination, on the performance of female- and male-owned firms.

We find a sequence of results following the loan application. First, males apply for loans more easily; however, after controlling for the credit score, the effect is economically moderate (1.2 percent higher probability for male entrepreneurs in our preferred specification). Second, after controlling for the differential probability between genders in the loan application, we find no significant gender gap in the probability of loan origination or in the loan amount and spread.

Subsequently, we examine the role of credit and its origination (or not) in the nexus between gender and future firm outcomes (three year after the bank's credit decision). We find that male-owned firms indeed have higher leverage three years after the bank's credit decision and this excess leverage contributes to somewhat higher default probability for their firms. However, male-owned firms also have higher returns on assets three years after the bank' credit decision. We show that this performance gap is mainly shaped by the tendency of male entrepreneurs to apply for credit more easily (higher demand) compared to female entrepreneurs.

The natural extension to our analysis and findings is to analyze more deeply the reasons behind the moderately higher credit demand by male entrepreneurs. The entrepreneurship literature proposes that any differences between genders in credit risk-taking and financial motivation (deviation from pure profit maximization) emanates from the propensity of female entrepreneurs to demand work autonomy (be self-employed and decide on their labor supply) and be risk-averse, especially when they are married and have children. Given our results that

credit origination is the key element affecting the gender-related performance gap, further analyzing the precise reasons behind the higher demand for credit by male entrepreneurs implies better understanding of this gap, which is especially important given our finding that male applicants more readily reapply for loans after being rejected. As our analysis already covers considerable ground, we leave this avenue for future research.

References

Adams, R., and D. Ferreira. "Women in the Boardroom and their Impact on Governance and Performance." *Journal of Financial Economics* 94 (2009), 291-309.

Adams, W.; L. Einav; and J. Levin. "Liquidity Constraints and Imperfect Information in Subprime Lending." *American Economic Review* 99 (2009), 49-84.

Ahl, H. "Why Research on Women Entrepreneurs Needs New Directions." *Entrepreneurship Theory and Practice* 30 (2006), 595-621.

Alesina, A.F.; F. Lotti; and P.E. Mistrulli. "Do Women Pay More for Credit? Evidence from Italy." *Journal of the European Economic Association* 11 (2013), 45-66.

Alsos, G.A.; E.J. Isaksen; and E. Ljunggren. "New Venture Financing and Subsequent Business Growth in Men and Women-led Businesses." *Entrepreneurship Theory and Practice* 30 (2006), 667-686.

Asiedu, E.; I. Kalonda-Kanyama; L. Ndikumana; and A. Nti-Addae. "Access to Credit by Firms in Sub-Saharan Africa: How Relevant is Gender?" *American Economic Review* 103 (2013), 293-297.

Baker, T.; H.E. Aldrich and N. Liou. "Invisible Entrepreneurs: the Neglect of Women Business Owners by Mass Media and Scholarly Journals in the USA." *Entrepreneurship and Regional Development* 9 (1997), 221-238.

Beck, T.; P. Behr; and A. Guettler. "Gender and Banking: Are Women Better Loan Officers?" *Review of Finance* 17 (2012), 1279-1321.

Bellucci, A.; A. Borisov; and A. Zazzaro. "Does Gender Matter in Bank-Firm Relationships? Evidence From Small Business Lending." *Journal of Banking and Finance* 34 (2010), 2968-2984.

Berg, T. "Got Rejected? Real Effects of Not Getting a Loan." *Review of Financial Studies* 31 (2018), 4912-4957.

Berger, A.N.; W. Goulding; and T. Rice. "Do Small Businesses Still Prefer Community Banks?" *Journal of Banking and Finance* 44 (2014), 264-278.

Berger A.N., and K. Schaeck. "Small and Medium-sized Enterprises, Bank Relationship Strength, and the Use of Venture Capital." *Journal of Money, Credit and Banking* 43 (2011), 461-490.

Bonfim, D.; Q. Dai; and F. Franco. "The Number of Bank Relationships and Borrowing Costs: the Role of Information Asymmetries." *Journal of Empirical Finance* 46 (2018), 191-209.

Calonico, S.; M.D. Cattaneo; and R. Titiunik. "Robust Nonparametric Confidence Intervals for Regression-discontinuity Designs." *Econometrica* 82 (2014), 2295-2326.

Calonico, S.; M.D. Cattaneo; and M.H. Farrell. "On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference." *Journal of the American Statistical Association* 113 (2018), 767-779.

Carter, S.; E. Shaw; W. Lam; and F. Wilson. "Gender, Entrepreneurship, and Bank Lending: the Criteria and Processes Used by Bank Loan Officers in Assessing Applications." *Entrepreneurship Theory and Practice* 31 (2007), 427-444.

Cattaneo, M.D.; M. Jansson; and X. Ma. "Manipulation Testing Based on Density Discontinuity." *Stata Journal* 18 (2018), 234-261.

Cattaneo, M.D.; R. Titiunik; and G. Vazquez-Bare. "Inference in Regression Discontinuity Design Under Local Randomization." *Stata Journal* 16 (2016), 331-367.

Chen, J.; W. Leung; and M. Goergen. "The Impact of Board Gender Composition on Dividend Payouts." *Journal of Corporate Finance* 43 (2016), 86-105.

Coleman, S. "The Role of Human and Financial Capital in the Profitability and Growth of Women-owned Small Firms." *Journal of Small Business Management* 45 (2007), 303-319.

Conroy, T.; and S. Weiler. "Does Gender Matter for Job Creation? Business Ownership and Employment Growth." *Small Business Economics* 47 (2016), 397-419.

Dawson, C., and A. Henley. "Gender, Risk, and Venture Creation Intentions." *Journal of Small Business Management* 53 (2015), 501-515.

Dean, H.; G. Larsen; J. Ford; and M. Akram. "Female Entrepreneurship and the Metanarrative of Economic Growth: a Critical Review of Underlying Assumptions." *International Journal of Management Reviews* 21 (2019), 24-49.

Delis, M.; F. Fringuellotti; and S. Ongena. "Credit and Income Inequality." Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3392497 (2019) .

Delis, M.D., and P. Papadopoulos. " Mortgage Lending Discrimination Across the U.S.: New Methodology and New Evidence." *Journal of Financial Services Research* 56 (2019), 341-368.

Dobbie, W.; A. Liberman; D. Paravisini; and V. Pathania. "Measuring Bias in Consumer Lending." *National Bureau of Economic Research*, working paper No. 24953, (2018).

European Commission [EC]. Statistical Data on Women Entrepreneurs in Europe. European Commission, Brussels, (2014).

Fang, L.H., and S. Huang. "Gender and Connections Among Wall Street Analysts." *Review of Financial Studies* 30 (2017), 3305-3335.

