Consumer Debt and Poverty: the Default Risk Gap*

Lucía Bertoletti   Fernando Borraz   Graciela Sanroman †

Abstract

This paper examines the disparity in default risk between vulnerable and non-vulnerable populations in consumer lending. We merge an exhaustive registry of loans granted in the financial system with microdata on vulnerable individuals applying for social programs. We estimate the sources of this disparity and how loan and individual characteristics influence the probability of default. We find that vulnerable individuals have a higher risk than non-vulnerable individuals. However, this difference is reduced when individual debt characteristics, particularly the interest rate, are considered. Specifically, interest rates explain at least 30 percent of the risk gap. We also find that the default probabilities faced by lending firms are higher than those faced by banks, but we show that this effect is partly due to interest rate divergences. Our study underscores the importance of considering individual characteristics, loan characteristics, and interest rates when assessing default risk. While recognizing their limitations, these results suggest the need for policy interventions to promote financial inclusion, fair interest rate practices, and financial education, especially for vulnerable populations.

Keywords: consumer lending, default, interest rate, poverty

JEL classification codes: G21, G23, G51

*We are grateful to Banco Central del Uruguay for providing us with the credit data and the Ministerio de Desarrollo Social for the vulnerable data. We thank CSIC UDELAR for funding our project. We also thank Elisa Failache and seminar participants at Universidad ORT Uruguay, Universidad de la Plata, and Universidad de la Repúblicafor for their helpful comments and suggestions.

†dECON Universidad de la República del Uruguay. Emails: lucia.bertoletti@cienciassociales.edu.uy, fernando.borraz@cienciassociales.edu.uy, graciela.sanroman@cienciassociales.edu.uy.

1 Introduction

For a long time, consumer lending has served as a crucial engine for economic growth, enabling individuals to access goods and services beyond their immediate means. Besides, consumer lending is one of the principal activities of financial institutions. The study of the consumer credit segment of the financial system is of great interest due to its impact on financial stability and people's well-being.

From financing durable and nondurable goods, entrepreneurship, and housing to consolidating debt and pursuing educational opportunities, consumer credit plays a significant role in shaping individual financial trajectories and overall well-being. Moreover, borrowing capacity is crucial to smooth consumption over the life cycle or when faced with adverse shocks.

However, the ease and accessibility of credit often come at a cost, manifested in the form of interest rates. While responsible lending practices and informed borrowing decisions can lead to positive outcomes, high interest rates can trap individuals in a cycle of debt, eroding their financial security and potentially exacerbating social inequalities.

The past 30 years have witnessed a significant surge in household borrowing. This phenomenon can be partially attributed to the growing acceptance of credit as a tool to manage financial fluctuations and ensure a stable standard of living even when faced with unexpected expenses or income disruptions. However, it is crucial to recognize that consumer credit is only sometimes beneficial, and its suitability as a financial product varies considerably across different population segments.

Overall, the literature on consumer lending for poor individuals highlights the importance of access to credit, the impact of high interest rates, the emergence of alternative lending options, behavioral factors influencing borrowing behavior, and potential policy interventions. These findings contribute to a better understanding of the challenges poor individuals face in the consumer lending landscape (Durkin and Elliehausen, 2014, Dobridge, 2016, Dobbie et al., 2021, Fuster et al., 2022, Gardner, 2022, Bartlett et al., 2022 and Becher et al., 2023).

In this paper, we focus on the performance of individuals in the consumer credit segment and its relationship with poverty using data from Uruguay. We investigate the sources of the gap in default risk between vulnerable and non-vulnerable borrowers, focusing on the role of loan characteristics and individual borrower traits. We combine two rich sources of information. On the one hand, we use data on debtors at the micro level, including their links with the financial system, credit rating, and payment history. On the other hand, we have data on those who have applied for social assistance plans and their socioeconomic status. We merge both datasets by their anonymized identity card number.

We distinguish between the two main types of financial institutions offering consumer credit in Uruguay: banks and lending firms. The former mainly has medium and high-income people as clients with higher entry barriers, and the latter provides credits more quickly and with fewer demands; however, it does so at a higher interest rate (Arnabal et al., 2023). Therefore, low-income people typically find the solution to their loan needs in lending firms. We estimate two default models: one in which the covariates are entered linearly and another, more flexible, selected using a CV-Lasso algorithm.

We find that vulnerable individuals exhibit significantly higher default probabilities than non-vulnerable people. We refer to this difference as the "risk gap" and aim to disentangle whether the observed unconditional gap is due to the borrower's characteristics or limited access to inferior financial products, specifically loans with higher interest rates.

Our findings have direct policy implications, as they highlight the importance of financial education for the most vulnerable and regulatory changes to improve the quality of information for individuals, demonstrating that usury limits play a crucial role.

Our work is related to two areas of the literature. The first is the relationship between consumer lending, credit score, bias, and discrimination (Ghent et al., 2014, Dobbie et al., 2021, Bartlett et al., 2022). Second, our work is related to lending and information where vulnerable households, poor consumers in particular, with lesser bargaining power, often make counterproductive financial decisions, are over-optimist, and end up with high-cost credits (Elliehausen and Lawrence, 2009, Durkin and Elliehausen, 2014, Gardner, 2022, Becher et al., 2023). However, not all effects are adverse. Dobridge, 2016 indicates that consumer lending improves the well-being of distressed households by helping them smooth consumption.

