PRÊMIO ANPEC-FEBRABAN
DE ECONOMIA BANCÁRIA

14ª EDIÇÃO | **2022**

Categoria A

Artigos Acadêmicos

1º lugar

São Paulo

2022

Gender gap and local economic diversity in microfinance: evidence from the microcredit program in Brazil

Abstract: The objective of the paper is to measure the gender gap in the microcredit programs, considering the dynamics of the local economy and the labor market. We used data of the CrediBahia program in the State of Bahia, in Brazil to realize the research. The literature indicates that the default rate for female borrowers is lower than male peers, even considering the difficulty of women to access the traditional credit market. In this context, the research contributes to literature by including the effects of local and labor market features on the ability to pay by woman. We applied logistic models, Propense score matching (PSM), and Instrumental variable estimation to test the existence of gender gap among borrowers in Brazilian microcredit program. The marginal effect shows that female borrowers' default probability was 21.5% lower than that of males. The female's expected profit was 30.8% higher than the sample mean. Female borrowers was associated with an expected loss 18.6% lower than the sample mean and the repayment ratio was 1.7% more than the sample mean. The default rate also is lower for older and married women. The marginal effect shows that married female borrowers' default probability was 7.6% lower than for single female and 5.3% lower than that of males. Each additional year of the woman, the female borrowers' default probability was 0.5% lower than for men. In addition, the dynamics of the local economy contributes woman loan performance. Thus, the results suggest that lending to females is more profitable and less risky.

Keywords: Microcredit; Gender gap; Local economy.
JEL Code: G20; G21; J16.

1 Introduction

The literature indicates that the development of microcredit institutions (MFI) promotes financial inclusion, improves consumption levels, and reduces poverty for the families (SCHROEDER, 2020; BROWN; GUIN; KIRSCHENMANN, 2016; BUERA; KABOSKI; SHIN, 2021). The impacts of microcredit programs in different locations should be more carefully analyzed due to their unclear impacts in the short-term (ATTANASIO et al., 2014; AUGSBURG et al., 2012; BANERJEE et al., 2015; CRÉPON et al., 2015; BANERJEE; JACKSON, 2017; BANERJEE et al., 2019; BANERJEE; DUFLO; SHARMA, 2020). In summary, there are many gaps about the impacts of microcredit on the financial inclusion, being the gender gap one of the most important to be studied. The literature consider that females are more credit crunch than males in the traditional banking system (MURAVYEV; TALAVERA; SCHÄFER, 2009; ASIEDU et al., 2013). Shahriar (2016) indicates that external factors related to familiar factors also may influence female's ability to pay their loans. However, research indicates that women are better customers than men (PITT; KHANDKER, 1998; SHAHRIAR; UNDA; ALAM, 2020; CHEN; HUANG; YE, 2020).

We investigate gender gap of the default rate, expected profit, expected loss and repayment ratio of the loans from the microcredit program in the state of Bahia, Brazil, which is composed by 417 municipalities. The CrediBahia is the microcredit program managed by Development Agency of the State of Bahia (Desenbahia). The program was implemented in 2002 to support access to credit for small businesses in the State of Bahia, Brazil. The database is composed by microdata of the borrowers from the Desenbahia distributed among the municipalities, municipal databases on income, labor market, educational level, and population from the Brazilian Institute of Geography and Statistics (IBGE), and data on violence from the Brazilian Department of Health Informatics (DATASUS).

The local information of the municipalities also led us to investigate local context around the borrowers, affecting the ability of payment, by exploring hypotheses of agency costs such as Degryse e Ongena (2005) and Zamore, Beisland e Mersland (2019), the futures of the local labor market based on Lee e Xu (2020), and local heterogeneities of the local economies based on Gyourko e Tracy (2014) and Gerardi et al. (2018).

The refining gender gap results in the context of the local economies and the labor market can creates the possibility of a new local policies and strategies of MFI. Thus, the research contributes to the microfinance agenda on gender gap in accessing credit market. We considered the local features of the economy and the labor market in the analysis of the loan performance by gender gap, interacting gender and economic variables in each municipality. The literature usually considers the estimation of the ability to pay with individual and business characteristics. Thus, we try to create a channel to improve the microcredit program locally, by the analysis of the loan performance in the context of change in gender and local economy.

To find the results, econometric models were estimated in four steps. Fist, to examine the loan performance given the characteristics of the individual and the business we estimated Logit and OLS regression. In the second step, we included the features of municipalities into these models. In the third step, the Logit and OLS regression estimated the interaction between loan performance and gender, individual and municipal features. Finally, the fourth step estimated Propensity Score Matching (PSM), OLS regression, Probit instrumental variable (IV Probit) and Two-Stage least squares (2SLS) models to analyze possible problem of sample selection bias. The robustness of results also was verified by the Probit model to re-estimate the impact of gender on default probability and we eliminate the observations from the database that are in the upper or lower 1st percentile to eliminate outliers and re-estimate the Logit model.

The remaining of the paper is organized as follows. Section 2 presents the literature reviews on gender in microfinance. Section 3 presents the features of the microcredit program and the features of the local economy. Section 4 describes the database and summary statistics. Section 5 presents the methodology and empirical results. Section 6 concludes the paper.

2 Gender in microfinance

The development of microfinance institutions enabled the financial inclusion of families excluded from the banking sector. The MFI started operationalizing as banks in some locations, and changes in the microfinance sector have similar impacts to those changes in the traditional banking sector (BROWN; GUIN; KIRSCHENMANN, 2016; CULL; DEMIRGÜÇ-KUNT; MORDUCH, 2009; WAGNER; WINKLER, 2013). In this perspective, microcredit institutions consider the payment capacity and the indebtedness of the creditors in different economic environments where the borrowers are placed ((SCHICKS, 2014).

The improvement of administrative and financial management in the microfinance sector can avoid unnecessary risks for MFIs (STEINWAND, 2000). The literature suggests that local rather than foreign managers, as well as women, can reduce financial risks in the governance process of microfinance institutions (MERSLAND; STRØM, 2009). Nevertheless, there is no standard procedure for governance of microfinance institutions (STRØM; D'ESPALLIER; MERSLAND, 2014). The access to microcredit by low-income families without collateral makes costs sensitive to the program's management model (MORDUCH, 1999). Thus, the MFI has stimulated the research agenda about default and solvency risks of the institutions (SCHULTE; WINKLER, 2019).

The growth of microfinance institutions improves credit supply for low-income families (ASSEFA; HERMES; MEESTERS, 2013). In this context, the increase of competition has adverse effects, such as the increase of interest or credit crunch (MCINTOSH; JANVRY; SADOULET, 2005). The MFI performance could be influenced from geographical diversification, once that it reduces the loan risks as well as promotes financial inclusion for society (STEINWAND, 2000; ZAMORE; BEISLAND; MERSLAND, 2019) and the growth in the number of MFI may increase the bank risk management and the agency cost (DEGRYSE; ONGENA, 2005). Therefore, the capillarity of the MFI can reduce the risk of default but may increase the cost of risk management and agency costs.

The literature on default and credit risk modeling is extensive and growing. Altman (1968) is a seminal study about business default risk. Since that, many studies have estimated default risk models (ALTMAN, 1973; ALTMAN, 1984; FRYDMAN; ALTMAN; KAO, 1985). In addition, the macroeconomics conditions for available the credit risk also were investigated (NICKELL; PERRAUDIN; VAROTTO, 2000; ALLEN; SAUNDERS, 2004). Carling et al. (2007) includes the correlation between the firm credit risk and macroeconomics variables. They suggest that macroeconomics conditions have significant effects from default risk. Therefore, it is important to consider economic volatility over time in models of default risk.

The information about payment capability of families, such as gender and familiar structure, have been incorporated to improve the management and monitoring of microfinance institutions. The research agenda has showed that women are better to perform payments than men, but in the presence of domestic violence, this this performance decreases (SHAHRIAR, 2016; SHAHRIAR; UNDA; ALAM, 2020). In this context, the studies report that women are more risk-averse than men in financial and investment decisions (CHEN; HUANG; YE, 2020; BELLUCCI; BORISOV; ZAZZARO, 2010; HUANG; KISGEN, 2013). These results led to conclude that the participation of women in the financial market is an important object to be investigated in microfinance.

Buvinic e Berger (1990) found that the allocation of credit to women by microfinance institutions were higher than the allocation of credit to men. Muravyev, Talavera e Schäfer (2009) concluded that female-managed companies were less likely to obtain bank loans than male-managed companies. Agier e Szafarz (2013) found no gender bias in refusing loans but that Brazilian women had a “glass ceiling” effect on loan scale. Aterido, Beck e Iacovone (2013) showed for sub-Saharan Africa countries that women use less formal financial services, which could be explained by gender differences in education, income level, and formal employment. In summary, the evidence of these researches was that female borrowers might have a credit crunch.

Beck, Behr e Guettler (2013) showed that the default rate of female loan officers is significantly lower than that of male loan officers. The default gender difference is attributed to women’s higher "moral responsibility"(CROSON; GNEEZY, 2009; HERTZBERG; LIBERTI; PARAVISINI, 2010). The findings of Chen, Huang e Ye (2020) showed that the lending to female borrowers is associated with better loan performance, including a lower default probability, a higher expected profit, and a lower expected loss than for their male peers. Consequently, the increase in the credit supply to women can reduce the default rate and generates financial inclusion.

There is a growing empirical literature in microfinance that can be used to investigate the local development. Some studies around the world show that female borrowers are better than men in microcredit market (ATTANASIO et al., 2014; FRIEDSON-RIDENOUR; PIEROTTI, 2019) and that MFIs can help fight the poverty trap of economic agents without access to the traditional credit market (PITT; KHANDKER, 1998; BANERJEE et al., 2015; AUGSBURG et al., 2012; CRÉPON et al., 2015; DESAI et al., 2013; BANERJEE; JACKSON, 2017; BANERJEE et al., 2019; BANERJEE; DUFLO; SHARMA, 2020). The expansion of microfinance programs would only have a small impact on per capita income, market labor and investments profit rate (BUERA; KABOSKI; SHIN, 2021; KABOSKI; J, 2012; MEAGER, 2019; FIALA, 2018). Therefore, the microfinance agenda has been looking for analyze the effects of MFIs on economic development.