Farinha, M.L.A., and J.A.C. Santos. "Switching From Single to Multiple Bank Lending Relationships: Determinants and Implications." *Journal of Financial Intermediation* 11 (2002), 124-151.

Huang, J., and D.J. Kisgen. "Gender and Corporate Finance: Are Male Executives Overconfident Relative to Female Executives?" *Journal of Financial Economics* 108 (2013), 822-839.

Imbens, G. W., and T. Lemieux. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142 (2008), 615-635.

Iyer, R., and M. Puri. "Understanding Bank Runs: the Importance of Depositor-bank Relationships and Networks." *American Economic Review* 102 (2012), 1414-1445.

Karavitis, P.; S. Kokas; and S.Tsoukas. Gender Board Diversity and the Cost of Bank Loans. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3453680 (2019).

Lee, D. S., and T. Lemieux." Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2010), 281-355.

Muravyev, A.; O. Talavera; and D. Schäfer. "Entrepreneurs' Gender and Financial Constraints: Evidence from International Data." *Journal of Comparative Economics* 37 (2009), 270-286.

Niederle, M., and L.Vesterlund. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* 122 (2007), 1067-1101.

OECD/European Union. Policy Brief on Women's Entrepreneurship. Luxembourg: Publications Office of the European Union (2017).

Ongena, S., and A. Popov. "Gender Bias and Credit Access." *Journal of Money, Credit and Banking* 48 (2016), 1691-1724.

Piacentini, M. "Women Entrepreneurs in the OECD: Key Evidence and Policy Challenges." OECD Social, Employment and Migration Working Papers, No. 147, OECD Publishing, Paris.

Qi, S.; S. Ongena; and H. Cheng. " When They Work With Women, Do Men Get All the Credit?" Swiss Finance Institute Research Paper No. 18-01 (2018).

Roper, S., and J.M. Scott. "Perceived Financial Barriers and the Start-up Decision: An Econometric Analysis of Gender Differences Using GEM Data ." *International Small Business Journal* 27 (2009), 149-171.

Sila, V.; A. Gonzalez; and J. Hagendorff. "Women on Board: Does Boardroom Gender Diversity Affect Firm Risk?" *Journal of Corporate Finance* 36 (2016), 26-53.

Schwartz-Ziv, M. "Gender and Board Activeness: The Role of a Critical Mass." *Journal of Financial and Quantitative Analysis* 52 (2017), 751-780.

Stefani, M.L., and V.P. Vacca. "Credit Access for Female Firms: Evidence from a Survey on European SMEs." Bank of Italy Occasional Paper No. 176 (2013).

Table 1. Data and Variable Definitions

Variable	Description
A. Dimension of the data	
Individuals	Loan applicants who have an exclusive relationship with the bank and are majority owners (own more than 50%) of a firm. These borrowers apply to the bank for one or more business loans during the period 2002-2016 and the loan is either originated or denied. Due to the exclusive relationship, the bank holds information on the applicants even outside the year of loan application.
Year	The years covering the period 2002-2017. Applications end in 2016 and we use one more year of firm financial ratios to examine future firm outcomes.
B. Variables	
Apply	A dummy variable equal to 1 if the individual applied for a loan in a given year and 0 otherwise.
Gender	A dummy variable equal to 1 if the applicant is a male and 0 otherwise.
Income	The euro amount of individuals' total annual income (in log).
Wealth	The euro amount of individuals' total wealth other than the assets of the firm and minus total debt (in log).
Education	An ordinal variable ranging between 0 and 5 if the individual completed the following education. 0: No secondary; 1: Secondary; 2: Post-secondary, non-tertiary; 3: Tertiary; 4: MSc, PhD or MBA.
Age	The applicant's age.
Marital status	A dummy variable equal to 1 if the applicant is married and 0 otherwise.
Dependents	The number of dependents.
Firm size	Total firm's assets (in log).
Firm leverage	The ratio of firm's total debt to total assets.
Firm ROA	The ratio of firm's after-tax profits to total assets.
Firm cash	The ratio of cash holdings to total assets.
Forward ROA	The mean FIRM_ROA in the three years after the year of the loan application.
Forward growth	The mean increase in FIRM_SIZE in the three years after the year of the loan application.
Forward leverage	The mean FIRM_LEVERAGE in the three years after the year of the loan application.
Credit score	The credit score of the applicant, as calculated by the bank. There is a 0 cutoff: positive values indicate that the loan is approved and negative values indicate that the loan is denied.
Applications	The number of applications to the same bank before the current loan application.
Granted	A dummy variable equal to 1 if the loan is originated (Credit score>0) and 0 otherwise (Credit score<0).
Default	A dummy variable equal to 1 if the firm defaults and 0 otherwise.
Loan amount	Log of the loan facility amount in thousands of euros.
Loan spread	The difference between the loan rate and the LIBOR (in basis points).
Maturity	Loan maturity in months.
Loan provisions	A dummy variable equal to 1 if the loan has performance pricing provisions and 0 otherwise.
Collateral	A dummy variable equal to 1 if the loan has collateral guarantees and 0 otherwise.
Average female entrepreneurs	The share of female entrepreneurs to total entrepreneurs by region, industry, and year, 15 years before the loan application.

Table 2. Summary Statistics

The table reports the number of observations, mean, standard deviation, minimum, and maximum for the variables use in the empirical analysis. The variables are defined in Table 1.

	Obs.	Mean	St. dev.	Min.	Max.
Panel A: Full sample					
Apply	357,056	0.267	0.442	0	1
Gender	357,056	0.802	0.379	0	1
Income	357,056	10.94	0.428	9.738	12.78
Wealth	357,055	12.07	0.615	7.481	14.29
Education	357,056	2.723	1.015	0	5
Age	357,056	43.95	15.86	21	76
Marital status	357,056	0.573	0.484	0	1
Dependents	357,056	1.851	1.472	0	7
Firm size	357,055	12.89	0.430	10.18	14.39
Leverage	357,055	0.206	0.125	0.126	0.831
ROA	357,056	0.079	0.100	-0.409	0.579
Cash	357,056	0.080	0.032	0.066	0.255
Credit score	357,056	0.637	0.603	-0.773	3.500
Applications	357,056	6.813	2.474	1	9
Granted	95,340	0.834	0.372	0	1
Default	357,056	0.020	0.094	0	1
Loan amount	95,340	3.546	2.008	0.703	10.96
Loan spread	79,470	338.3	245.8	33.45	985.7
Maturity	95,340	46.6	37.07	4	271
Loan provisions	79,470	0.396	0.456	0	1
Collateral	79,470	0.692	0.488	0	1
Av. female entrepreneurs	357,056	0.197	0.045	0.036	0.300
Application probability	357,056	0.267	0.015	0.146	0.603
Panel B: Female applicants (Gender = 0)					
Income	62,033	10.80	0.418	9.738	12.55
Wealth	62,032	11.88	0.630	7.554	14.23
Education	62,033	2.713	1.003	0	5
Firm size	62,032	12.40	0.416	10.62	14.32
Leverage	62,032	0.189	0.024	0	0.761
ROA	62,033	0.068	0.088	-0.218	0.497
Cash	62,033	0.072	0.013	0.070	0.255
Apply	62,033	0.255	0.438	0	1
Granted	16,097	0.853	0.379	0	1
Age	62,033	42.88	15.20	24	70
Dependents	62,033	1.737	1.288	0	4
Credit score	62,033	0.655	0.511	-0.629	3.500
Applications	62,033	5.128	1.485	1	8
Default	62,033	0.018	0.078	0	1
Loan amount	16,097	3.379	1.987	0.717	10.12
Loan spread	13,370	325.1	230.3	37.20	852.5
Maturity	16,097	44.8	36.85	9	264
Loan provisions	13,370	0.368	0.460	0	1
Collateral	13,370	0.661	0.455	0	1
Panel C: Male applicants (Gender = 1)					
Income	295,023	11.01	0.426	9.818	12.78
Wealth	295,023	12.13	0.612	7.481	14.29
Education	295,023	2.729	1.011	0	5