This paper delves into the multifaceted impact of consumer lending on individuals, particularly the critical interplay between its importance and the often-hidden cost associated with interest rates for poor individuals. We analyze the potential pitfalls of high-interest debt and its detrimental effects on personal well-being and economic mobility. Through a critical analysis of the rich individual-level datasets, this paper aims to shed light on the complex relationship between consumer lending and individual lives, sparking further discussion and promoting responsible lending practices and informed borrowing decisions.

The rest of the paper is organized as follows. Section 2 presents the related literature. Section 3 describes the data used in this paper and the selected sample. Section 4 presents the empirical strategy. Section 5 reports the results of estimating the default risk gap and the influence of interest rates. Section 6 describes the theoretical insights to interpret the empirical results. Finally, section 7 provides a discussion and conclusion.

2 Literature

Despite there being no consensus about the definition of consumer lending, it is generally the purchase payment with an advance of cash from a financial institution or retailer. Consumer lending is one of the principal activities of financial institutions.

Our work is related to two areas of the literature. The first studies the relationship between consumer lending, credit score, bias, and discrimination. (Ghent et al., 2014, Dobbie et al., 2021, Fuster et al., 2022, Bartlett et al., 2022). Second, our work is related to loans and information where vulnerable households, poor consumers in particular, with lesser bargaining power, often make counterproductive financial decisions, are over-optimist, and end up with high-cost credits (Durkin and Elliehausen, 2014, Elliehausen and Lawrence, 2009, Becher et al., 2023, Gardner, 2022). This population allocates a significant proportion of their income to commodities (Kaplan et al., 2014); they are "hand-to-mouth" consumers, also characterized by their lack of access to consumption-smoothing technologies (Camara, 2022). However, not all effects are adverse. Dobridge (2016) indicate that consumer lending improves the well-being of distressed households by helping them smooth consumption.

As mentioned above, our work is related to the relationship between consumer lending, credit score, bias, and discrimination. There are several ways to measure and address bias in consumer lending, each offering its strengths and limitations. Dobbie et al. (2021) focuses on methodologies for measuring bias in consumer lending, and Bartlett et al. (2022) analyzes discrimination in consumer lending through fintech. Using machine learning, Fuster et al. (2022) concluded that the statistical technology used to evaluate creditworthiness has a negative distributional impact on black and white Hispanics.

The statistical tests analyze loan approval rates, interest rates, and other lending outcomes across different demographic groups, controlling for relevant factors like credit score, income, and loan amount. For example, studies by Bhutta et al. (2022) use this approach to identify racial and gender disparities in mortgage lending. Another strand of the literature uses machine learning algorithmic models in loan decisions, checking for unfair biases built into the data or the algorithms themselves. Dobbie et al. (2021) found that some credit scoring algorithms exhibited discrimination bias against immigrants and older applicants. Matched comparison involves comparing similar borrowers from different demographic groups who have applied for the same loan, controlling for all other relevant factors. A study by Butler et al. (2022) used this method to find evidence of discrimination and racial bias in auto loan approvals. The field experiments are uncommon and costly, and different loan offers are randomly assigned to borrowers to see if their demographics affect their choices and outcomes. Hanson et al. (2016) used this approach to show that lenders offered higher interest rates to Black borrowers than White borrowers with similar creditworthiness.

The cost of borrowing is typically expressed as the annual percentage rate, which includes the stated interest rate and any additional fees charged by the lender. However, for poor borrowers, who often rely on high-cost credit options, other factors also influence the actual cost of borrowing: short-term loans that create a cycle of debt as borrowers struggle to repay within the short timeframe (Bhuta et al. (2015) and Spector (2008)), fees and penalties and deceptive marketing, bait-and-switch tactics, and hidden costs can trap borrowers in unfair loan terms that exacerbate their financial difficulties. Dobridge (2016), and Elliehausen and Lawrence (2009) analyze consumers' use of payday loans and the potential positive and negative impacts of high-cost credit, and Gardner (2022) explores the future of high-cost credit and considers alternative solutions like payday lending reform.

For a developing country, Castellanos et al. (2020) find that despite rising as a de facto financial tool for low-income first-time borrowers in Mexico, credit cards pose a risk due to high unpredictable defaults. Standard solutions like raising minimum payments or interest rates prove ineffective, while unemployment significantly impacts repayment. This highlights the need for alternative risk-management strategies and the complexity of financial inclusion for vulnerable populations.

In a randomized control trial, Brune et al. (2021) show the effectiveness of a simple savings program in promoting saving, enabling purchases of durable goods, and potentially leading to long-term improvements in participants' lives. It also highlights the need for readily available and effective savings options, particularly for those struggling with self-control.

The only related study for Uruguay is Arnabal et al. (2023), which analyzes individual debt from a descriptive point of view without addressing the issue of consumer bias discrimination and the cost of lending.

To address bias, the proposed solutions are regulatory interventions (limitations on credit scoring methods and requirements for lenders to document their decision-making processes), algorithmic fairness tools, and financial education and counseling that help borrowers understand creditworthiness, compare loan offers, and manage debt that can reduce disparities in borrowing outcomes.