The literature of microfinance also let a venue to investigate local context around the borrowers. The spatial distribution of economic activities affects the decision of borrowers and the ability of payment, on demand side microcredit marked. On the other hand, the distance from microfinance branches can increase agency costs in the supply side (DEGRYSE; ONGENA, 2005; ZAMORE; BEISLAND; MERSLAND, 2019). Lee e Xu (2020) also showed that regional heterogeneity can influence the characteristics of the labor market of municipalities. High unemployment rates can affect both the decision to borrow and the credit default rate (GYOURKO; TRACY, 2014; GERARDI et al., 2018). Thus, the local economy and the labor market features may lead to heterogeneity in the income levels and the ability to pay of borrowers. We consider that this element can used to find more refined results regarding the gap gender, once that the dynamic of the local economy, can affect the decision of women to access the microcredit to generate income by small business. The composition and details of the dataset also encourages to contribute to gender gap literature considering the possibility of a new context of results for the local policies and strategies of MFI.

3 The microcredit program and the local economy

The CrediBahia is a microcredit program in the State of Bahia, in Brazil, which started to work in 2002. It is an initiative of the State government by the joint management of Secretariat of Labor, Employment, Income and Sports (SETRE), the State Development Agency of Bahia SA (Desenbahia), and the Brazilian Micro and Small Business Support Service (Sebrae) and Municipal governments. The program was designed to provide microcredit to micro and small business in all municipalities in the State of Bahia, under the support of Sebrae to reach micro and small business, and the coordination of SETRE and Desenbahia, considering local priorities and potentialities (COSTA; LIMA; SOUZA, 2016).

The beneficiaries are owners of micro or small businesses that demonstrates capacity to develop economic activities and pay the loans. The business must have been in operation for at least 6 months and the entrepreneur must have lived in the municipality for 1 year at least. There also is no gender discrimination and no interest rate differentials for the participates of the program. The classification criteria for credit operations and the rules for establishing credit provisions follow the resolutions of the Central Bank of Brazil.[1]

The credit flow of CrediBahia follows four stages. In the first step, the loan officers visit the business to conducts a survey locally. The baseline survey includes questions on assets, investments, production in business, as well as household income, age, and marital status of the entrepreneur. This survey allows to proceed the evaluation of the customer's trustworthiness, suitability, and entrepreneurship capacity. Second, the loan officers proceed a feasibility study, identifying the necessity of technical support for the business, and report their opinion. Third, the loan officers present the financing requests to the municipal Microcredit Committee,[2] joint with the socioeconomic information survey and his technical opinion. The approved loans by the Microcredit Committee are sent to MFI (Desenbahia) to carry the final approval. Finally, in the fourth stage the Desenbahia analyzes the request, the proposed guarantees, the amounts already disbursed under the program, the client's payment history. After analyzing all these factors, it may approve or reject the requested credit.

Table 1 shows the annual evolution of the CrediBahia program. In 2002, only five cities had CrediBahia' offices with an average ticket per contract of around R$ 971.60. The maximum value of the program financing and the loan term increased during the program. Over the years, the program became more robust and already was available in 163 municipalities in 2014, amounting about R$ 22.9 million, and the average credit ticket around R$ 978.92 per contract. In the end of 2019, these

[1] Resolution Nº 2.682/99.

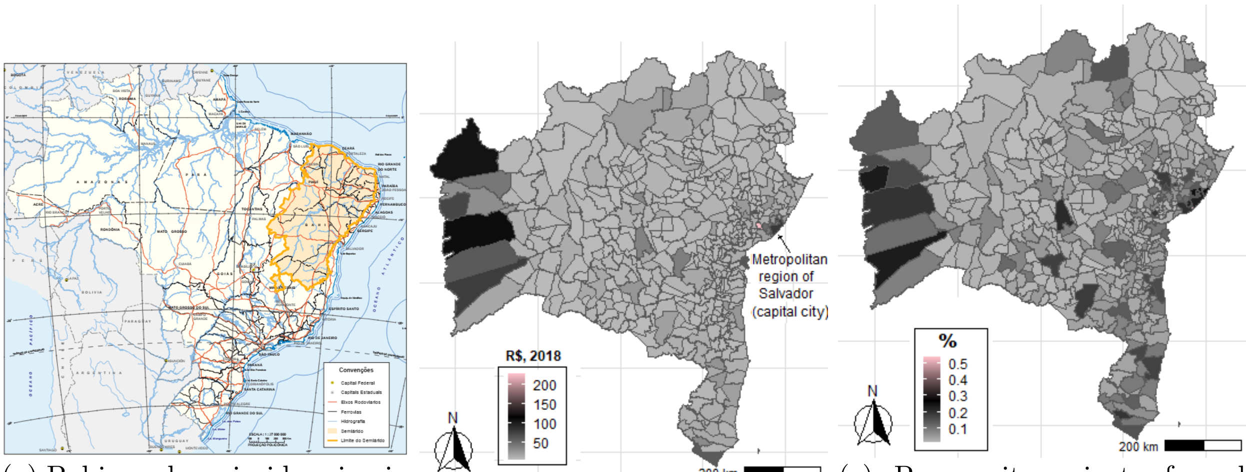
[2] The Credit Committee was formed by at least three titular members and their respective substitutes, from agencies and entities defined by consensus by SETRE, Desenbahia, and the Municipality. The committee analyzes and approves the liberation or not of the credits requested, based on studies of economic viability and technical opinions/letters.

values were R$ 11.6 million, with 10,630 credit contracts and the average credit ticket to R$ 1,089. The loan interest rate was considerable flat in the period. It was about 1.8% per month from 2002 to 2012, decreased to 0.41% from January 2014 to April 2015, and increased to 2% since the end of 2017.

About the local economy, the state of Bahia is located in the Northeast region of Brazil. The population was about 14.9 million inhabitants in 2018, with 72.0% of this population in the urban areas and 28.0% in the rural areas, distributed in 417 municipalities. The city of Salvador, with 2.9 million inhabitants, is the capital city of the State. As can be seen in figure 01, the state borders with 8 states of four macroregions in Brazil: Sergipe, Alagoas, Pernambuco and Piauí (Northeast Region), Tocantins (North Region), Goiás (Midwest Region), Minas Gerais and Espírito Santo (Southeast Region). It is washed to the east, by the Atlantic Ocean, in a coastal strip of 922 km. It has the fifth largest territorial extensions in Brazil, covering 564,733.177 km². The socioeconomic heterogeneity in the state is mainly reinforced by the fact that 2/3 of its area is on the Brazilian semi-arid region. The state GDP about R$ 303 billion in 2019 represents 4.1% of national GDP.

According to figure 1, chart (a), a large portion of semiarid region of Brazil is on area of the state of Bahia. For this reason, 61.9% of the municipalities of the state belong to the semiarid region. The regional distribution of income is considerably heterogeneous in the state, with the Metropolitan Region of Salvador concentrating 41.8% of GDP and few other large municipalities concentrating large amounts of income. As a consequence, the chart (b) shows few municipalities with high per capita GDP and a large portion of municipalities with low per capita GDP. The service sector (including the public sector) represented 69.1% of the value added, followed by industry, with 23.7% and agriculture, with 7.2% (SANTOS et al., 2019). Thus, despite to be location in the poorest region of Brazil and concentrated in the Metropolitan Region of Salvador, this state economy is considerably industrialized, with indicators from industrial sectors close to the more developed regions of Brazil. Finally, regarding the labor market, the informality rate in the labor market of 66.0% is one of the highest rates among Brazilian states (SANTOS; RIBEIRO; CERQUEIRA, 2020). In the chart (c) shows there, few private formal employment opportunities in the state and the most of the private formal employment are concentrated in few municipalities.

Figure 1: Location of the State of Bahia, per capital GDP, and private formal jobs, 2018.



(a) Bahia and semiarid region in Brazil

(b) Per capita GDP, R$ 2018

(c) Per capita private formal jobs, 2018

Source: Brazilian Institute of Geography and Statistics (IBGE); and Annual Social Information Report Ministry of Labor and Employment (RAIS).

The regional features of the local economy and its labor market are important elements to be considered in the decisions to access the credit and the ability to pay, and so, can be used by the MFI to design economic mechanisms or incentives. There are many poor, richer and industrialized municipalities with different development patterns. Therefore, we investigate gender gap of the default rate and expected profit of the loans at CrediBahia from individual, business, and municipalities variables.

4 Data and econometric modeling

In this section, we describe our main data sources and the construction of our measures of loan performance.

4.1 Credit bureau data and key variables

Our sample includes 220,183 microcredit loans, of which 30,512 were not paid by the end of 2019. We track the payment performances of all loan listings. In addition, the variables are obtained from the Desenbahia,IBGE, and DATASUS database. To analyze the gender gap in CrediBahia, we have collected three categories of database. The first category of loan characteristics is formed by the interest rate, and the amount and number of loans from each borrower. The second category is the borrowers' characteristics and includes age, marital status, income, monthly operating cash flow, and enterprise performance indexes: current ratio, working capital, and return on investment. These two categories were abtained from the Desenbahia database.

The third category includes the municipal variables from IBGE about labor market conditions, educational level, and GDP per year of each city where credit was released from IBGE. In addition, we use the violence rate from DATASUS as proxy variable to represent the institutional level of the municipality (DEMETRIADES; LAW, 2006; DEMETRIADES; FIELDING, 2012). The definition of each variable is summarized in Table 2.

4.1.1 Expected profit

Following the variable applied by Chen, Huang e Ye (2020), we defined the expected profit variable to test the impact of gender on loan performance. This variable aims to measure the expected profit for each loan. It assumes that for each loan, if the borrower repays the loan, the MFI receives $(1+r)$, where r is the interest rate. This means that the MFI makes a net profit of if the borrower repays the loan and loses the full amount if the borrower does not repay the loan. Therefore, the expected profit is $E(\pi) = (1-\sigma)r - \sigma$, where σ is the default probability (DP). To obtain the DP, we estimate the following equation using the Probit model:

$$Pr(DEFAULT_i) = \alpha_0 + \alpha_1 Female_i + \alpha_2 X_i + u_t + E_i \tag{1}$$

In this equation, the binary dependent variable, equals 1 if the borrower defaults of loan i. $Female_i$, the main variable of interest, also is binary and represents the gender of a borrower, which equals 1 if the borrower is female. X_i is a vector of control variables, including characteristics of the loans, the borrowers, and the municipalities.

The potential macroeconomic factors are controled by using a dummy year (u_t) (CARLING et al., 2007; CHEN; HUANG; YE, 2020). E_i refers to the error term. The coefficients estimated from Equation 1 are then used to predict the default probability for each loan listing. With the default probability and the interest rate, we can measure the expected profit for each loan. The expected profit is estimated using the Ordinary Least Squares (OLS) regression:

$$Exprofit_i = \alpha_0 + \alpha_1 Female_i + \alpha_2 X_i + u_t + E_i \tag{2}$$

The dependent variable indicates expected profit for each loan i. $Female_i$ is the binary variable of interest (gender of a borrower) and is equals 1 if the borrower is female. X_i is a vector of control variables, including characteristics of the loan, the borrower, and the municipality. u_t to capture year effect. E_i to the error term.