37

Firm size	295,023	13.01	0.429	10.18	14.39
Leverage	295,023	0.215	0.075	0	0.831
ROA	295,023	0.082	0.093	-0.349	0.579
Cash	295,023	0.082	0.044	0.0661	0.254
Apply	295,023	0.277	0.443	0	1
Granted	79,243	0.827	0.368	0	1
Age	295,023	44.11	15.99	21	76
Dependents	295,023	1.889	1.429	0	7
Credit score	295,023	0.649	0.603	-0.773	3.500
Applications	295,023	6.215	1.474	1	9
Default	295,023	0.021	0.097	0	1
Loan amount	79,243	3.594	1.997	0.703	10.96
Loan spread	66,100	353.9	246.4	33.45	985.7
Maturity	79,243	47.4	37.12	4	271
Loan provisions	66,100	0.418	0.446	0	1
Collateral	66,100	0.710	0.475	0	1

Table 3. Probability of Loan Application

The table reports marginal effects and standard errors clustered by individual (in parentheses) from the estimation of the probability that individuals apply for a loan during our sample period. Dependent variable is the binary variable APPLY, and all variables are defined in Table 1. Specifications 1 and 2 are estimated using a probit model, specification 3 using a two-stage IV probit model (probit in both stages), and specification 4 with OLS. AVERAGE_FEMALE_ENTREPRENEURS is the instrumental variable for GENDER in specification 3 and its effect in the first stage is given after the second-stage results. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

	1	2	3	4
Dependent variable:	Apply	Apply	Apply	Apply
Income	0.030***	0.036***	0.032***	0.030***
	(0.004)	(0.004)	(0.006)	(0.003)
Wealth	-0.001	0.002*	-0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)
Education	0.024***	0.038***	0.025***	0.022***
	(0.007)	(0.008)	(0.007)	(0.006)
Age	0.000	0.003***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Dependents	0.001	0.002	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.005)
Firm size	0.001	0.007***	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Leverage	0.285***	0.389***	0.285***	0.299***
	(0.032)	(0.041)	(0.032)	(0.034)
ROA	0.024**	0.028***	0.025**	0.024**
	(0.010)	(0.009)	(0.011)	(0.010)
Cash	-0.978***	-1.822***	-0.975***	-0.942***
	(0.358)	(0.406)	(0.360)	(0.327)
Applications	0.001***	0.002***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Credit score	0.298***		0.293***	0.260***
	(0.029)		(0.031)	(0.032)
Gender	0.008***	0.027***	0.012***	0.008***
	(0.002)		(0.004)	(0.002)
First-stage results				
Av. female entrepreneurs			0.546***	
			(0.055)	
Year fixed effects	Yes		Yes	Yes
Regional fixed effects	Yes		Yes	Yes
Observations	357,055		357,055	357,055
R-squared				0.712

Table 4. Probability of Loan Origination

The first two specifications report marginal effects and standard errors clustered by individual (in parentheses) from the estimation of the probability that a loan is originated (against the probability that the loan is denied). The last two specifications report coefficient estimates and standard errors (in parentheses) of a credit score equation. Dependent variable in the first two specifications is the binary variable GRANTED and in the last two specifications the variable CREDIT_SCORE; all variables are defined in Table 1. Specification 1 is estimated using a probit model, specification 2 using a two-stage IV probit model (probit in both stages), specification 3 with OLS and specification 4 with 2SLS. AVERAGE_FEMALE_ENTREPRENEURS is the instrumental variable for GENDER in specifications 2 and 4 and its effect in the first stage is given after the second-stage results. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

	1	2	3	4
Dependent variable:	Granted	Granted	Credit score	Credit score
Income	0.262***	0.268***	0.627***	0.632***
	(0.005)	(0.006)	(0.004)	(0.006)
Wealth	0.013***	0.013***	0.036***	0.036***
	(0.002)	(0.003)	(0.002)	(0.002)
Education	0.023**	0.025**	0.025**	0.027**
	(0.011)	(0.012)	(0.012)	(0.013)
Age	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Dependents	0.001	0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.003)	(0.002)
Firm size	-0.001	-0.001	0.001	0.002
	(0.003)	(0.003)	(0.002)	(0.002)
Leverage	-0.147***	-0.147***	-0.405***	-0.395***
	(0.041)	(0.041)	(0.039)	(0.040)
ROA	0.069***	0.069***	0.113***	0.111***
	(0.015)	(0.015)	(0.012)	(0.013)
Cash	0.760**	0.765**	1.099***	1.239***
	(0.312)	(0.315)	(0.423)	(0.449)
Applications	0.057***	0.057***	0.053***	0.053***
	(0.005)	(0.007)	(0.006)	(0.008)
Application probability			4.110***	4.126***
			(1.320)	(1.355)
Gender	-0.005*	-0.005	-0.027	-0.035
	(0.003)	(0.004)	(0.040)	(0.045)
First-stage results				
Av. female entrepreneurs		0.274***		0.182***
		(0.071)		(0.033)
Year fixed effects	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes
Observations	95,340	95,340	357,055	357,055
R-squared			0.616	

Table 5. Loan Amount and Spread

The table reports coefficient estimates and standard errors clustered by individual (in parentheses) from the estimation of loan amount (first two specifications) and loan spread (latter two specifications) equations. Results are obtained from the sample of originated loans. The dependent variable is noted on the first line of table; all variables are defined in Table 1. Specifications 1 and 3 are estimated using OLS; specifications 2 and 4 are estimated using 2SLS. The lower part of the table denotes the rest of the control variables (same as in Table 3), fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