By understanding how various factors contribute to the cost of borrowing and exploring strategies to address bias and predatory practices, we can work towards ensuring fair and equitable access to credit for all borrowers, especially those experiencing financial vulnerability. However, whether the market solution is viable for the highly vulnerable group is still being determined. One option is affordable access to credit for this group in addition to the public lump sum transfers. This is particularly important in the Uruguayan context, as the cost of committing default includes not being able to contract some essential services, difficulties in obtaining rental guarantee, and limited access to credit, which can lead to borrowing on the informal market, leading to a poverty trap for the most vulnerable population.

3 Data

Our empirical evidence is derived from two highly reliable administrative records sources: the Credit Bureau of the Uruguayan Central Bank (CB) and data from the Ministry of Social Development (MIDES). The former provides comprehensive loan data, while the latter offers valuable insights into vulnerable households at the individual level, with data linked through national identifier numbers.

3.1 Data from Credit Bureau of the Uruguayan Central Bank

The loan data from the Uruguayan Central Bank Credit Bureau is a comprehensive record of debt granted to Uruguayan individuals in the financial system. This exhaustive public information, provided monthly on a mandatory basis by regulated financial institutions, allows for a detailed analysis of the loan product level. The dataset spans from January 2014 to April 2023, with loans from lending companies becoming available from 2014 onwards.

In this paper, we focus on the performance of individuals in the consumer credit segment, so we take only consumer loans, including traditional and credit card loans. Hence, this dataset includes all persons who hold credit cards even when not using them in a particular month. Also, it is worth mentioning that if somebody has an old unpaid debt, she appears every month until she pays the complete amount that she owes[1].

The information contained in this dataset includes the total amount of the loan, the amount of the loan on each status (current debt, matured debt, debt under management, overdue debt, and written-off debt), the provider institution, the currency of the loan, the preapproved debt (contingency), and the individual credit rating category assigned by each institution according to the regulation. The credit rating category varies from 1 (strong payment capacity) to 5 (irrecoverable).

The credit rating category, a key determinant of a debtor's financial health, is primarily influenced by the time in arrears on debt payments. In our research, we define the *default* condition as having a credit category 4 or 5, indicating a delinquency of 90 days or more. This categorization is crucial as it provides a clear benchmark for assessing the financial stability of debtors.

As Table 1 shows, around 2.5 million individuals had at least one consumer debt during the analyzed period, approximately the number of persons age 18 and older (the Uruguayan population is around 3.4 million). Most had bank debt and a credit card, and 55% had lending firm debt. Moreover, 49% of the debtors have defaulted at least once.

While the interest rate for each loan is unavailable on this dataset, we have the monthly average rate for all the financial systems and the monthly limit for usury established by the Central Bank from public Central bank records. The limit of usury is calculated based on the average rate of the last three months. Lending firms tend to fix their interest rates at the limit, while banks set lower interest rates. We define a proxy for the interest rate for each type of institution as follows: the 75% of the average monthly rate for banks, and the limit for usury for lending firms[2]. During the analysis period, the proxy of the bank rate is around 56% while the lending firm's rate is around 130%. Figure 2 presents information on the average interest rate and usury limits, demonstrating significant variability over the period analyzed.

Furthermore, since July 2023, we have collected the interest rate for each bank and lending firm from its websites. Although financial institutions must provide the interest rate monthly on their website, the Central Bank does not provide this information. Moreover, we can not

[1]For example, someone takes a loan on January 2014 and stops paying it on July 2014 but is not yet sold, she appears on the April 2023 records unless the financial institution gives it as a loss and stops to reported it

[2]The limit for usury is established by the Central Bank at the 190% of the average interest rate of the three previous months

build a time series backward dataset for the institution rates, but we have the information since July 2023. This data supports our proxy for bank and lending firms' rates previously explained.

3.2 Vulnerable households data

Our second dataset comes from MIDES and contains vulnerable household data. MIDES is the central government agency in charge of public social assistance in Uruguay. Two of its more crucial non-contributory programs are *Asignaciones Familiares - Plan de Equidad* (AFAM) and *Tarjeta Uruguay Social* (TUS).

The AFAM programs target poor families with children under age 18 and pregnant women, and the amount of the transfer depends on the number of children in the household and whether they are enrolled in school. The transfer is conditioned on the minors' school attendance and regular health checks. To become eligible for this program, households must meet the two criteria. First, households must have a score (ICC) above a certain threshold. This score is based on an application completed by the household that captures an array of socioeconomic data. The second condition they need is to earn less than a specific monthly formal per capita income. If a household surpasses this threshold for three consecutive months, it loses its eligibility status. The cash transfer amount depends non-linearly on the number of children in each household (up to a maximum of four) [3].

The TUS program aims to "provide food support to people in extreme poverty" (DINEM, 2011). It targets households (with or without minors) in extreme socio-economic vulnerability conditions, as measured by the ICC. Additionally, a household receives double the amount if MIDES concludes it is within the poorest families in the country (double TUS).

Therefore, households can enter or exit the programs mainly by receiving a household visit from a MIDES agent. There are two types of visits: area visits, conducted to all households in a specific geographical area, and targeted visits, conducted to current beneficiaries to check whether they are still eligible or to households that request to become new recipients.

We define the vulnerable group as individuals eligible for the AFAM program and the highly vulnerable group as individuals eligible for the TUS program. Notably, households enrolled in the TUS program may also qualify for AFAM benefits. In such cases, we classify them as part of the AFAM group if they are eligible only for the AFAM program and not for TUS

[3] See Lagomarsino (2021) for more details of the MIDES programs

benefits.