4.1.2 Expected loss

The literature on credit risk management defines the expected loss (EXPLOSS) of a loan listing as the product of loss given default (LGD) and DP, i.e., $EXPLOSS = LGD \times DP$ (BESSIS, 2011).

Chen, Huang e Ye (2020) define LGD as the fraction of the principal amount remaining if the borrower defaults at time t. We assume that all loan listings are fully amortized. The borrower pays off the debt with a fixed monthly repayment schedule in equal installments so that the loan will be fully paid off at maturity. According to Hayre e Mohebbi (1992), LGD can be computed as follows:

$$LGD = 1 - \frac{((1+r^m)^t - 1)}{((1+r^m)^n - 1)} \tag{3}$$

The variable r^m is the monthly rate (i.e., the annualized rate divided by 12) and the term n is quoted in months. For the loan listings fully repaid at maturity, t = n, and hence LGD = 0. After computing LGD, we can get the repayment ratio (RR) for problematic loans as RR = 1 – LGD.

The EXPLOSS and RR variables includes 163,150 microcredit loans. The reduction of the sample is associated with the missing information for listings of loans fully repaid at maturity for some old contracts in the microcredit program.

4.2 Summary statistics

Table 3 presents summary statistics and the distribution of loan status.The database values were deflated using IBGE's implicit deflator (Monetary values of 2002 BRL). As can be seen, 13.9% of the loans defaulted. The desegregation by sex show that 12.7% of the loans for females defaulted and 16% for males. The females account for 65% of all borrowers, similar features to literature (ATTANASIO et al., 2014; AUGSBURG et al., 2012; BANERJEE et al., 2015; CRÉPON et al., 2015; BANERJEE; JACKSON, 2017; BANERJEE et al., 2019; BANERJEE; DUFLO; SHARMA, 2020) but it is not for finance studies (BELLUCCI; BORISOV; ZAZZARO, 2010; AGIER; SZAFARZ, 2013; STRØM; D'ESPALLIER; MERSLAND, 2014; SHAHRIAR; UNDA; ALAM, 2020; CHEN; HUANG; YE, 2020).

The average amount of microcredit released was R$ 936 at an annual interest cost of 19.2%. There is evidence that borrowers are supported by the MFI because they have difficulty accessing the conventional credit market for micro and small businesses. The average number of loans requested per borrower at CrediBahia was 2.8 times. The loans are short term at CrediBahia. The loans are usually financed with an average term of 9 months.

Borrowers have an average age of 39 years, 27.9% are married, and have an income around R$ 791.18. In addition, the monthly operating cash flow is R$ 598.54 and the enterprise performance indexes show that the businesses financed by CrediBahia have an average current ratio of R$ 7.182, an average working capital of R$ 1,930, and an average return on investment of R$ 41.31. Thus, the most of the borrowers' income comes from their businesses and the entrepreneurs have limited capacity to improve their activities.

Local heterogeneity among the municipalities that joined the microcredit program also are important. Table 3 shows that there is a high standard deviation of the GDP, and population of the municipalities of R$ 2,153 thousand, and 309,686 habitants, respectively. In the labor market, the average of private formal jobs per capita is 7.5% and the maximum 70.5%., after the exclusion of public sector formal jobs. The educational level cannot be measured per borrower because there is no information in the Desenbahia database. Therefore, we have used a proxy of the ratio of employees with higher education level in the formal labor market. We observed that the average number of formal employees with higher education in the municipalities is 16.3%.

Table 4 compares the descriptive statistics between male and female loans and presents the mean difference test results. The default rate of female borrowers is 12.7%, 3.3% lower than that of male borrowers. In addition, other indicators also show significant differences between male and female borrowers. For example, male business has better working capital and business flow than female borrowers. However, women are slightly older and single, and live in more populous cities and highest GDP. Women have shorter financing terms, earn a higher income, and more expected profit than their male peers. In terms of employment, loans to female micro entrepreneurs are in municipalities with better levels of education and formal employment rates.

5 Empirical strategy and results

Econometric models were used to estimate the default rate, the expected profit, the expected loss and the repayment ratio of lenders, considering the characteristics of the individual, the business, and the municipality. The first econometric model followed the Logit and OLS regression. The second model was the Logit and OLS estimation varying across different cohorts. The third model we estimated a Propensity Score Matching (PSM) to analyze possible problem of sample selection bias. In the fourth model we estimated gender gap with instrumental variables for the OLS, IV Probit and 2SLS models.

We applied two approaches to robustness checks. In the first, the Probit model was applied to re-estimate the impact of gender on default probability. In the second, we eliminate from the database whose AGE, INCOME, FLOW, C_RATIO, W_CAPITAL, and R_INVEST are in the upper or lower 1st percentile to eliminate outliers and re-estimate the Logit model.

5.1 Gender and loan performance

The literature indicates that female borrowers are associated with a lower default rate and a higher expected profit than male peers. Table 5 examines the effect of gender on the microcredit default rate using Equation (1). Columns (1) and (2) of Table 5 presents the models by the Logit model. The results in Column (1) show that female borrowers have a lower default rate than men. Adding municipal variables, Column (2) confirms the result. These results are also consistent with the literature (CHEN; HUANG; YE, 2020; SHAHRIAR; UNDA; ALAM, 2020). We also analyzed the marginal effect estimated by Logit method. The marginal effect shows that female borrowers' default probability was 21.5% (0.02988/0.139)[3] lower than that of males.

The highest default rates were associated with loans that had a larger loan amount of borrowing, a higher interest rate and longer term of maturity. In addition, the lower default rate was associated with loans from older, married, and higher income borrowers. Moreover, the larger the number of

[3] 0.02988 is the marginal effect of Female in Column (2) of Table 5. We do not show the marginal effects of all variables for the sake of brevity. They are available upon request.

loans from the borrower and the larger the volume of contracts per capita in the municipality, the lower the default rate.

Columns (3) and (4) of Table 5 report OLS estimation of Equation (2) for expected profit. The results found in Column (3) showed that the expected profit was 0.028% higher for female loans than for male loans or 30.8% higher than the sample mean. Columns (4) support the results after adding municipal variables. The highest expected profit was associated in municipalities with the highest formal employment in the private sector per capita. Therefore, there was evidence that female microcredit borrowers have lower default rates and higher expected profits than male microcredit borrowers. Columns (5) and (6) of Table 5 present the OLS estimation results on the expected loss. The coefficient on FEMALE was negative and statistically significant, suggesting that lending to female borrowers was associated with an expected loss 0.008% lower than that for males in Column (5), or 18.6% lower than the sample mean. The coefficient on FEMALE was positive and significant in the regression for repayment ratio in Columns (7) and (8) of Table 5. The repayment ratio was 1.7% (0.016/0.947) more than the sample mean in Column (7). Thus, the results suggest that lending to females is more profitable and less risky.

Furthermore, the business performance indicators did not present strong impacts on the default rate and expected profit. The variables C_RATIO, W_CAPITAL, and R_INVEST presented values close to zero. W_CAPITAL and R_INVEST were not statistically significant in Columns (1) and (2) of Table 5 and C_RATIO was statistically significant in Columns (1) and (2). For expected profit, these variables were statistically significant and close to zero in Columns (3) and (4) of Table 5. In the regression for expected loss and repayment ratio, these variables were not statistically significant in Columns (5), (6), (7), and (8) of Table 5.

5.2 Gender gap across different cohorts

This section presents the Credibahia's analysis default rate using Logit estimation across different cohorts. These econometric models have interactions between the variables "FEMALE and AGE", "FEMALE and MARRIED", "FEMALE and INCOME", "FEMALE and GDP", "FEMALE and JOBS", "FEMALE and POP", and "FEMALE and VIOLENCE"(AARONSON et al., 2021; COLLINS; HEMBRE; URBAN, 2020).

The results in Table 6 show that female borrowers have lower loan default rates than male borrowers. Columns (1) and (2) of Table 6 show a negative and statistically significant interaction. The marginal effect shows that married female borrowers' default probability was 7.6% lower than for single female and 5.3% lower than that of males. Each additional year of the woman, the female borrowers' default probability was 0.5% lower than for men. However, Columns (3), (4), (5), (6) and (7) of Table 6 showed not statistically significant interactions. Thus, Columns (1) and (2) suggest that older and married women have lower loan default rate than men.

We also analyze the results by age cohorts in Table 7. We found that younger women - between 18 and 25 years old - have opposite results to those found in Table 5. The marginal effect shows that women's default rate in this age group was 4.0% higher than that of males, the expected profit was 4.4% higher than the sample mean, the expected loss was 9.3% higher than the sample mean, and the repayment ratio was 1.0% lower than the sample mean.

In the subsequent age cohorts, women are associated with lower default and expected loss rates, and higher expected profit and repayment ratios than their male peers. For example, women between 46 and 55 years old, the marginal effect shows that female's default rate in this age group was 30.8% lower than that of males, the expected profit was 45.1% higher than the sample mean, the expected loss was 27.9% lower than the sample mean, and the repayment ratio was 2.4% higher than the sample mean. And, women over 56 years old, the marginal effect shows that female's default rate was 38.2%

lower than men, the expected profit was 57.1% higher than the sample average, the expected loss was 37.2% lower than the sample average, and the repayment ratio was 3.1% higher than the sample average.

5.3 Endogeneity

The analysis of CrediBahia's loan performance indicate the existence of sample selection bias in the econometric model, once that borrowers auto-select to participate in the program. There are a number of significant methodological issues to be addressed when examining the impact of gender on loan performance. First, the literature suggests using the PSM to mitigate the potential estimation bias that might arise from the difference in sample size by gender (CHEN; HUANG; YE, 2020). Second, some unobserved or omitted variables may to bias our estimation results. We implemented the IV Probit estimator and 2SLS regression to solve the selection bias.

5.3.1 Propensity score matching estimation

The credit bureau data shows that the number of loans from female borrowers is higher than from male borrowers. Table 3 shows that females account for 65% of all borrowers. In order to analyze the problem of sample selection bias, we conducted an experiment with pairs of loans that are identical, except gender. The observed difference in loan performance across pairs would then be a robust estimate of the effect of gender.