Dependent variable:	1 Loan amount	2 Loan amount	3 Loan spread	4 Loan spread
Gender	0.064*	0.060	-1.248	-1.828
	(0.034)	(0.039)	(1.432)	(2.689)
Maturity	0.005***	0.005***	0.478***	0.489***
	(0.001)	(0.001)	(0.092)	(0.105)
Loan provisions	0.243***	0.250***	-22.310***	-22.755***
	(0.051)	(0.056)	(3.109)	(3.188)
Collateral	0.020	0.021	-4.060	-4.048
	(0.014)	(0.015)	(3.103)	(3.119)
Loan amount			-5.241***	-5.170***
			(1.125)	(1.239)
Loan spread	-0.003***	-0.003***		
	(0.000)	(0.000)		
First-stage results				
Av. female entrepreneurs		0.287***		0.278***
		(0.099)		(0.097)
Other controls	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes
Observations	79,470	79,470	79,470	79,470
R-squared	0.855	0.693	0.858	0.834

Table 6. Probability of Reapplication after Rejection

The table reports marginal effects and standard errors clustered by individual (in parentheses) from the estimation of the probability that individuals reapply for a loan after facing a rejection from the bank. Dependent variable is the binary variable REAPPLY, which takes the value 1 when rejected applicants reapply within a specific time frame (and 0 otherwise). All variables are defined in Table 1. Specifications 1 and 2 are estimated using a probit model and Reapply referring to a one-year and two-year periods, respectively; specifications 3 and 4 report the equivalent results using a two-stage IV probit model (probit in both stages). AVERAGE_FEMALE_ENTREPRENEURS is the instrumental variable for GENDER in specifications 3 and 4, and its effect in the first stage is given after the second-stage results. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

Dependent variable:	1 Reapply one year	2 Reapply two years	3 Reapply one year	4 Reapply two years
Income	0.035***	0.032***	0.034***	0.035***
	(0.008)	(0.008)	(0.009)	(0.010)
Wealth	0.002	0.002	0.004	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Education	0.017**	0.021***	0.025**	0.027***
	(0.007)	(0.008)	(0.010)	(0.010)
Age	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)
Dependents	0.002	0.002	0.002	0.002
	(0.004)	(0.004)	(0.005)	(0.005)
Firm size	0.004**	0.006***	0.005**	0.009***
	(0.002)	(0.002)	(0.002)	(0.003)
Leverage	0.311***	0.328***	0.403***	0.400***
	(0.068)	(0.076)	(0.088)	(0.093)
ROA	0.020*	0.032**	0.035**	0.038**
	(0.011)	(0.014)	(0.013)	(0.016)
Cash	-0.860***	-0.822***	-0.983***	-0.992***
	(0.229)	(0.206)	(0.237)	(0.247)
Applications	0.002***	0.002***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Credit score	0.401***	0.411***	0.393***	0.407***
	(0.049)	(0.052)	(0.047)	(0.052)
Gender	0.027**	0.035**	0.030**	0.042**
	(0.013)	(0.015)	(0.015)	(0.018)
First-stage results				
Av. female entrepreneurs			0.447***	0.447***
			(0.083)	(0.083)
Year fixed effects	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes
Observations	15,826	15,826	15,826	15,826

Table 7. Probability of Firm Default in the General Sample

The table reports marginal effects and standard errors clustered by individual (in parentheses) from the estimation of the probability that a firm defaults using our full sample. Dependent variable is the binary variable DEFAULT; all variables are defined in Table 1. Specification 1 is estimated using a probit model, specification 2 using a two-stage IV probit model (probit in both stages), specification 3 with OLS and specification 4 with 2SLS. AVERAGE_FEMALE_ENTREPRENEURS is the instrumental variable for GENDER in specifications 2 and 4 and its effect in the first stage is given after the second-stage results. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

	1	2	3	4
Dependent variable:	Default	Default	Default	Default
Income	-0.009	-0.009	-0.015**	-0.018
	(0.011)	(0.012)	(0.007)	(0.014)
Wealth	-0.015**	-0.016**	-0.025***	-0.026***
	(0.006)	(0.007)	(0.008)	(0.009)
Education	-0.009***	-0.009***	-0.012***	-0.012***
	(0.002)	(0.002)	(0.003)	(0.004)
Age	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Dependents	-0.002	-0.001	-0.000	-0.000
	(0.007)	(0.007)	(0.006)	(0.008)
Firm size	-0.003	-0.003	-0.005	-0.006
	(0.004)	(0.004)	(0.005)	(0.007)
Leverage	0.154***	0.153***	0.199***	0.202***
	(0.025)	(0.027)	(0.033)	(0.043)
ROA	-0.216***	-0.223***	-0.218***	-0.219***
	(0.023)	(0.024)	(0.022)	(0.022)
Cash	-0.350***	-0.339***	-0.453***	-0.402***
	(0.100)	(0.101)	(0.141)	(0.154)
Applications	-0.001**	-0.001**	-0.002***	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)
Application probability	-0.123*	-0.128*	-0.139*	-0.145*
	(0.069)	(0.072)	(0.075)	(0.082)
Credit score	-0.141***	-0.148***	-0.155***	-0.152***
	(0.022)	(0.031)	(0.025)	(0.037)
Gender	0.013***	0.020***	0.014***	0.025**
	(0.004)	(0.005)	(0.003)	(0.011)
First-stage results				
Av. female entrepreneurs		0.340***		0.207***
		(0.052)		(0.040)
Year fixed effects	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes
Observations	357,055	357,055	357,055	357,055
R-squared			0.513	

Table 8. Probability of Firm Default when Individuals Apply for a Loan

The table reports marginal effects for GRANTED × GENDER (all four combinations) from the estimation of the probability that a firm defaults. Results are obtained from the full sample (aka balanced panel of individuals irrespective of whether there is a loan application). Dependent variable is the binary variable DEFAULT; all variables are defined in Table 1. Specification 1 is estimated using a probit model, and includes the interaction term GRANTED × GENDER (along with the relevant main terms) and CREDIT_SCORE as a control variable. Specification 2 is estimated using a probit model and includes the interaction term GRANTED × GENDER × CREDIT_SCORE (along with the relevant main terms and double interactions). Specifications 3 and 4 replicate the first two specifications, respectively, but are estimated with OLS. Specification 5 is estimated using a two-stage IV probit model (probit in both stages), with AVERAGE_FEMALE_ENTREPRENEURS being the instrumental variable. The lower part of the table denotes the control variables (same as in Table 3), fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