The MIDES dataset, a comprehensive collection of responses to the questionnaire applied by MIDES agents, spans a significant period from 2011 to 2023. This extensive data collection effort involved visits to 292,580 households (942,423 persons). Table 3 provides a snapshot of the dataset, revealing that around 33% of these households are based in the capital city (Montevideo), and almost all have at least one person below 18 years old. From these vulnerable households, 25% are eligible for AFAM (vulnerable), and 36% are eligible for TUS (highly vulnerable). Those who are not eligible are not considered in the analysis.

Our analysis of the credit market behavior of the vulnerable and highly vulnerable populations focuses on individuals aged 18 and above. Table **??** provides an overview of key demographic characteristics within these two vulnerable groups, AFAM and TUS. The AFAM program extends to approximately 155,000 adults, while the TUS program encompasses nearly 290,000 adults. Approximately half of the population in both groups identify as female, with an average age hovering around 36 years old. The predominant level of educational attainment across both groups is complete elementary school, with lower levels of education observed among those highly vulnerable (eligible for TUS). About 50% of vulnerable individuals are engaged in either salaried employment or self-employment, while this proportion decreases to 37% highly vulnerable. Additionally, both groups exhibit a relatively high proportion of inactive individuals.

Table 4 gives some insight into the indebtedness characteristics of this population. More than 70% of them have had a consumer debt during the analysis period. Those vulnerable (AFAM) have a higher rate of indebtedness with banks and lending firms than the most vulnerable (TUS). The percentage of people who defaulted at least once during the analysis is notably high for both groups, standing at 75.4% for the vulnerable group (AFAM) and 82.8% for the highly vulnerable group (TUS). In stark contrast, the incidence of default among non-vulnerable individuals is notably lower, standing at 44%.

Furthermore, distinct preferences emerge when comparing vulnerable and highly vulnerable individuals with non-vulnerable counterparts regarding access to loans. Vulnerable and highly vulnerable individuals seek loans from lending firms, whereas non-vulnerable individuals obtain loans from traditional banking institutions. Figure 1 shows that, for the non-vulnerable group, the proportion of people that have only bank debt or only lending firms lend has been similar, around 30%. However, for the vulnerable and highly vulnerable groups, the proportion of people with only bank debt has been 10%. In contrast, the proportion of people with only a lending firm's debt has been around 60%.

3.3 Sample: relationships between debtors and institutions

This paper focuses on the performance of vulnerable and highly vulnerable individuals in the consumer credit segment. In particular, we aim to investigate the disparity in default probabilities between these groups and the rest of the population. We built a sample at the product level (a loan) to do so. There is no identifier for each loan in the CB data, but we can identify a new relationship between a debtor and an institution. We call *spell* a continuous relationship between a debtor and institution[4]. We study the behavior of the debtor during it.

We exclude *spells* starting before March 2014 because our dataset begins in January 2014; hence, we can not determine whether the debtor-institution relationship started in January or originated from a pre-existing arrangement established in the months preceding January. Additionally, we omit *spells* starting after April 2021 because loans haven't matured sufficiently to analyze the behavior of debtors, given that the typical duration of a consumer loan is up to 24 months.

We consider that there is a default on a *spell* if the debtor was classified by the institution with a credit rating of 4 or 5 on some month during it. We build some indebtedness variables at the *spell* level; for example, if the *spell* is the first loan for the debtor, the number of institutions when the *spell* begins or if the debtor has only lending firms debt when the *spell* begins.

Table 5 summarizes descriptive statistics for this sample. Over the period from January 2014 to December 2022, a total of 5,397,208 spells were recorded. Among these, approximately 5% correspond to individuals in the vulnerable group, while approximately 7% pertain to individuals in the highly vulnerable group. The average loan amount differs significantly between non-vulnerable and vulnerable individuals, with non-vulnerable individuals obtaining an average loan amount of 30,242 pesos, compared to approximately 20,000 pesos for the vulnerable group and 17,000 pesos for the highly vulnerable group[5]. Similarly, the contingency for non-vulnerable individuals is higher, averaging around 50,000 pesos, while it ranges between 12,193 pesos on average for the vulnerable group and 7,882 for the highly

[4]If we do find a gap in this relationship that lasts less than three months, we consider it a misreporting mistake and take the whole relationship as one spell.

[5]At constant prices of December 2021, 1 US dollar = 44.3 pesos

vulnerable group. Although there are no differences in the number of institutions in which a person has debt when taking the loan in the spell considered, significant variations exist across groups in the proportion of defaulted loans and the institutions they operate with. While the proportion of defaulted loans for non-vulnerable individuals is 20.7%, individuals in the vulnerable group defaulted in 46% of the spells, and individuals in the highly vulnerable group defaulted in 59% of the spells. Significantly, in the non-vulnerable group, most spells consist of bank debt. In contrast, for vulnerable and highly vulnerable individuals, bank debt is much less observed, and most of the loans are issued by lending companies (see Figure 1).

4 Empirical strategy

This section explores alternative models to assess the disparity in default probabilities between vulnerable groups and the rest of the population. We have two primary objectives: (i) to identify the source of the observed disparity - whether it stems from higher interest rates charged to these groups or their inherent characteristics, and (ii) to investigate how specific loan features influence default probability.