According to Rosenbaum e Rubin (1983) and Heckman et al. (1998), propensity score matching (PSM) performs the function of construct good, matched samples based on the observed characteristics for the experiment. Thus, following Chen, Huang e Ye (2020), we use matching strategy in which for each loan borrowed by a woman, we adopt the closest match to choose the n (n = 1, 2, 3, 4, or 5) loans with the closest propensity score, and then compare their arithmetic mean loan performance. On average, the treatment group was a 20.7% lower default rate than the control group, female's expected profit was 38.3% higher than male, expected loss was 17.5% lower than men and repayment ratio was 1.9% lower than male borrowers. Thus, the results found in Table 8 suggest that the default rate of female borrowers is lower than that of male borrowers. In addition, loans to female borrowers have a higher expected profit and repayment ratio, and a lower expected loss than loans to male peers.

Heckman et al. (1998) points out that the identification of PSM estimators requires that the selection of treatment and non-treatment can be considered random after matching. This means that selection bias is caused by both observable and unobserved variables. We implemented two sensitivity analyses outlined by Mantel e Haenszel (1959) and Rosenbaum (2002) to estimate the extent to which the selection of unobserved may bias our inference about the gender difference in the default rate.

By supposing we have a pair of lists of loans i and j, if there is some hidden bias, two lists with the same observed parameter will have different odds of defaulting. In contrast, without hidden bias, i and j will have the same odds of being defaulted, meaning that their odds ratio Γ is equal to 1 (CHEN; HUANG; YE, 2020). Table 9 reports the Wilcoxon and Mantel-Haenszel statistics for the default rate. These two statistics indicate that the PSM estimate of default is free of hidden bias. Thus, the estimator suggests that there is an effect of gender on the default rate of the CrediBahia program.

5.3.2 Instrumental variable estimation

The finance and microcredit literature reports bias problems associated with simultaneity (SHAHRIAR; UNDA; ALAM, 2020; CLARK et al., 2020). One alternative is to use lag in the explanatory variable like Shamshur e Weill (2019). However, the explanatory variable in this research is binary.

Thus, we constructed the IV called IVGENDER, i.e., the ratio of the female labor force participation rate to the male labor force participation rate for each municipality as an instrumental variable (CHEN; LEUNG; GOERGEN, 2017).

Table 10 shows the estimation results. Staiger e Stock (1994) demonstrated that critical F-value is 8.96 when the number of instruments is one. The F-statistic in Columns (1) and (2) are 39.001 and 37.284, respectively, and we can reject the null hypothesis that the coefficient on the instrument is insignificantly different from zero at the 1% level. Thus, excluding the concern of a weak instrument. The OLS regression result presented in Columns (1) and (2) of Table 10 shows a negatively significant coefficient on the instrumental variable IVGENDER, meaning that the likelihood of a woman applying for microcredit is less in places with a higher ratio of the female labor force participation rate to the male labor force participation rate.

The IV Probit regression on the default probability in Columns (3) and (4) of Table 10 are not in line with the baseline estimations. These results shows that the effect of gender is not statistically significant. The 2SLS regression results are reported in Columns (5), (6), (7), (8), (9) and (10) of Table 10. The results in Column (6) show that the expected profit was 0.136% higher for female loans than for male loans. The estimation for expected loss and repayment ratio in Columns (7), (8), (9) and (10) of Table 10 showed that female borrowers were not in line with previous results. The female borrowers was associated with an expected loss 0.070% higher than that males peers in Column (7) and inconclusive result for repayment ratio in Columns (9) and (10) of Table 10. Therefore, the results suggest that lending to female borrowers has a higher expected profit than male borrowers.

5.4 Robustness checks

The previous results showed the use of instrumental variables to correct bias problems. In addition to this instrument, we demonstrate other methods for robustness checks. The first method is to remove outliers. We eliminated the observations whose variables AGE, INCOME, FLOW, C_RATIO, W_CAPITAL and R_INVEST in the 1st upper or lower percentile of the database and re-estimate the Logit model. The results without outliers present a reduction in the default rate, and an increase in lender profitability from financing female borrowers. These results are similar to those found in Table 5.

The second robustness check was the re-estimation of the models by the Probit method for the default rate and by cohorts. The results show that being an older and married woman reduces the default rate of microcredit. The results for the Probit models do not change the signs found in the Logit model results. Therefore, it brings arguments that the female gender is a better client and more profitable for the MFI than the male borrower. We do not show the robustness checks results for the sake of brevity. They are available upon request.

6 Conclusion

Gender disparities, associated with the economic activity of municipalities in Bahia, Brazil, can affect economic development in the region where the MFI is located. The better payment capacity of women in the program motivated us to investigate the causes for the lower default rate of women compared to men in CrediBahia. Using data from Desenbahia, IBGE and DATASUS, we showed that microcredit loan requested by female borrowers are associated with a lower default probability and expected loss and a higher expected profit and repayment ratio than male borrowers. The results showed that there was gap of gender in the ability to pay in the program, which are in line with the current existing literature on the subject. We considered that are important local variation in the features of the state economy that allowed to test how the gap of gender might be influenced by these features. Thus, the results also showed that the region's economic dynamics influences women's ability to pay. This is an important result to be considered by the MFI to improve the program, and also by the social and economic policy makers to design policies to reduce inequalities.

Future research may contribute to CrediBahia and the microfinance literature. It is possible to investigate the interaction between CrediBahia and the federal social programs (Bolsa Família program, for example), once those women also are in the target group of these programs, allowing to observe effects on the female borrower default rate, income, and consumption of low-income families.

This research indicates the existence of sample selection bias in the econometric model. We applied PSM and instrumental variables estimation as an alternative, but other methods can be explored. In addition to sample selection bias, subjective information could be analyzed for ability to pay. In this way, new information could be investigated to improve our model and the microfinance literature.

Acknowledgments

The authors are grateful to the Desenbahia administration that made possible this analysis and to Paola Gallo, Rodrigo Calabrich, Vinicius Mendes and Vinicius Felipe for valuable comments and inputs throughout the long gestation of this project. The findings, interpretations, and conclusions in this article are entirely those of the authors and they do not necessarily reflect the view of the Desenbahia.

References

AARONSON, D. et al. The long-run effects of the 1930s holc "redlining" maps on place-based measures of economic opportunity and socioeconomic success. *Regional Science and Urban Economics*, Elsevier, v. 86, p. 103–622, 2021.

AGIER, I.; SZAFARZ, A. Microfinance and gender: Is there a glass ceiling on loan size? *World Development*, Elsevier, v. 42, p. 165–181, 2013.

ALLEN, L.; SAUNDERS, A. Incorporating systemic influences into risk measurements: A survey of the literature. *Journal of financial services research*, Springer, v. 26, n. 2, p. 161–191, 2004.

ALTMAN, E. I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, JSTOR, v. 23, n. 4, p. 589–609, 1968.

ALTMAN, E. I. Predicting railroad bankruptcies in america. *The Bell Journal of Economics and Management Science*, JSTOR, p. 184–211, 1973.

ALTMAN, E. I. The success of business failure prediction models: An international survey. *Journal of Banking & Finance*, Elsevier, v. 8, n. 2, p. 171–198, 1984.

ASIEDU, E. et al. Access to credit by firms in sub-saharan africa: How relevant is gender? *American Economic Review*, v. 103, n. 3, p. 293–97, 2013.

ASSEFA, E.; HERMES, N.; MEESTERS, A. Competition and the performance of microfinance institutions. *Applied Financial Economics*, Taylor & Francis, v. 23, n. 9, p. 767–782, 2013.

ATERIDO, R.; BECK, T.; IACOVONE, L. Access to finance in sub-saharan africa: is there a gender gap? *World development*, Elsevier, v. 47, p. 102–120, 2013.

ATTANASIO, O. et al. *Group lending or individual lending? Evidence from a randomised field experiment in Mongolia.* [S.l.], 2014.

AUGSBURG, B. et al. *Microfinance, poverty and education*. [S.l.]: National Bureau of Economic Research Cambridge, MA, 2012.

BANERJEE, A. et al. *Can microfinance unlock a poverty trap for some entrepreneurs?* [S.l.], 2019.

BANERJEE, A. et al. The miracle of microfinance? evidence from a randomized evaluation. *American economic journal: Applied economics*, v. 7, n. 1, p. 22–53, 2015.

BANERJEE, A.; DUFLO, E.; SHARMA, G. *Long-term effects of the targeting the ultra poor program*. [S.l.], 2020.

BANERJEE, S. B.; JACKSON, L. Microfinance and the business of poverty reduction: Critical perspectives from rural bangladesh. *Human relations*, SAGE Publications Sage UK: London, England, v. 70, n. 1, p. 63–91, 2017.

BECK, T.; BEHR, P.; GUETTLER, A. Gender and banking: are women better loan officers? *Review of Finance*, Oxford University Press, v. 17, n. 4, p. 1279–1321, 2013.

BELLUCCI, A.; BORISOV, A.; ZAZZARO, A. Does gender matter in bank–firm relationships? evidence from small business lending. *Journal of Banking & Finance*, Elsevier, v. 34, n. 12, p. 2968–2984, 2010.

BESSIS, J. *Risk management in banking*. [S.l.]: John Wiley & Sons, 2011.

BROWN, M.; GUIN, B.; KIRSCHENMANN, K. Microfinance banks and financial inclusion. *Review of Finance*, Oxford University Press, v. 20, n. 3, p. 907–946, 2016.

BUERA, F. J.; KABOSKI, J. P.; SHIN, Y. The macroeconomics of microfinance. *The Review of Economic Studies*, Oxford University Press, v. 88, n. 1, p. 126–161, 2021.

BUVINIC, M.; BERGER, M. Sex differences in access to a small enterprise development fund in peru. *World Development*, Elsevier, v. 18, n. 5, p. 695–705, 1990.

CARLING, K. et al. Corporate credit risk modeling and the macroeconomy. *Journal of banking & finance*, Elsevier, v. 31, n. 3, p. 845–868, 2007.

CHEN, J.; LEUNG, W. S.; GOERGEN, M. The impact of board gender composition on dividend payouts. *Journal of Corporate finance*, Elsevier, v. 43, p. 86–105, 2017.

CHEN, X.; HUANG, B.; YE, D. Gender gap in peer-to-peer lending: Evidence from china. *Journal of Banking & Finance*, Elsevier, v. 112, p. 105–633, 2020.

CLARK, B. et al. Bank loan renegotiation and credit default swaps. *Journal of Banking & Finance*, Elsevier, p. 105936, 2020.

COLLINS, J. M.; HEMBRE, E.; URBAN, C. Exploring the rise of mortgage borrowing among older americans. *Regional Science and Urban Economics*, Elsevier, v. 83, p. 103524, 2020.

COSTA, P. de O.; LIMA, C. L. C.; SOUZA, M. M. F. de. Programa de microcrédito do estado da bahia (credibahia): Uma análise do arranjo organizacional de operação. *IX Congresso CONSAD de gestão pública. Brasília/DF, Brasil*, 2016.