	1	2	3	4	5
Dependent variable:	Default	Default	Default	Default	Default
Granted=0 x Gender=0	0.038***	0.023**	0.040***	0.026**	
	(0.004)	(0.025)	(0.004)	(0.022)	
Granted=0 x Gender=1	0.043***	0.024***	0.045***	0.025**	
	(0.002)	(0.006)	(0.002)	(0.022)	
Granted=1 x Gender=0	0.008***	0.006***	0.008***	0.008***	
	(0.001)	(0.001)	(0.001)	(0.001)	
Granted=1 x Gender=1	0.025***	0.023***	0.025***	0.025***	
	(0.001)	(0.001)	(0.001)	(0.001)	
Gender					0.021**
					(0.012)
First-stage results					
Av. female entrepreneurs					0.293***
					(0.105)
Control variables	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	95,340	95,340	95,340	95,340	95,340

Table 9. Credit Decision, Gender, and Future Firm Outcomes: Parametric Results

The table reports marginal effects for GRANTED $\times$ GENDER (all four combinations) from the estimation of the probability that a firm defaults. Results are obtained from the full sample (aka balanced panel of individuals irrespective of whether there is a loan application). The dependent variable is noted on the first line of table; all variables are defined in Table 1. Specifications 1 and 3 are estimated using OLS, and include the interaction term GRANTED $\times$ GENDER (along with the relevant main terms) and CREDIT_SCORE as a control variable. Specifications 2 and 4 are estimated using OLS and include the interaction term GRANTED $\times$ GENDER $\times$ CREDIT_SCORE (along with the relevant main terms and double interactions). The lower part of the table denotes the control variables (same as in Table 3), fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

Dependent variable:	1 Forward ROA	2 Forward ROA	3 Forward leverage	4 Forward leverage
Granted=0 x Gender=0	0.041***	0.036***	0.164***	0.160***
	(0.000)	(0.001)	(0.000)	(0.001)
Granted=0 x Gender=1	0.040***	0.037***	0.190***	0.188***
	(0.000)	(0.001)	(0.000)	(0.000)
Granted=1 x Gender=0	0.073***	0.070***	0.195***	0.195***
	(0.000)	(0.000)	(0.000)	(0.000)
Granted=1 x Gender=1	0.084***	0.084***	0.211***	0.209***
	(0.000)	(0.000)	(0.000)	(0.000)
Control variables	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes
Observations	82,119	82,119	82,119	82,119

45

Table 10. Credit Decision, Gender, and Future Firm Outcomes: Nonparametric Results

The table reports coefficients and standard errors clustered by individual (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the bias-corrected RD estimates with robust variance estimator. The *** and ** marks denote statistical significance at the 1% and 5% levels. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	1	2	3	4	5	6
	Female applicants			Male applicants		
Dependent variable:	Default	Forward ROA	Forward leverage	Default	Forward ROA	Forward leverage
Granted	-0.189***	0.020***	0.009*	-0.259***	0.031***	0.028***
	(0.043)	(0.005)	(0.005)	(0.054)	(0.006)	(0.006)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,097	16,097	16,097	79,243	79,243	79,243
Eff. obs. left of cutoff	1,740	1,820	1,677	4,318	4,529	4,116
Eff. obs. right of cutoff	1,781	1,911	1,692	4,429	4,782	4,368
BW estimate	37.80	38.14	35.98	57.32	59.48	53.91
BW bias	68.94	70.12	66.28	88.47	90.22	86.85

Table 11. Credit Decision, Gender, and *Forward ROA*: Controlling for Application and Default Probabilities

The table reports coefficients and standard errors clustered by individual (in parentheses). The dependent variable is FORWARD_ROA and, compared to Table 9, both specifications include as controls the APPLICATION_PROBABILITY and the DEFAULT_PROBABILITY. Estimation method is the local linear regression with triangular kernel. For each specification, we report the bias-corrected RD estimates with robust variance estimator. The *** and ** marks denote statistical significance at the 1% and 5% levels. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	Female applicants	Male applicants
Granted	0.018***	0.020***
	(0.005)	(0.006)
Control variables	Yes	Yes
Year fixed effects	Yes	Yes
Regional fixed effects	Yes	Yes
Observations	16,097	79,243
Eff. obs. left of cutoff	1,820	4,529
Eff. obs. right of cutoff	1,911	4,782
BW estimate	38.14	59.48
BW bias	70.12	90.22

Figure 1. Our Bank Compared to other Systemic European Banks

Figure 1a shows the regression between the ratio of market to book value of equity of our bank in 2017 over the period 1985-2018 and the average of the 32 European systemic banks (32 observations). Figures 1b and 1c show the equivalent regressions for liquidity and ROA. The coefficient estimates of all three lines are statistically significant at the 1% level and correlation coefficients are 0.43, 0.34, and 0.35, respectively.
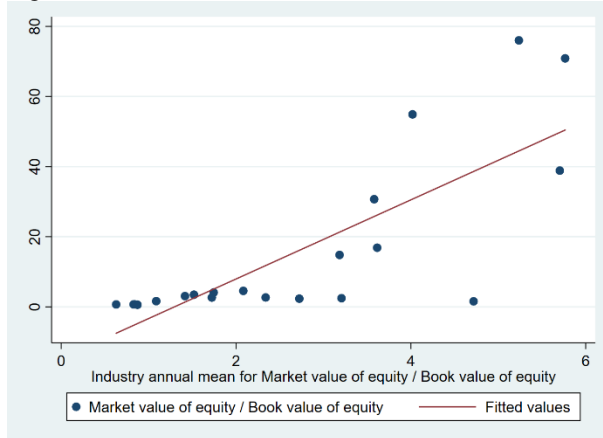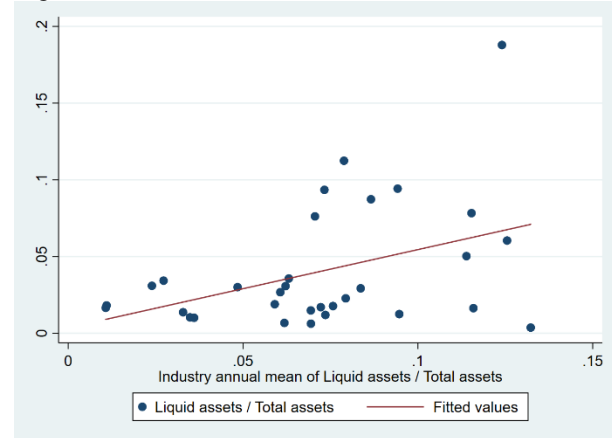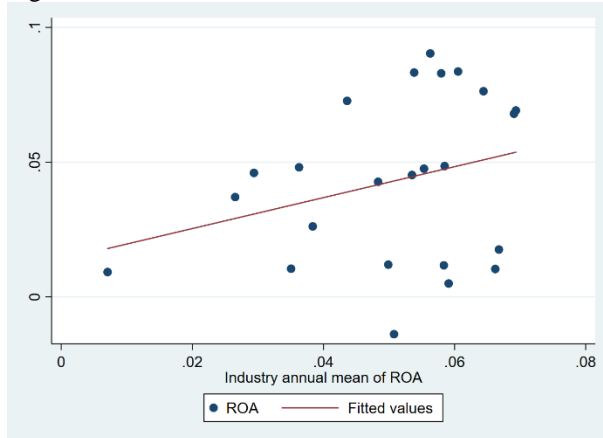
Figure 1a