We employ Linear Probability Models (LPMs) with three specifications to achieve these goals. The first model excludes interest rates, allowing us to isolate the impact of other factors. The second model incorporates a third-degree polynomial of a proxy for interest rates, enabling a more intricate analysis of their relationship with default probabilities. Finally, the third model introduces firm and time fixed effects, providing greater flexibility in capturing the dynamic interplay between interest rates and default probability across different firms and over time. Due to the inclusion of high-dimensional fixed effects, the LPM remains a suitable choice compared to non-linear models like Logits, which might face computational challenges.

We define a binary variable G indicating membership in a vulnerable group. Its coefficient will estimate the default risk gap between this group and the rest of the population. We model the probability of default as a function of G alongside three additional groups of variables: i) individual demographic characteristics (X), ii) observable characteristics of the individual's indebtedness (W), and iii) the loan interest rate (R). Therefore, the conditional probability of default is represented as $P(G, X, W; R) = Pr(Default \mid G, X, W; R)$. However, since demographic data (X) are only available for people applying for social programs, when the sample includes non-vulnerable people we will be able to estimate models only for

$P(G, W; R)$.

To define the binary indicator G, we classify people into three groups using MIDES data. As explained in section 3, the MIDES panel dataset includes all individuals in households who have ever applied for cash transfer subsidies. Based on an exhaustive set of household characteristics, a deprivation index ICC is computed to determine whether or not the household is eligible for social programs.

The first group we define as vulnerable is those households with children eligible for the conditional cash transfer AFAM but not for the TUS program. The second group is the most vulnerable, i.e., households eligible for the TUS program. This group encompasses the most vulnerable individuals, those in poverty and extreme poverty.

The analysis includes a third group, labeled "non-vulnerable", for individuals not in the MIDES dataset. We assume their income and economic means place them significantly above poverty. Separate models are estimated for each group within the target population $G = AFAM, TUS$, and compared directly with the "non-vulnerable" group.

Thus, our LPM is given by:

$$Pr(Default_{ij} \mid G_i, W_{ij}, R_{ij}) = \alpha + \gamma G_i + W'_{ij}\beta + f(R_{ij}, \pi) + \epsilon_{ij} \tag{1}$$

where i is individual, j is loan, and G is a binary indicator for vulnerability (AFAM) or high vulnerability (TUS).

Interest rates are crucial in our estimations, determining the impact on individual default probability through the function $f(R_{ij}, \pi)$. Unfortunately, individual loan-level interest rates $R_{i,j}$ are absent from the dataset. To overcome this limitation, we employ two alternative approaches. The first one utilizes a proxy for the individual interest rate, incorporating the average and cap rates for consumer loans the financial system offers. Figure 2 illustrates the significant variability of these indicators over the study period. It is worth mentioning that when proxies for interest rates are included, clustered standard errors are computed.[6]

The second approach, considered more flexible, employs fixed effects. By interacting firm dummies with monthly time dummies, we can isolate the influence of interest rates while

[6] Clusters are defined to reflect the aggregation level of the proxy for the interest rate, which is given by the type of the lender (banks or lending firms) interacted with time dummies. The number of clusters is 172. This option potentially overestimates the true standard errors; however, all coefficients are statistically significant at the 1 percent level.

accounting for potential differences across lending firms and periods. This strategy is appropriate if each lender charges the same interest rate to all borrowers in a given month. As far as we know, this assumption is not overly restrictive in our context. Firstly, in the consumer segment of the Uruguayan financial system, credit ratings influence lending decisions but not individual interest rates. Secondly, evidence suggests that Uruguayan banks and credit companies adjust interest rates monthly, usually at the beginning of each month. In particular, the Central Bank updates usury limits on a monthly frequency.

The set of observable characteristics of the individual's indebtedness (W) encompasses seven variables. The first two are binary variables indicating whether all loans were obtained from banks or lending firms. The remaining covariates include income proxy (contingencies), a binary variable indicating if the individual has at least one credit card, binary variables indicating if the loan is the first for the individual, and if an institution issued the loan without a prior relationship with the borrower and the total number of firms the individual has loans with.

Finally, the group of socio-demographic characteristics of the vulnerable people (X) includes sex, age, city of residence, highest level of education, and occupation[7]. We represent the level of education with a set of dummies: completed elementary school, incomplete or complete secondary school, and the omitted group, incomplete elementary school. Similarly, occupation is represented through a set of dummies indicating employment status: unemployed, inactive, self-employed, pensioner, and salaried employed are the omitted groups.

We estimate two alternative specifications: one in which the covariates enter linearly and another, more flexible, selected using a CV-Lasso algorithm.

5 Estimating default risk gaps and the influence of interest rates

While access to financial services has become more widespread, empirical evidence demonstrates a persistent disparity in access to loan options with reasonable interest rates for individuals from lower socioeconomic backgrounds. In addition, the literature extensively addresses the phenomenon that disadvantaged minorities and people with lower socioeconomic status are more likely to default on consumer loans. This discrepancy, termed "risk

[7]Socioeconomic data comes from the MIDES dataset. Hence, we do not have socioeconomic data for non-vulnerable people.

gap" herein, often puts them at higher interest rates than wealthier borrowers. This inequity raises concerns about fairness and equitable access to financial services and its potential detrimental effects on economic development and social mobility. By analyzing empirical data, this section contributes to a deeper understanding of the factors perpetuating this risk gap and its impact on borrowing decisions and default risk. This knowledge can be crucial for informing policy interventions to promote financial inclusion, reduce economic inequality, and ultimately contribute to a more equitable financial landscape.