CRÉPON, B. et al. Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in morocco. *American Economic Journal: Applied Economics*, v. 7, n. 1, p. 123–50, 2015.

CROSON, R.; GNEEZY, U. Gender differences in preferences. *Journal of Economic literature*, v. 47, n. 2, p. 448–74, 2009.

CULL, R.; DEMIRGÜÇ-KUNT, A.; MORDUCH, J. Microfinance meets the market. *Journal of Economic perspectives*, v. 23, n. 1, p. 167–92, 2009.

DEGRYSE, H.; ONGENA, S. Distance, lending relationships, and competition. *The Journal of Finance*, Wiley Online Library, v. 60, n. 1, p. 231–266, 2005.

DEMETRIADES, P.; FIELDING, D. Information, institutions, and banking sector development in west africa. *Economic Inquiry*, Wiley Online Library, v. 50, n. 3, p. 739–753, 2012.

DEMETRIADES, P.; LAW, S. H. Finance, institutions and economic development. *International journal of finance & economics*, Wiley Online Library, v. 11, n. 3, p. 245–260, 2006.

DESAI, J. et al. On the impact of microcredit: Evidence from a randomized intervention rural ethiopia. *Barcelona Graduate School of Economics Working Papers*, v. 741, 2013.

FIALA, N. Returns to microcredit, cash grants and training for male and female microentrepreneurs in uganda. *World Development*, Elsevier, v. 105, p. 189–200, 2018.

FRIEDSON-RIDENOUR, S.; PIEROTTI, R. S. Competing priorities: Women's microenterprises and household relationships. *World Development*, Elsevier, v. 121, p. 53–62, 2019.

FRYDMAN, H.; ALTMAN, E. I.; KAO, D.-L. Introducing recursive partitioning for financial classification: the case of financial distress. *The journal of finance*, Wiley Online Library, v. 40, n. 1, p. 269–291, 1985.

GERARDI, K. et al. Can't pay or won't pay? unemployment, negative equity, and strategic default. *The Review of Financial Studies*, Oxford University Press, v. 31, n. 3, p. 1098–1131, 2018.

GYOURKO, J.; TRACY, J. Reconciling theory and empirics on the role of unemployment in mortgage default. *Journal of Urban Economics*, Elsevier, v. 80, p. 87–96, 2014.

HAYRE, L.; MOHEBBI, C. Mortgage mathematics. *published in Handbook of Mortgage-Backed Securities, Frank Fabozzi (ed.), Chicago, Illinois: Probus Publishing Company*, p. 1223–1231, 1992.

HECKMAN, J. J. et al. *Characterizing selection bias using experimental data*. [S.l.]: National bureau of economic research Cambridge, Mass., USA, 1998.

HERTZBERG, A.; LIBERTI, J. M.; PARAVISINI, D. Information and incentives inside the firm: Evidence from loan officer rotation. *The Journal of Finance*, Wiley Online Library, v. 65, n. 3, p. 795–828, 2010.

HUANG, J.; KISGEN, D. J. Gender and corporate finance: Are male executives overconfident relative to female executives? *Journal of financial Economics*, Elsevier, v. 108, n. 3, p. 822–839, 2013.

KABOSKI, J.; J, R. T. A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica*, Wiley Online Library, v. 79, n. 5, p. 1357–1406, 2012.

LEE, J.; XU, J. Why do businesses grow faster in urban areas than in rural areas? *Regional Science and Urban Economics*, Elsevier, v. 81, p. 103521, 2020.

MANTEL, N.; HAENSZEL, W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the national cancer institute*, Oxford University Press, v. 22, n. 4, p. 719–748, 1959.

MCINTOSH, C.; JANVRY, A. D.; SADOULET, E. How rising competition among microfinance institutions affects incumbent lenders. *The economic journal*, Oxford University Press Oxford, UK, v. 115, n. 506, p. 987–1004, 2005.

MEAGER, R. Understanding the average impact of microcredit expansions: A bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics*, v. 11, n. 1, p. 57–91, 2019.

MERSLAND, R.; STRØM, R. Ø. Performance and governance in microfinance institutions. *Journal of Banking & Finance*, Elsevier, v. 33, n. 4, p. 662–669, 2009.

MORDUCH, J. The microfinance promise. *Journal of economic literature*, v. 37, n. 4, p. 1569–1614, 1999.

MURAVYEV, A.; TALAVERA, O.; SCHÄFER, D. Entrepreneurs' gender and financial constraints: Evidence from international data. *Journal of comparative economics*, Elsevier, v. 37, n. 2, p. 270–286, 2009.

NICKELL, P.; PERRAUDIN, W.; VAROTTO, S. Stability of rating transitions. *Journal of Banking & Finance*, Elsevier, v. 24, n. 1-2, p. 203–227, 2000.

PITT, M. M.; KHANDKER, S. R. The impact of group-based credit programs on poor households in bangladesh: Does the gender of participants matter? *Journal of political economy*, The University of Chicago Press, v. 106, n. 5, p. 958–996, 1998.

ROSENBAUM, P. R. Overt bias in observational studies. In: *Observational studies*. [S.l.]: Springer, 2002. p. 71–104.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, Oxford University Press, v. 70, n. 1, p. 41–55, 1983.

SANTOS, G. F. et al. Análise da trajetória tendencial e choques de investimento em equilíbrio geral dinâmico para o estado da bahia. *Revista Econômica do Nordeste*, v. 50, n. 2, p. 183–203, 2019.

SANTOS, G. Ferreira dos; RIBEIRO, L. C. de S.; CERQUEIRA, R. Barbosa de. The informal sector and covid-19 economic impacts: The case of bahia, brazil. *Regional Science Policy & Practice*, Wiley Online Library, v. 12, n. 6, p. 1273–1285, 2020.

SCHICKS, J. Over-indebtedness in microfinance–an empirical analysis of related factors on the borrower level. *World development*, Elsevier, v. 54, p. 301–324, 2014.

SCHROEDER, E. The impact of microcredit borrowing on household consumption in bangladesh. *Applied Economics*, Taylor & Francis, v. 52, n. 43, p. 4765–4779, 2020.

SCHULTE, M.; WINKLER, A. Drivers of solvency risk–are microfinance institutions different? *Journal of Banking & Finance*, Elsevier, v. 106, p. 403–426, 2019.

SHAHRIAR, A. Z. M. Exposure to intimate partner violence and repayment of microcredit: Evidence from field experiments in bangladesh. *Journal of Banking & Finance*, Elsevier, v. 72, p. 1–14, 2016.

SHAHRIAR, A. Z. M.; UNDA, L. A.; ALAM, Q. Gender differences in the repayment of microcredit: The mediating role of trustworthiness. *Journal of Banking & Finance*, Elsevier, v. 110, p. 105685, 2020.

SHAMSHUR, A.; WEILL, L. Does bank efficiency influence the cost of credit? *Journal of Banking & Finance*, Elsevier, v. 105, p. 62–73, 2019.

STAIGER, D. O.; STOCK, J. H. *Instrumental variables regression with weak instruments*. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1994.

STEINWAND, D. A risk management framework for microfinance institutions. *GTZ, Financial Systems Development, Eschborn, Germany*, p. 1–70, 2000.

STRØM, R. Ø.; D'ESPALLIER, B.; MERSLAND, R. Female leadership, performance, and governance in microfinance institutions. *Journal of Banking & Finance*, Elsevier, v. 42, p. 60–75, 2014.

WAGNER, C.; WINKLER, A. The vulnerability of microfinance to financial turmoil–evidence from the global financial crisis. *World Development*, Elsevier, v. 51, p. 71–90, 2013.

ZAMORE, S.; BEISLAND, L. A.; MERSLAND, R. Geographic diversification and credit risk in microfinance. *Journal of Banking & Finance*, Elsevier, v. 109, p. 105665, 2019.

Table 1 : The evolution of the CrediBahia

Year	Municipalities with Credibahia offices in operation	Amount released deflated	Number of contracts	Average value per contract	Maximum value of the program's financing	Maximum loan term of the program (months)	Interest rate
2002	5	R$ 390,584	402	R$ 971.60	R$ 5,000	12 months	1.8% per month
2003	14	R$ 1,401,860	1,581	R$ 886.69	R$ 5,000	12 months	1.8% per month
2004	32	R$ 2,432,191	2,734	R$ 889.61	R$ 5,000	12 months	1.8% per month
2005	61	R$ 3,872,626	4,494	R$ 861.73	R$ 5,000	12 months	1.8% per month
2006	116	R$ 10,639,007	11,660	R$ 912.44	R$ 5,000	12 months	1.8% per month
2007	127	R$ 9,812,173	11,379	R$ 862.31	R$ 5,000	12 months	1.8% per month
2008	134	R$ 9,111,109	10,817	R$ 842.30	R$ 5,000	Until May: 12 months After May: 18 months	1.8% per month
2009	144	R$ 9,213,142	12,120	R$ 760.16	Until August: R$ 5,000; After August: R$ 10,000	18 months	1.8% per month
2010	152	R$ 12,871,918	16,442	R$ 782.87	R$ 10,000	18 months	1.8% per month
2011	149	R$ 14,407,882	16,635	R$ 866.12	R$ 10,000	18 months	1.8% per month
2012	149	R$ 13,247,581	14,245	R$ 929.98	R$ 10,000	18 months	1.8% per month
2013	150	R$ 17,056,975	16,489	R$ 1,034.45	R$ 15,000	24 months	Until August: 1.2% per month. After August: Loans of up to R$ 3,000: 0.64% per month; Loans above R$ 3,000: 1.0% per month.
2014	163	R$ 20,806,797	20,545	R$ 1,012.74	R$ 15,000	24 months	0.41% per month
2015	162	R$ 22,965,525	23,460	R$ 978.92	R$ 15,000	24 months	Until April: 0.41% per month; After April: Loans of up to R$3,000: 0.55% per month; Loans of over R$3,000: 1.30% per month.
2016	157	R$ 17,926,004	18,836	R$ 951.69	R$ 15,000	24 months	Loans up to R$ 3,000: 1.2% per month; Loans above R$ 3,000: 1.8% per month.
2017	146	R$ 14,981,354	14,619	R$ 1,024.79	R$ 15,000	24 months	Until October: Loans up to R$ 3,000: 1.2% per month; Loans above R$ 3,000: 1.8% per month. After October: 2.0% per month.
2018	149	R$ 13,368,487	13,095	R$ 1,020.88	R$ 15,000	24 months	2.0% per month
2019	152	R$ 11,584,924	10,630	R$ 1,089.83	Until October: R$ 15.000 After October: R$ 21,000	24 months	2.0% per month