Figure 1b

Figure 1c

Figure 2. Percentage of Rejected Loans to Small Firms in the Eurozone and by our Bank

The figure plots the annual mean of rejected loans to small and micro firms in the euro area (data from SAFE) and the equivalent for the loans in our sample.
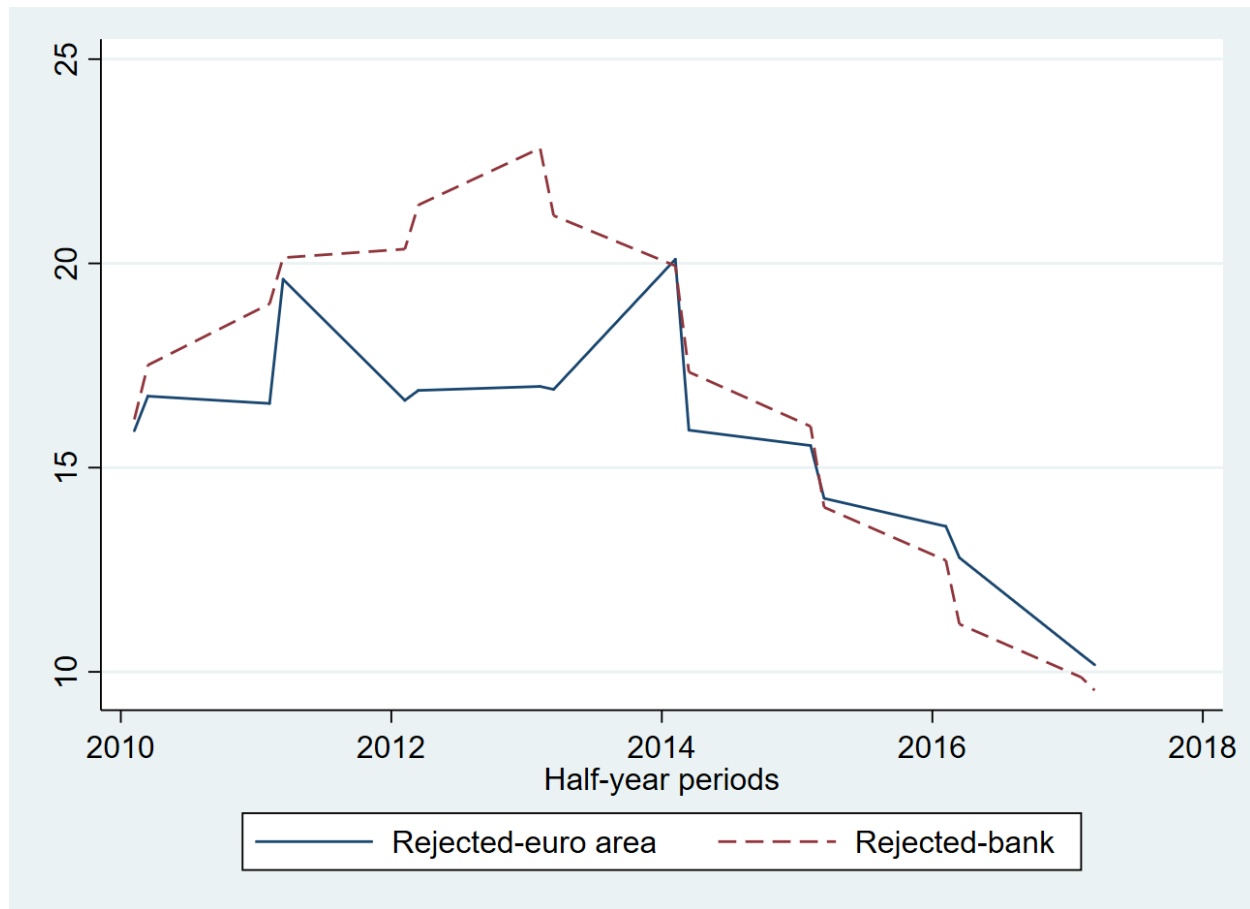

Figure 3. Leverage and ROA in Northern European Small Firms vs. our Sample

The figure plots the annual mean of leverage (Figure 3a) and ROA (Figure 3b) of small and micro firms in Austria, Belgium, Denmark, France, Germany, and the Netherlands (solid lines) and the equivalent for firms in our sample (dashed lines).


49
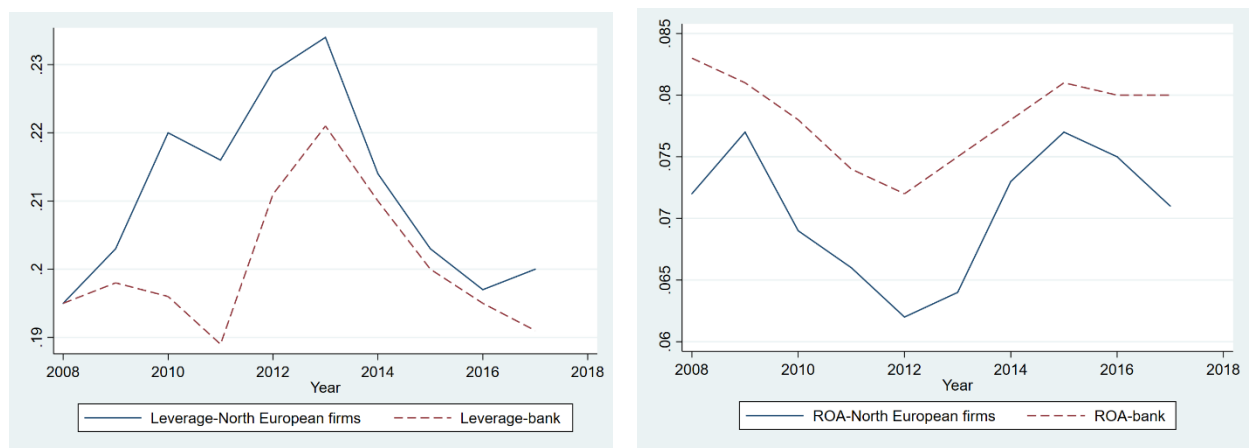
Figure 4. Responses of FORWARD_ROA at the Credit Score's Cutoff