5.1 Quantifying the risk gap: vulnerable vs. non-vulnerable

We estimate the default risk gap according to the vulnerability for two groups: the vulnerable, defined as those who receive the AFAM social benefit from MIDES, and the highly vulnerable, who receive the TUS benefits.

Table 6 presents a set of results for the vulnerable group; thus, the estimated sample includes people from this group and individuals classified as non-vulnerable. The first three columns correspond to LPM that do not include any measure of interest rates; that is, $f(R_{ij}, \pi)$ of equation (1) is omitted. All models include a squared time trend and a constant, except for the fixed effects specifications. In column (1), the coefficient associated with vulnerability measures the unconditional risk gap between the poorest and the non-vulnerable is as high as 25.5 percentage points (p.p.). In turn, columns (2) and (3) estimate this parameter by adding controls for the selected characteristics of the individual's indebtedness, by linearly entering those covariates (column 2) and using a chosen flexible specification by a CV-Lasso algorithm (column 3). Incorporating explanatory variables significantly boosts the model's performance, with substantial improvements across metrics like R^2, AUPRC, AUROC, and the Brier score.

Notably, the risk gap substantially falls to 17.4 and 17.5 p.p., respectively, indicating that debt characteristics can explain eight p.p. of the gap. These characteristics capture access to better-quality financial products and, therefore, show us that the poor access credit with worse conditions.

The subsequent models incorporate a proxy for interest rates, as shown in columns (4) to (6).[8] Comparing columns (1) and (4) reveals that higher interest rates faced by vulnerable individuals explain four p.p. of the initial risk gap. Interestingly, upon controlling for factors

[8]The subsection 5.5 discusses in depth the results on the relationship between interest rates and default probabilities.

related to indebtedness, introducing the interest rate as a variable does not significantly alter the previously estimated risk gap.

Our final approach utilizes a flexible model with high-dimensional fixed effects to control for the influence of the interest rate. These fixed effects are constructed from the interaction of lender and time binary indicators (37 firms x 85 months = 3,145 variables). As we have argued in the previous section, this strategy effectively isolates the impact of the interest rate from that of group membership.

Column (7) in Table 6 indicates that the risk gap is 17 p.p., a significant reduction of eight p.p. compared to the unconditional gap of column (1). This finding suggests that interest rate disparities explain roughly 30% of the initial gap between vulnerable and non-vulnerable individuals. Additionally, incorporating individual indebtedness characteristics into the fixed effects model reduces the estimated risk gap by nearly 12 p.p. to 14 p.p. and 13 p.p. when these characteristics are entered linearly using the selected CV-Lasso specification.

When examining the highly vulnerable group, the findings hold, but the disparities are notably more significant. As shown in Table 7, the gap in default likelihood between the highly vulnerable group and the non-vulnerable group stands at 38.9 percentage points (p.p.), which is higher than the gap observed for the previously mentioned vulnerable group (AFAM group). This disparity shrinks further as the model incorporates characteristics of individual indebtedness and the influence of interest rates. Moreover, the importance of interest rate explaining the estimation of the default gap is greater (40%) for highly vulnerable than for vulnerable individuals.

In summary, our results document the existence of a non-negligible risk gap between vulnerable and non-vulnerable individuals. They also highlight the importance of the interest rate when analyzing the default of the vulnerable. After controlling for interest rate and loan characteristics, there is an additional default risk of 14.5 p.p. and 22.4 p.p. for the vulnerable and highly vulnerable, respectively.

5.2 Quantifying the risk gap: banks vs. lending firms

Individuals borrowing only from lending firms exhibit a significantly higher probability of default compared to those borrowing from both banks and lending firms or solely from banks (columns (2) and (3) in Tables 6 and 7) without controlling for interest rates. Lending firms' typically higher interest rates can partially explain this disparity. Specifically, individuals with only lending firm debt have an 18 p.p. higher default likelihood than those with loans from banks and lending firms. Conversely, those with only bank debt experience a five p.p. lower default probability. This 23 p.p. difference between the two groups underscores the potential influence of loan source on default risk by type of institution, where lending firms usually charge higher interest rates than banks.

However, accounting for interest rates significantly reduces the impact of loan sources. When a proxy for interest rates is included, the "Only banks" coefficient becomes non-significant. The coefficient of "Only lending firms" also decreases eight percentage points (p.p.), falling to 10 p.p..

In the fixed effect model, the effect of "Only bank" is statistically significant but is very small in magnitude. The coefficient of "Only lending firms" decreases by ten p.p., falling to eight p.p. Therefore, including the interest rate lowers the effect of "Only banks" in magnitude and reduces more than 50% the effect of "Only lending firms".

The higher default probability for individuals relying solely on lending firms can be explained by their focus on higher-risk borrowers. However, in this analysis, we control for group affiliation, mitigating this potential effect.

5.3 Quantifying the risk gap: other indebtedness characteristics

This subsection contributes to a deeper understanding of factors influencing individuals' default probability by exploring characteristics beyond vulnerability and interest rates, including loan source and indebtedness features.