Table 2 : Variables and definitions

Variable	Name	Definition	Data source
Probability of Default	DEFAULT	1 if the funded loan has been defaulted, and 0 otherwise	Desenbahia
Expected Profit	EXPROFIT	Expected profit of a loan listing	Desenbahia
Expected Loss	EXPLOSS	Expected loss of a loan listing	Desenbahia
Repayment ratio	RR	Repayment ratio of a loan listing	Desenbahia
Marital Status	MARRIED	1 if a borrower is married, and 0 otherwise	Desenbahia
Gender of Borrower	FEMALE	1 if a borrower is female, and 0 otherwise	Desenbahia
Age	AGE	Age of a borrower (in years)	Desenbahia
Operating cash flow	FLOW	Monthly operating cash flow deflated of the business	Desenbahia
Loan Term (months)	MONTHS	Loan term requested by a borrower (in months)	Desenbahia
Loan Amount	AMOUNT	Loan amount released deflated by the borrower	Desenbahia
Borrower income	INCOME	Monthly borrower income deflated	Desenbahia
Number of loans applied	T_LOANS	The number of loans from the each borrower	Desenbahia
Interest Rate (%)	INTEREST	The annual interest rate that a borrower pays on the loan	Desenbahia
Current ratio	C_RATIO	Current ratio to measure a company's liquidity or ability to pay off short-term debts for each borrower	Desenbahia
Working Capital	W_CAPITAL	Working Capital measures an organization's currently available assets to meet short-term financial obligations for each borrower	Desenbahia
Return on Investment	R_INVEST	Return on Investment (ROI) is a performance measure used to evaluate the efficiency or profitability of an investment or compare the efficiency of a number of different investments for each borrower	Desenbahia
GDP of the municipality	GDP	Gross Domestic Product deflated for each municipality	IBGE
Loans per capita	CONT_MUN	Total number of contracts per capita for each municipality	IBGE
Instrumental variable	IVGENDER	The ratio of the female labor force participation rate to the male labor force participation rate for each municipality	IBGE
Formal employment per capita	JOBS	Formal employment in the private sector per capita for each municipality	IBGE
Higher education	HEDUC	The ratio of employees with higher education to the total number of employees in the formal labor market for each municipality	IBGE
Population	POP	Municipality population	IBGE
Violence	VIOLENCE	The number of homicides per municipality	DATASUS

Table 3 : Summary statistics and distribution of loan status

	Summary statistics									
Variable	N	Mean	St. Dev.	Min	Max	Variable	N	Mean	St. Dev.	Min	Max
DEFAULT	220,183	0.139	0.346	0	1	Year = 2002	220,183	0.002	0.043	0	1
EXPROFIT	220,183	0.091	0.126	-1	0.268	Year = 2003	220,183	0.007	0.084	0	1
EXPLOSS	163,150	0.043	0.203	0	1	Year = 2004	220,183	0.012	0.111	0	1
RR	163,150	0.947	0.224	0	1	Year = 2005	220,183	0.02	0.141	0	1
MARRIED	220,183	0.279	0.448	0	1	Year = 2006	220,183	0.053	0.224	0	1
FEMALE	220,183	0.651	0.477	0	1	Year = 2007	220,183	0.052	0.221	0	1
AGE	220,183	39.622	12.421	18	118	Year = 2008	220,183	0.049	0.216	0	1
FLOW	220,183	598.542	1,395.432	24.267	421,239	Year = 2009	220,183	0.055	0.228	0	1
MONTHS	220,183	9.123	4.076	1	24	Year = 2010	220,183	0.075	0.263	0	1
AMOUNT	220,183	936.00	641.476	53.854	9,102	Year = 2011	220,183	0.076	0.264	0	1
INCOME	220,183	791.180	1,411.692	0	311,900	Year = 2012	220,183	0.065	0.246	0	1
T_LOANS	220,183	2.848	2.341	1	26	Year = 2013	220,183	0.075	0.263	0	1
INTEREST	220,183	0.192	0.075	0.05	0.268	Year = 2014	220,183	0.093	0.291	0	1
C_RATIO	220,183	7.182	159.520	0	19,480	Year = 2015	220,183	0.107	0.309	0	1
W_CAPITAL	220,183	1,930.271	7,911.048	-84,074	1,066,935	Year = 2016	220,183	0.09	0.28	0	1
R_INVEST	220,183	41.310	275.874	0	113,867	Year = 2017	220,183	0.066	0.249	0	1
GDP	209,553	638,947	2,153,204	8,907	20,397,578	Year = 2018	220,183	0.059	0.237	0	1
CONT_MUN	220,183	0.006	0.005	0.00001	0.051	Year = 2019	220,183	0.048	0.214	0	1
IVGENDER	202,502	0.885	0.327	0.071	2.427						
JOBS	220,183	0.075	0.073	0.0005	0.705						
HEDUC	220,183	0.163	0.078	0.006	0.95						
POP	220,183	98,277	309,686	3,108	2,998,056						
VIOLENCE	220,183	41.118	151.301	1	1,854						

Distribution of loan status					
Loan status	N	Female	Proportion (%	Male	Proportion (%)
Funded	220,183	143,248	65%	76,935	35%
Default	30,512	18,219	60%	12,293	40%
Disaggregated default by sex		12.70%		16.00%	

Table 4 : Difference test

Difference test (Male vs. Female: Full Sample)			
Variable	Male	Female	Mean Diff
DEFAULT	0.16	0.127	0.033***
EXPROFIT	0.072	0.101	-0.029***
EXPLOSS	0.048	0.041	0.007***
RR	0.937	0.952	-0.015***
MARRIED	0.291	0.272	0.019***
AGE	39.495	39.689	-0.194***
FLOW	635	579	55.402***
MONTHS	9.357	8.997	0.36***
AMOUNT	930	939	-8.564**
INCOME	803	785	-18.863***
T_LOANS	2.668	2.945	-0.277***
INTEREST	0.193	0.192	0.001***
C_RATIO	7.365	7.083	0.282***
W_CAPITAL	2,144.88	1,815.01	329.867***
R_INVEST	36.907	43.674	-6.767***
GDP	580,385	670,526	-90.141***
CONT_MUN	0.006	0.006	0.000***
IVGENDER	0.902	0.876	0.026***
JOBS	0.069	0.078	-0.009***
HEDUC	0.162	0.164	-0.002***
POP	94,080	100,531	-6,451***
VIOLENCE	38.942	42.287	-3.345**

(a) Note: This table compares the descriptive statistics between male and female loan applicants and reports the mean difference test results in the column of "Mean Diff." *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5 : Logit regression result on default probability across different cohorts

	(1) DEFAULT	(2) DEFAULT	(3) EXPROFIT	(4) EXPROFIT	(5) EXPLOSS	(6) EXPLOSS	(7) RR	(8) RR
FEMALE	-0.287***	-0.285***	0.028***	0.028***	-0.008***	-0.007***	0.016***	0.015***
	(0.014)	(0.014)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
LNINCOME	-1.234***	-1.246***	0.121***	0.123***	-0.026***	-0.027***	0.055***	0.058***
	(0.016)	(0.016)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
AGE	-0.015***	-0.015***	0.001***	0.001***	-0.0002***	-0.0001***	0.001***	0.001***
	(0.001)	(0.001)	(0.00001)	(0.00001)	(0.00004)	(0.00004)	(0.00005)	(0.00005)
LNAMOUNT	0.147***	0.153***	-0.006***	-0.007***	0.001	0.003**	0.002	0.0003
	(0.018)	(0.019)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
LNFLOW	0.348***	0.353***	-0.035***	-0.036***	0.007***	0.007***	-0.014***	-0.014***
	(0.013)	(0.013)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
MARRIED	-0.234***	-0.238***	0.023***	0.024***	-0.003**	-0.002*	0.010***	0.010***
	(0.017)	(0.017)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
MONTHS	0.071***	0.072***	-0.007***	-0.007***	0.003***	0.002***	-0.005***	-0.004***
	(0.002)	(0.002)	(0.00003)	(0.00003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
T_LOANS	-0.105***	-0.106***	0.004***	0.005***	-0.004***	-0.003***	0.004***	0.004***
	(0.005)	(0.005)	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
INTEREST	0.577**	0.475	0.884***	0.876***	0.032	0.033*	-0.054**	-0.060***
	(0.291)	(0.292)	(0.004)	(0.004)	(0.020)	(0.019)	(0.022)	(0.021)
C_RATIO	-0.0001**	-0.0001**	0.00002***	0.00002***	0.00001	0.00001	-0.00001	-0.00001
	(0.0001)	(0.0001)	(0.00000)	(0.00000)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
W_CAPITAL	-0.00000	-0.00000	-0.00000**	-0.00000***	0.00000	-0.000	-0.00000	-0.00000
	(0.00000)	(0.00000)	(0.000)	(0.000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
R_INVEST	0.00002	0.00002	-0.00001***	-0.00001***	-0.00000	-0.00000	-0.00000	-0.00000
	(0.00002)	(0.00002)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
HEDUC		0.147		0.011***		0.0004		0.030**
		(0.143)		(0.002)		(0.012)		(0.013)
LNGDP		0.106**		0.002**		-0.013***		0.006
		(0.053)		(0.001)		(0.005)		(0.005)
LNPOP		-0.038		0.013***		0.020		0.004
		(0.084)		(0.001)		(0.020)		(0.022)
LNCONT_MUN		-0.055***		-0.003***		0.001		-0.001
		(0.015)		(0.0002)		(0.001)		(0.002)
JOBS		0.172		-0.005		0.007		0.048
		(0.514)		(0.007)		(0.052)		(0.057)
LNVIOLENCE		-0.011		-0.0003*		-0.001		-0.0003
		(0.013)		(0.0002)		(0.001)		(0.001)
Constant	-0.534	-0.120	-0.338***	-0.533***	0.202***	0.106	0.509***	0.420**
	(0.421)	(0.907)	(0.004)	(0.012)	(0.018)	(0.192)	(0.020)	(0.214)
Year	YES	YES	YES	YES	YES	YES	YES	YES
Municipality	YES	YES	YES	YES	YES	YES	YES	YES
Observations	220,183	209,553	220,183	209,553	163,150	152,574	163,150	152,574
R2	0.1161	0.1092	0.899	0.9	0.02	0.013	0.037	0.035
Adjusted R2			0.899	0.9	0.018	0.012	0.035	0.034
Log Likelihood	-78,153.350	-76,888.600						
Akaike Inf. Crit.	156,808.700	154,285.200						
Residual Std. Error			0.040	0.039	0.201	0.193	0.220	0.215
F Statistic			7,856.052***	7,454.089***	14.626***	9.153***	28.016***	24.982***

Note: (1) This table reports the regression results on loan performance and default probability. Columns (1) and (2) are estimated by Logit regression while Columns (3), (4), (5), (6), (7) and (8) are estimated by OLS regression. Year – a dummy controlling year effect. Municipality – a dummy variable reflecting the municipality in which a borrower is located. The prefix "LN"means the use of the natural log in the variable. (2) *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% levels, respectively. Ordinary standard errors are used, and Z-statistics are reported in parentheses. R2 is pseudo R-square.