The figures show the responses of FORWARD_ROA (y-axis) at the credit score's cutoff value (=0 on the x-axis). The first figure uses the full sample of loan applicants, the second is for male applicants, and the third for female applicants. The points represent local sample means of the applicant's income for a set of disjoint bins of control and treatment units spanning the full sample. We select evenly spaced bins that mimic the underlying variability of the data using spacings estimators. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of applicants' income below and above the cutoff.
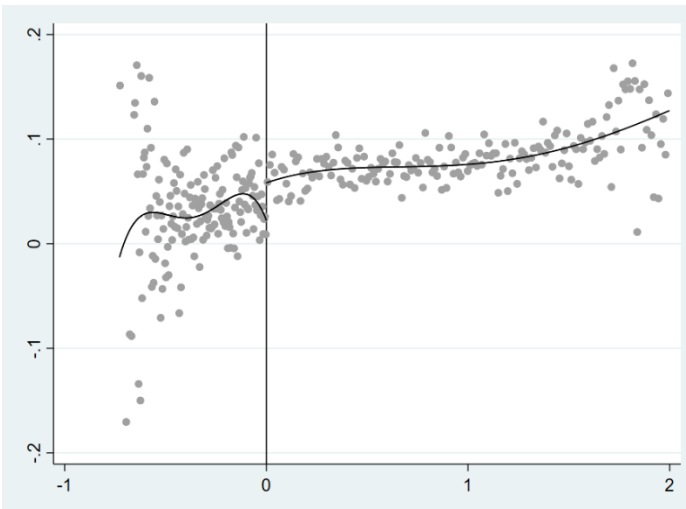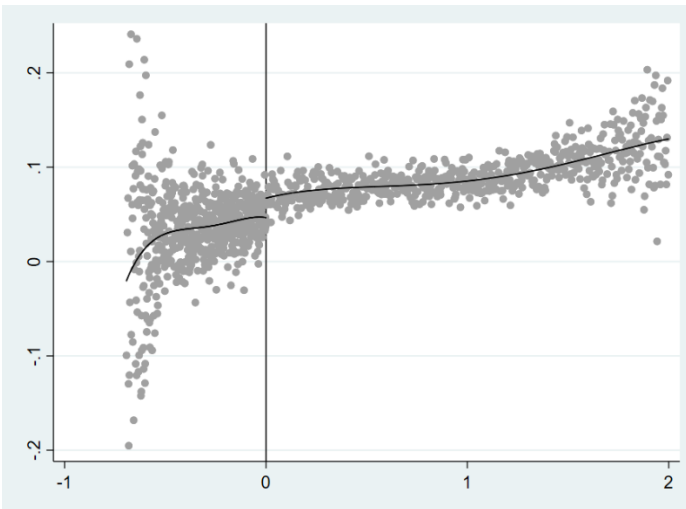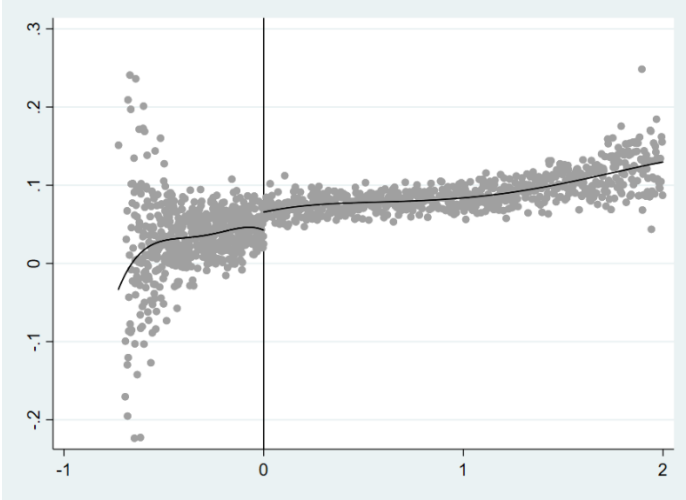
Figure 5. Responses of FORWARD_LEVERAGE at the Credit Score's Cutoff

The figures show the responses of FORWARD_LEVERAGE (y-axis) at the credit score's cutoff value (=0 on the x-axis). The first figure uses the full sample of loan applicants, the second is for male applicants, and the third for female applicants. The points represent local sample means of the applicant's income for a set of disjoint bins of control and treatment units spanning the full sample. We select evenly spaced bins that mimic the underlying variability of the data using spacings estimators. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of applicants' income below and above the cutoff.
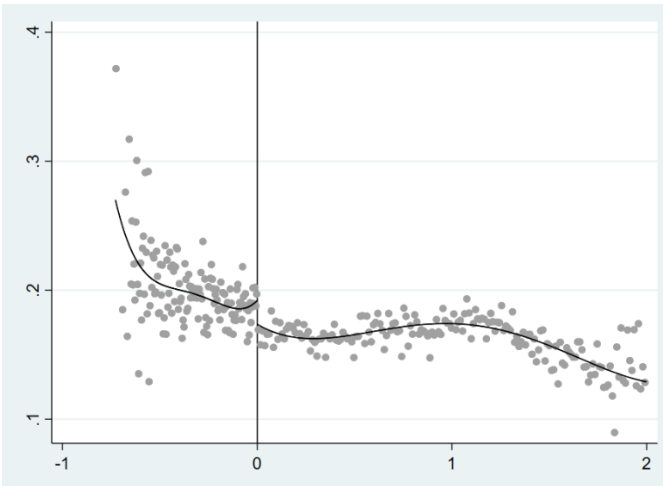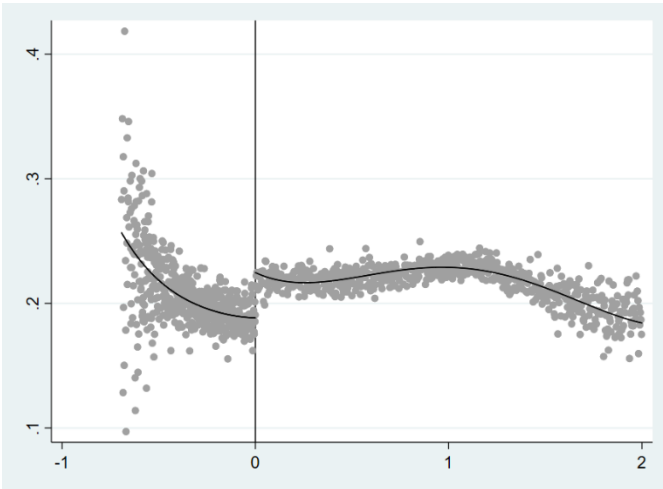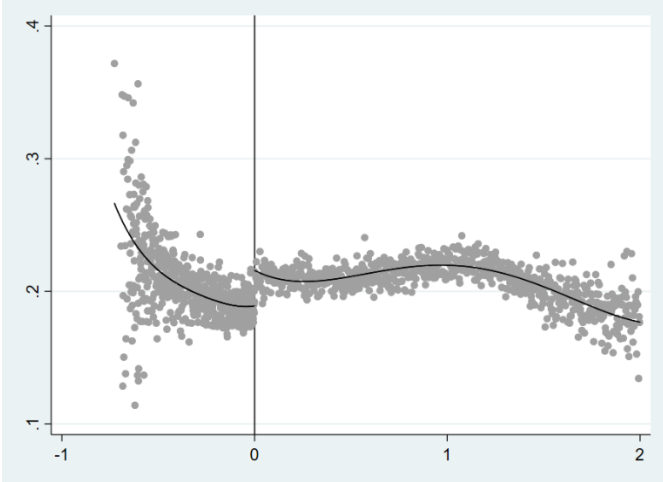
Figure 6. Manipulation Test