The *Contingency* variable is a proxy for individual income, primarily based on credit card limits. This variable exhibits a significant negative correlation with default probability, indicating that lower-income individuals are more likely to default. This effect remains consistent across all models with a magnitude of approximately one p.p. (Tables 6 and 7). Notably, these findings are similar for vulnerable and highly vulnerable and persist regardless of the chosen model specification (linear vs. CV-Lasso).

Including a binary variable indicating credit card ownership presents a multifaceted interpretation. Credit card ownership signifies improved access to financial products. According to the 2017 Survey of Household Finances in Uruguay, 40% of households lacked credit cards, and ownership strongly correlates with economic status (Olivieri et al., 2022). However, our data reveals a concerning portion of individuals in default with small, credit card-related debts.

Because the firms have low incentives to recover these small debts, potentially leaving individuals under-informed and vulnerable due to limited financial literacy, which is often prevalent among disadvantaged populations. A credit card increases the probability of default by between three and six p.p. for vulnerable and non-vulnerable regardless of the model specification.

Finally, three additional variables with statistically significant positive coefficients are included: i) indicating whether the loan is the person's first one in the observed period, ii) capturing whether the loan was issued by a firm with which the individual had no prior relationship, and iii) representing the total number of firms the individual has a loan. Again, the effect of these variables on the probability of default is quite similar, regardless of whether the sample includes vulnerable or highly vulnerable individuals.

5.4 Quantifying the risk gap among vulnerable people

We further analyze the risk gap between highly vulnerable and vulnerable individuals. In addition to the loan-level variables discussed above, we incorporate the individual's socioeconomic characteristics, such as age, gender, area of residence, schooling, employment status, and, in particular, the deprivation index (ICC), which is a continuous measure of household vulnerability.

Considering models that only include a binary variable (Highly vulnerable) to control vulnerability, the risk gap between the highly vulnerable and the vulnerable people is 13.5 p.p. without controls, and 9 p.p. when fixed effects that control interest rates are included. An interesting result is that when ICC (and other characteristics) are added, the binary variable capturing group membership becomes not significantly different than zero.[9] However, the difference in the probability of default for an average highly vulnerable person and a vulnerable person remains close to 10 p.p.[10] It is worth noting that the former estimate uses a binary indicator for vulnerability classification, underscoring the consistency of the risk gap even when employing different measurement approaches. Furthermore, this result suggests the presence of risk heterogeneity among individuals within each group.

[9]Except in the CV-Lasso fixed effects model, where it is 1 p.p.

[10]This risk gap is calculated by subtracting the average predicted probability for highly vulnerable individuals (considering their average deficiency index) from the average predicted probability for vulnerable individuals (considering their average deficiency index).

Regarding other covariates, as expected, the probability of default decreases with age and education. In all models, the gender gap in default is negative and significant, but narrow, ranging from -0.5 p.p. to -1.5 p.p. (see Table 8). In addition, default rates are higher in the capital than in the rest of the country. It is worth noting that including controls for the interest rate does not substantially modify the effect of the exploratory variables on the probability of default.

5.5 Default risk and the interest rate

Corporate finance has heavily scrutinized the trade-off between expected returns and risk, but a similar level of analysis needs to be included in the domain of personal debt. This is surprising because a comprehensive model to determine consumer loan demand is noticeably absent. Recent groundbreaking work by Chatterjee et al. (2023) and others signal the potential for exciting advancements in this field in the coming years. Conversely, recent empirical research in the consumer credit market has focused on analyzing factors determining the probability of default. However, interest rates have yet to be a central focus of these studies.

This subsection presents empirical evidence on the relationship between default risk and interest rates, acknowledging limitations in the analysis. Figure 3 plots the estimated default probabilities for each interest rate for the three groups: non-vulnerable, vulnerable, and highly vulnerable. To compute these estimated default probabilities, we set the OLS model's covariates with controls at their average values. Note that we are using a third-order interest rate polynomial for the estimations. As expected, the probability of default increases over the relevant range of interest rates. It is important to note that our estimation does not capture the behavior of interest rates at low values since the interest rate proxy we use does not take values below 41%. This underscores the caution needed when extrapolating our results to low-interest rate scenarios since the default probability function is decreasing as a function of the interest rate at low values of this variable.

Our estimates show considerable variation in the probability of default across interest rates for each group analyzed. However, the variation between groups exceeds the variability across interest rates, especially in the case of vulnerable groups. In addition, introducing loan-level controls leads to a consistent decrease in the probability of default at each interest rate and a flattening of all curves.

While existing research analyzes consumer loan default probabilities, it often ignores interest rates. We present evidence of a positive correlation between default risk and interest rates, with the effect being more pronounced for vulnerable populations. However, our results have shortcomings because we do not observe interest rates but use a proxy for them. The study calls for greater focus on the relationship between interest rates and personal debt, similar to the analysis done in corporate finance.

6 Theoretical insights

This section lays out theoretical insights to build a foundation for interpreting our empirical results. We consider a simplified scenario: a credit market with two borrowers, two lenders, and a regulator.

Inspired by Fuster et al. (2022), we represent credit demand through a reduced form, directly linking the interest rate to the probability of default, simplifying the model by assuming borrowers only decide whether to default, neglecting any potential elasticity of demand to the interest rate.