Table 6 : Logit regression result on default probability across different cohorts

	(1) DEFAULT	(2) DEFAULT	(3) DEFAULT	(4) DEFAULT	(5) DEFAULT	(6) DEFAULT	(7) DEFAULT
FEMALE	-0.025	-0.259***	-0.280*	-0.262**	-0.275***	-0.360***	-0.285***
	(0.043)	(0.016)	(0.165)	(0.109)	(0.019)	(0.136)	(0.023)
LNINCOME	-1.246***	-1.248***	-1.246***	-1.246***	-1.246***	-1.246***	-1.246***
	(0.016)	(0.016)	(0.023)	(0.016)	(0.016)	(0.016)	(0.016)
FEMALE_AGE	-0.007***						
	(0.001)						
FEMALE_MARRIED		-0.109***					
		(0.031)					
FEMALE_INCOME			-0.001				
			(0.026)				
FEMALE_GDP				-0.002			
				(0.009)			
FEMALE_JOBS					-0.139		
					(0.186)		
FEMALE_POP						0.007	
						(0.013)	
FEMALE_VIOLENCE							0.0002
							(0.008)
AGE	-0.011***	-0.015***	-0.015***	-0.015***	-0.015***	-0.015***	-0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
LNAMOUNT	0.154***	0.154***	0.153***	0.153***	0.153***	0.153***	0.153***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
LNFLOW	0.354***	0.352***	0.353***	0.353***	0.353***	0.353***	0.353***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
MARRIED	-0.246***	-0.171***	-0.238***	-0.238***	-0.238***	-0.238***	-0.238***
	(0.017)	(0.025)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
MONTHS	0.072***	0.072***	0.072***	0.072***	0.072***	0.072***	0.072***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
T_LOANS	-0.106***	-0.106***	-0.106***	-0.106***	-0.106***	-0.106***	-0.106***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
INTEREST	0.479	0.477	0.475	0.475	0.475	0.475	0.475
	(0.292)	(0.292)	(0.292)	(0.292)	(0.292)	(0.292)	(0.292)
C_RATIO	-0.0001**	-0.0001**	-0.0001**	-0.0001**	-0.0001**	-0.0001**	-0.0001**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
W_CAPITAL	-0.00000	-0.00000	-0.00000	-0.00000	-0.00000	-0.00000	-0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
R_INVEST	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
HEDUC	0.143	0.146	0.147	0.147	0.147	0.147	0.147
	(0.143)	(0.143)	(0.143)	(0.143)	(0.143)	(0.143)	(0.143)
LNGDP	0.107**	0.106**	0.106**	0.107**	0.106**	0.106**	0.106**
	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)
LNPOP	-0.038	-0.039	-0.038	-0.038	-0.035	-0.043	-0.038
	(0.084)	(0.084)	(0.084)	(0.084)	(0.085)	(0.085)	(0.084)
LNCONT_MUN	-0.055***	-0.055***	-0.055***	-0.055***	-0.055***	-0.055***	-0.055***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
JOBS	0.167	0.165	0.172	0.173	0.275	0.171	0.172
	(0.514)	(0.514)	(0.514)	(0.514)	(0.532)	(0.514)	(0.514)
LNVIOLENCE	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)
Constant	-0.291	-0.123	-0.123	-0.136	-0.149	-0.075	-0.120
	(0.908)	(0.907)	(0.914)	(0.911)	(0.908)	(0.911)	(0.907)
Year	YES	YES	YES	YES	YES	YES	YES
Municipality	YES	YES	YES	YES	YES	YES	YES
R2	0.1094	0.1093	0.1093	0.1093	0.1093	0.1093	0.1092
Observations	209,553	209,553	209,553	209,553	209,553	209,553	209,553
Log Likelihood	-77,019.380	-77,033.270	-77,039.250	-77,039.230	-77,038.970	-77,039.100	-77,039.250
Akaike Inf. Crit.	154,548.800	154,576.500	154,588.500	154,588.500	154,587.900	154,588.200	154,588.500

Note: (1) This table reports the Logit regression results on default probability across different cohorts. Year – a dummy controlling year effect. Municipality – a dummy variable reflecting the municipality in which a borrower is located. The prefix "LN"means the use of the natural log in the variable. (2) *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% levels, respectively. Ordinary standard errors are used, and Z-statistics are reported in parentheses. R2 is pseudo R-square.

Table 7 : Loan performance across age cohorts

	(1) DEFAULT	(2) DEFAULT	(3) EXPROFIT	(4) EXPROFIT	(5) EXPLOSS	(6) EXPLOSS	(7) RR	(8) RR
LNINCOME	-1.240***	-1.252***	0.122***	0.124***	-0.025***	-0.027***	0.055***	0.058***
	(0.016)	(0.016)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
FEMALE_AGE25	0.062***	0.063***	0.003***	0.004***	0.005***	0.004**	-0.010***	-0.009***
	(0.023)	(0.023)	(0.0004)	(0.0004)	(0.002)	(0.002)	(0.002)	(0.002)
FEMALE_AGE2635	-0.202***	-0.200***	0.019***	0.019***	-0.005***	-0.004***	0.013***	0.013***
	(0.018)	(0.018)	(0.0003)	(0.0003)	(0.001)	(0.001)	(0.002)	(0.002)
FEMALE_AGE3645	-0.346***	-0.345***	0.031***	0.031***	-0.009***	-0.007***	0.018***	0.017***
	(0.020)	(0.020)	(0.0003)	(0.0003)	(0.001)	(0.001)	(0.002)	(0.002)
FEMALE_AGE4655	-0.497***	-0.492***	0.040***	0.041***	-0.013***	-0.012***	0.024***	0.023***
	(0.023)	(0.023)	(0.0003)	(0.0003)	(0.002)	(0.002)	(0.002)	(0.002)
FEMALE_AGE56	-0.668***	-0.674***	0.050***	0.052***	-0.018***	-0.016***	0.030***	0.029***
	(0.032)	(0.033)	(0.0004)	(0.0004)	(0.002)	(0.002)	(0.002)	(0.002)
LNAMOUNT	0.157***	0.162***	-0.007***	-0.007***	0.001	0.003**	0.002	-0.0001
	(0.018)	(0.018)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
LNFLOW	0.351***	0.357***	-0.035***	-0.036***	0.007***	0.007***	-0.014***	-0.015***
	(0.013)	(0.013)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
MARRIED	-0.269***	-0.274***	0.026***	0.027***	-0.003**	-0.002*	0.011***	0.010***
	(0.017)	(0.017)	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)
MONTHS	0.071***	0.072***	-0.007***	-0.007***	0.003***	0.002***	-0.005***	-0.004***
	(0.002)	(0.002)	(0.00003)	(0.00003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
T_LOANS	-0.112***	-0.113***	0.005***	0.006***	-0.003***	-0.003***	0.004***	0.004***
	(0.005)	(0.005)	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
INTEREST	0.595**	0.494*	0.883***	0.875***	0.032	0.033*	-0.055**	-0.061***
	(0.291)	(0.292)	(0.004)	(0.004)	(0.020)	(0.019)	(0.022)	(0.021)
C_RATIO	-0.0001**	-0.0001**	0.00002***	0.00002***	0.00001	0.00001	-0.00001	-0.00001
	(0.0001)	(0.0001)	(0.00000)	(0.00000)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
W_CAPITAL	-0.00000	-0.00000	-0.000	-0.00000	0.00000	-0.000	-0.00000	-0.00000
	(0.00000)	(0.00000)	(0.000)	(0.000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
R_INVEST	0.00002	0.00002	-0.00001***	-0.00001***	-0.00000	-0.00000	-0.00000	-0.00000
	(0.00002)	(0.00002)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
HEDUC		0.132		0.012***		0.0005		0.030**
		(0.143)		(0.002)		(0.012)		(0.013)
LNGDP		0.107**		0.002**		-0.013***		0.007
		(0.053)		(0.001)		(0.005)		(0.005)
LNPOP		-0.039		0.013***		0.020		0.004
		(0.084)		(0.001)		(0.020)		(0.022)
LNCONT_MUN		-0.053***		-0.003***		0.001		-0.001
		(0.015)		(0.0002)		(0.001)		(0.002)
JOBS		0.180		-0.006		0.006		0.048
		(0.514)		(0.007)		(0.052)		(0.057)
LNVIOLENCE		-0.011		-0.0003		-0.001		-0.0004
		(0.013)		(0.0002)		(0.001)		(0.001)
Constant	-1.108***	-0.683	-0.286***	-0.484***	0.190***	0.096	0.539***	0.452**
	(0.421)	(0.907)	(0.004)	(0.013)	(0.018)	(0.192)	(0.020)	(0.214)
Year	YES	YES	YES	YES	YES	YES	YES	YES
Municipality	YES	YES	YES	YES	YES	YES	YES	YES
Observations	220,183	209,553	220,183	209,553	163,150	152,574	163,150	152,574
R2	0.1165	0.1083	0.888	0.888	0.020	0.014	0.037	0.036
Adjusted R2			0.888	0.888	0.019	0.012	0.036	0.034
Log Likelihood	-78,223.540	-76,957.110						
Akaike Inf. Crit.	156,955.100	154,428.200						
Residual Std. Error			0.042	0.041	0.201	0.193	0.220	0.215
F Statistic			6,905.929***	6,460.090***	14.796***	9.283***	28.028***	24.986***

Note: (1) This table reports the Logit regression results on default probability across age cohorts. Columns (1) and (2) are estimated by Logit regression while Columns (3), (4), (5), (6), (7) and (8) are estimated by OLS regression. FEMALE_AGE25 represents age between 18 and 25 years old; FEMALE_AGE2635 represents age between 26 and 35 years old; FEMALE_AGE3645 represents age between 36 and 45 years old; FEMALE_AGE4655 represents age between 46 and 55 years old; FEMALE_AGE56 represents age above 56 years old. Year – a dummy controlling year effect. Municipality – a dummy variable reflecting the municipality in which a borrower is located. The prefix "LN"means the use of the natural log in the variable. (2) *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% levels, respectively. Ordinary standard errors are used, and Z-statistics are reported in parentheses. R2 is pseudo R-square.