The figure reports results from the manipulation testing procedure using the local polynomial density estimator proposed by Cattaneo et al. (2018). To perform this test, we rely on the local quadratic estimator with cubic bias-correction and triangular kernel.
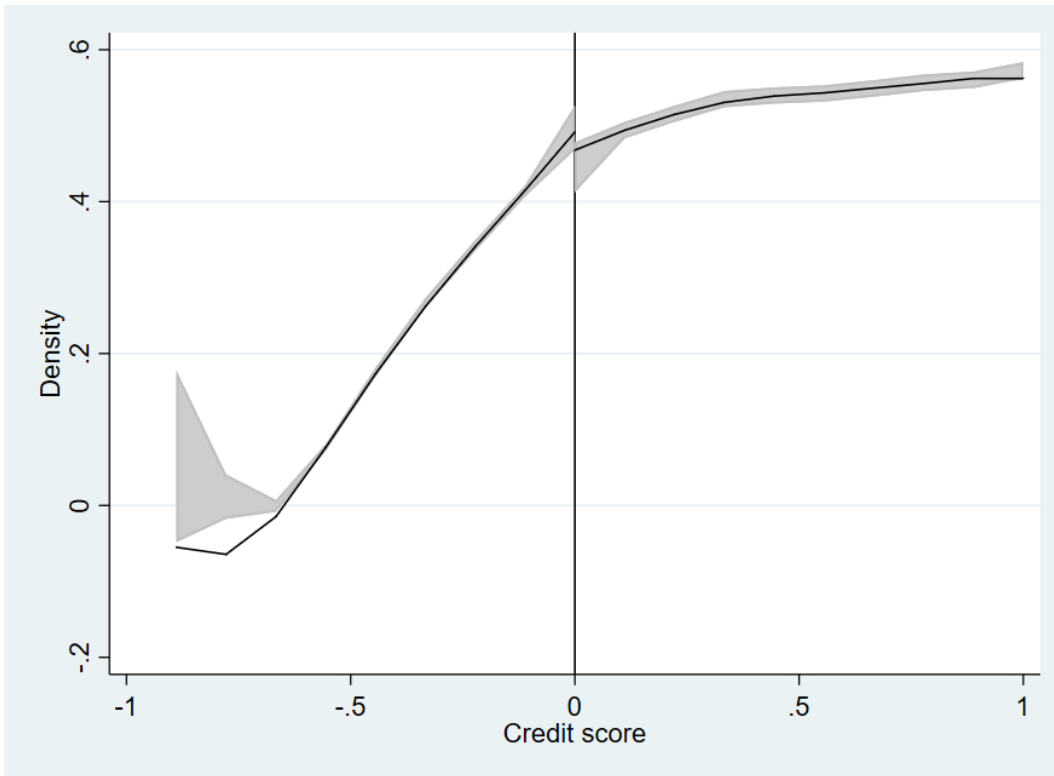
Appendix

In this appendix, we provide additional information on the validity of our identification methods. First, we provide evidence that our IV, AVERAGE_FEMALE_ENTREPRENEURS, does not directly explain the outcome variables in the IV regressions. Next, we show that the density of the credit score and other individual characteristics used as control variables in the RDD, do not significantly jump at the cutoff.

Table A1. Directly Controlling for AVERAGE_FEMALE_ENTREPRENEURS

The table reports specifications that replicate our baseline models but also include *Average female entrepreneurs* (our IV) directly in the models to show that this variable does not significantly correlate with our outcome variables. Specification 1 replicates the results of specification 1 of Table 3. Specification 2 replicates the results of specification 3 of Table 4. Specification 3 replicates the results of specification 1 of Table 5. Specification 4 replicates the results of specification 3 of Table 5. Specification 5 replicates the results of specification 1 of Table 6. Specification 6 replicates the results of specification 5 of Table 8. All variables are defined in Table 1. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

Dependent variable:	1 Apply	2 Credit score	3 Loan amount	4 Loan spread	5 Reapply one year	6 Default
Average female entrepreneurs	0.041	-0.024	-0.020	2.450	0.011	-0.016
	(0.105)	(0.163)	(0.144)	(4.801)	(0.118)	(0.201)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	357,055	357,055	79,470	79,470	15,826	95,340
R-squared		0.601	0.840	0.848		

Table A2. Additional Fixed Effects

Specification 1 reports marginal effects and standard errors clustered by individual (in parentheses) from the probability (probit model) that a loan is originated (against the probability that the loan is denied). Specification 2 reports coefficient estimates and standard errors (in parentheses) from an OLS regression on the credit score equation. Specifications 3 and 4 report coefficient estimates and standard errors clustered by individual (in parentheses) from the estimation of loan amount and loan spread equations using OLS. Specification 5 reports marginal effects and standard errors clustered by individual (in parentheses) from the estimation of the probability (probit model) that individuals reapply for a loan in the year after facing a rejection from the bank. Specification 6 reports marginal effects and standard errors clustered by individual (in parentheses) from the estimation of the probability (probit model) that a firm defaults using our full sample. All specifications include the control variables of the equivalent specifications in Tables 3 to 7. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). Compared to baseline results, the specifications include fixed effects for the head of the risk committee and the addition of a new female director in the bank's board. The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

Dependent variable:	1 Apply	2 Credit score	3 Loan amount	4 Loan spread	5 Reapply one year	6 Default
Gender	0.007***	-0.032	0.060	-1.327	0.029**	0.013***
	(0.002)	(0.041)	(0.037)	(1.740)	(0.012)	(0.004)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Head of risk dep. fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Female director fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	357,055	357,055	357,055	357,055	15,826	357,055
R-squared		0.652	0.860	0.865		

Figure A1. The Credit Score around the Cutoff