Table 8 : PSM estimation results

Variables	Sample	Treated: Female = 1	Control: Female = 0	ATT
Panel A	One-to-one matching			
DEFAULT	After match	0.1272	0.1602	-0.034528***
EXPROFIT	After match	0.1010	0.0734	0.0276044***
EXPLOSS	After match	0.0410	0.0502	-0.0091785***
RR	After match	0.9520	0.9336	0.01840315***
Panel B	One-to-two matching			
DEFAULT	After match	0.1272	0.1610	-0.03325003***
EXPROFIT	After match	0.1010	0.0730	0.02795167***
EXPLOSS	After match	0.0410	0.0501	-0.0090583***
RR	After match	0.9520	0.9348	0.01715994***
Panel C	One-to-three matching			
DEFAULT	After match	0.1272	0.1601	-0.0336991***
EXPROFIT	After match	0.1010	0.0732	0.02784303***
EXPLOSS	After match	0.0410	0.0500	-0.0090337***
RR	After match	0.9520	0.9338	0.0181505***
Panel D	One-to-four matching			
DEFAULT	After match	0.1272	0.1600	-0.03357115***
EXPROFIT	After match	0.1010	0.0726	0.02839154***
EXPLOSS	After match	0.0410	0.0493	-0.0082518***
RR	After match	0.9520	0.9344	0.01755509***
Panel E	One-to-five matching			
DEFAULT	After match	0.1272	0.1606	-0.03357813***
EXPROFIT	After match	0.1010	0.0730	0.02804703***
EXPLOSS	After match	0.0410	0.0491	-0.0081284***
RR	After match	0.9520	0.9347	0.01732355***

Note: (1) We use the nearest neighbor matching of 1:1, 1:2, 1:3, 1:4, and 1:5. (2) The variables used for matching include: LNINCOME – natural log of monthly borrower income deflated; AGE – the age of a borrower expressed in years; LNAMOUNT – natural log of loan amount requested by the borrowers; LNFLOW - natural log of monthly operating cash flow deflated of the business; MARRIED – a dummy variable taking a value of 1 if a borrower is married, and 0 otherwise; MONTHS - loan term requested by a borrower (in months); T_LOANS – the number of times that a borrower has applied for a loan; INTEREST – the interest rate that the borrower pays on the loan; C_RATIO - current ratio to measure a company's liquidity or ability to pay off short-term debts for each borrower; W_CAPITAL - working Capital measures an organization's currently available assets to meet short-term financial obligations for each borrower; R_INVEST - Return on Investment (ROI); HEDUC - the ratio of employees with higher education to the total number of employees in the formal labor market for each municipality; LNGDP - natural log of Gross Domestic Product deflated for each municipality; LNPOP - natural log of municipality population; LNCOUNT_MUN - natural log of total number of contracts per capita for each municipality; JOBS - formal employment per capita for each municipality; LNVIOLENCE - natural log of the number of homicides per municipality; Year – a dummy controlling year effect. Municipality – a dummy variable reflecting the municipality in which a borrower is located. (3) The treatment group is female borrowers.

Table 9 : Hidden bias in propensity score matching results

Panel A	Default: One-to-one matching			
	Wilcoxon Statistics		Mantel-Haenszel Statistics	
Γ	Upper bound	Lower bound	Upper bound	Lower bound
1.0000	0.0000	0.0000	0.0000	0.0000
1.0500	0.0000	0.0000	0.0000	0.0000
1.1000	0.0000	0.0000	0.0000	0.0000
1.1500	0.0000	0.0000	0.0000	0.0000
1.2000	0.0000	0.0000	0.0000	0.0000
1.2500	0.0000	0.0000	0.0000	0.0132
1.3000	0.0000	0.0000	0.0000	0.2929
1.3500	0.0000	0.0000	0.0000	0.0006
1.4000	0.0000	0.0000	0.0000	0.0000
1.4500	0.0000	0.0000	0.0000	0.0000
1.5000	0.0000	0.0000	0.0000	0.0000
1.5500	0.0000	0.0000	0.0000	0.0000
1.6000	0.0000	0.0000	0.0000	0.0000
1.6500	0.0000	0.0000	0.0000	0.0000
1.7000	0.0000	0.0000	0.0000	0.0000
1.7500	0.0000	0.0000	0.0000	0.0000
1.8000	0.0000	0.0000	0.0000	0.0000
1.8500	0.0000	0.0000	0.0000	0.0000
1.9000	0.0000	0.0000	0.0000	0.0000
1.9500	0.0000	0.0000	0.0000	0.0000
2.0000	0.0000	0.0000	0.0000	0.0000

Table 10 : Instrumental variable regression results

	(1) FEMALE	(2) FEMALE	(3) DEFAULT	(4) DEFAULT	(5) EXPROFIT	(6) EXPROFIT	(7) EXPLOSS	(8) EXPLOSS	(9) RR	(10) RR
IVGENDER	-0.021***	-0.018**								
	(0.008)	(0.009)								
FEMALE			0.197	-14.54	0.135***	0.136**	0.070***	0.029	-0.056**	0.445***
			(1.23)	(0.82)	(0.050)	(0.061)	(0.022)	(0.107)	(0.024)	(0.163)
LNINCOME	-0.063***	-0.061***	-0.548***	-1.267	0.156***	0.158***	-0.012***	-0.018**	0.035***	0.075***
	(0.002)	(0.002)	(53.34)	(1.49)	(0.003)	(0.004)	(0.002)	(0.008)	(0.002)	(0.011)
AGE	-0.00002	-0.0001	-0.00798***	-0.00940***	0.002***	0.002***	-0.0002***	-0.0001***	0.001***	0.001***
	(0.0001)	(0.0001)	(25.43)	(5.32	(0.00001)	(0.00001)	(0.00004)	(0.00005)	(0.00005)	(0.0001)
LNAMOUNT	0.033***	0.034***	0.0701***	0.643	-0.009***	-0.010***	-0.002	0.002	0.005***	-0.022**
	(0.003)	(0.003)	(5.93)	(0.93)	(0.002)	(0.002)	(0.002)	(0.006)	(0.002)	(0.009)
LNFLOW	-0.066***	-0.067***	0.215***	-0.716	-0.037***	-0.038***	0.013***	0.010	-0.019***	0.016
	(0.002)	(0.002)	(20.03)	(0.64)	(0.003)	(0.004)	(0.002)	(0.008)	(0.002)	(0.012)
MARRIED	-0.012***	-0.012***	-0.158***	-0.311	0.033***	0.034***	-0.005***	-0.004***	0.013***	0.015***
	(0.003)	(0.003)	(17.76)	(1.58)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
MONTHS	-0.008***	-0.008***	0.0362***	-0.124	-0.009***	-0.010***	0.003***	0.002**	-0.005***	0.0004
	(0.0004)	(0.0004)	(15.60)	(0.62)	(0.0004)	(0.001)	(0.0003)	(0.001)	(0.0003)	(0.002)
T_LOANS	0.014***	0.014***	-0.0704***	0.152	0.007***	0.008***	-0.005***	-0.004***	0.006***	-0.001
	(0.001)	(0.001)	(20.43)	(0.56)	(0.001)	(0.001)	(0.0004)	(0.002)	(0.0004)	(0.002)
INTEREST	0.153***	0.156***	0.529**	4.209	0.776***	0.764***	0.015	0.026	-0.039*	-0.132***
	(0.047)	(0.048)	(3.15)	(0.91)	(0.010)	(0.011)	(0.021)	(0.027)	(0.023)	(0.041)
C_RATIO	0.00001**	0.00001**	-0.0000458	0.0000253	0.00002***	0.00002***	-0.00001	-0.00001	0.00001	0.00001
	(0.00001)	(0.00001)	(1.83)	(0.20)	(0.00000)	(0.00000)	(0.00002)	(0.00002)	(0.00002)	(0.00003)
W_CAPITAL	-0.00000***	-0.00000***	-0.00000122*	-0.0000109	-0.00000	-0.00000	0.00000	0.000	-0.00000	0.00000*
	(0.00000)	(0.00000)	(1.98)	(0.88)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
R_INVEST	0.00002***	0.00002***	0.0000358	0.000311	-0.00001***	-0.00001***	-0.00000	-0.00000	-0.00000	-0.00003***
	(0.00000)	(0.00000)	(1.59)	(0.88)	(0.00000)	(0.00000)	(0.00000)	(0.00001)	(0.00000)	(0.00001)
HEDUC		0.003		0.0674		0.011***		0.018*		0.004
		(0.027)		(0.16)		(0.003)		(0.010)		(0.015)
LNGDP		0.012		0.499		-0.001		-0.003		-0.014**
		(0.009)		(0.77)		(0.001)		(0.004)		(0.005)
LNPOP		-0.038***		-0.539		0.014***		0.004		0.016***
		(0.015)		(0.86)		(0.003)		(0.003)		(0.005)
LNCONT_MUN		-0.016***		-0.255		-0.001		-0.002		0.013***
		(0.003)		(1.12)		(0.001)		(0.002)		(0.003)
JOBS		-0.038		1.457		0.017		0.018		-0.073***
		(0.081)		(1.12)		(0.010)		(0.016)		(0.025)
LNVIOLENCE		0.006***		0.107		-0.001**		0.001**		-0.006***
		(0.002)		(1.23)		(0.0004)		(0.001)		(0.001)
Constant	1.291***	1.432***	-0.334*	14.27	-1.128***	-1.358***	0.058**	0.049	0.808***	0.321**
	(0.037)	(0.154)	1.97	(0.86)	(0.069)	(0.094)	(0.024)	(0.106)	(0.027)	(0.161)
Year	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Municipality	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	202,502	191,872	202,502	191,872	202,502	191,872	150,294	139,718	150,294	139,718
R2	0.044	0.045			0.840	0.838	-0.017	0.002	0.002	-0.830
Adjusted R2	0.043	0.044			0.840	0.838	-0.017	0.001	0.002	-0.831
Residual Std. Error	0.466	0.466			0.060	0.058	0.205	0.194	0.224	0.295
F Statistic	39.001***	37.284***								

Note: (1) This table reports the OLS regression on female, IV Probit regression results on default probability and the 2SLS regression results on expected profit. Columns (1), and (2) are estimated by OLS regression, Columns (3) and (4) are estimated by IV Probit and Columns (5), (6), (7), (8), (9) and (10) are estimated by 2SLS regression. Year – a dummy controlling year effect. Municipality – a dummy variable reflecting the municipality in which a borrower is located. The prefix "LN"means the use of the natural log in the variable. (2) *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% levels, respectively. Ordinary standard errors are used, and Z-statistics are reported in parentheses. R2 is pseudo R-square.