In this theoretical framework, we posit that borrowers are differentiated solely based on their intertemporal preferences, which are reflected by their subjective discount rates (impatience) and risk aversion. This aligns with the established theoretical foundation in consumer lending, where individual heterogeneity in preferences is recognized as a crucial factor driving borrowing behavior.

Specifically, Borrower 2 is inherently riskier than Borrower 1, meaning their default probability is higher at any given interest rate. To simplify, the risk gap between these borrowers remains constant across different interest rates. Under these assumptions, the probability of default can be expressed as:

$$Pr(\text{Default} \mid G, R) = Pr(\text{Default} \mid R) + \gamma G \tag{2}$$

where $G = 1, 2$ indicates the lowest and highest risky borrowers, respectively, and R is the gross interest rate. The parameter γ is the default risk gap between Borrower 1 and Borrower 2.

Lending firms issue consumer loans to maximize the excess return over the risk-free asset. Profits are given by the following standard equation (see Fuster et al. (2022)):

$$\pi = [R(1 - h\,\mathrm{Pr}) - \rho](1 - c\,\mathrm{Pr}) \tag{3}$$

where R is $(1 + r)$, r is the interest rate, $\mathrm{Pr} = Pr(\text{Default} \mid G, R)$ is the probability of default, c captures costs (mainly provisions and capital requirements). The parameter h is the loss rate of defaulted loans, and ρ is the gross risk-free return rate. Parameters c and h would depend on Pr and thus on R; however, we consider the simplest case where they are constant.

The regulator plays a crucial role by setting the parameter c and establishing limits on usury rates. We assume that both types of firms face the same c and cap rates. The parameter h represents a firm's ability to recover defaulted loans, introducing heterogeneity among them. We assume that Lender 1 is less efficient than Lender 2 at recovering defaulted loans. Our understanding of the Uruguayan credit market suggests that banks generally tend to be less efficient than lending firms in this particular aspect. The loss rate of defaulted loans, also known as loss-given default, represents the percentage of the original loan amount a lender expects to lose when a borrower defaults.

Our analysis starts with a benchmark model designed explicitly for the parameter values of both equations, capturing the essential features of the consumer segment in Uruguay's credit market. Those parameters are:

Regulator:	Limit for usury 130 percent; $c = 0.4$
Lenders:	Type 1: $h = 0.95$
	Type 2: $h = 0.40$
Borrowers:	Type 1: $\mathrm{Pr}\,1 = (1 + exp(2 - 0.025 * r))^{-1} * 0.9 - 0.08$
	Type 2: $\mathrm{Pr}\,2 = \mathrm{Pr}\,1 + 0.2$

In Figure 5, we plot default probabilities using the nonlinear function described above for $Pr(\text{Default} \mid R)$, that is assuming a fixed risk gap γ of 20 percentage points. Figure 4 further illustrates the excess return profiles of lenders 1 and 2, respectively, when facing borrowers 1 or 2.

Our benchmark analysis highlights several key findings. For Lender 1, interest rates maximizing their excess return are internal solutions (below the regulatory cap) and lead to higher profits for Borrower 1 than Borrower 2. For lender 2, the optimal rates for maximizing their excess return are corner solutions (at the cap rates) for both borrowers. While Borrower 1 remains generally more profitable at any rate, in this specific scenario, Lender 2 earns a higher excess return by lending to Borrower 2 at the cap rate compared to lending to Borrower 1 at the lower rate offered by Lender 1.

In a non-competitive scenario, each lender with $1 decides whether to lend to Borrower 1 or 2, aiming to maximize profit through their chosen interest rate. In this setting, equilibrium occurs when both lenders lend to their respective borrowers (Lender 1 to Borrower 1 and Lender 2 to Borrower 2). This equilibrium results in the riskier borrower receiving the loan at a significantly higher interest rate. Notably, this equilibrium persists even if, in a potential clash between lenders, profits from lending to Borrower 1 at a lower rate outweigh those from lending to Borrower 2 at the cap rate. This scenario resonates with the unique characteristics of the Uruguayan credit market, which features a limited number of private banks that also own lending firms. In particular, data collected from bank and lending company websites since last July reveals distinct lending practices. Over this period, banks typically charge interest rates ranging from 30% to 85%, while lending companies operate within a broader range, extending from 87% to 150%—remarkably close to the usury limits established by the Central Bank.

Our empirical findings demonstrably align with the theoretical predictions outlined earlier. Figure 6 depicts the excess returns for both lenders when utilizing the estimated default probabilities while maintaining all other parameters as previously stated. Interestingly, the optimal interest rates for both lenders differ. Lender 1 achieves its highest profit (excess return), significantly exceeding the observed average market rate (90% versus 58%) but still below the cap rate. Notably, this optimal rate falls within a range of possible values, as Figure 6 indicates. Conversely, Lender 2 experiences its highest profit only at the maximum allowed rate (corner solution). It is worth noting that, regardless of the chosen interest rate, lending to borrowers deemed highly vulnerable (Borrower 2) by Lender 1 consistently results in negative excess returns.

Figure 7 sheds light on the two main factors contributing to the unconditional risk gap, the difference in default rates between different borrower groups regardless of loan details. These two factors are i) higher inherent risk, potentially due to lower income and limited resources and information, and ii) disparities in interest rates; vulnerable groups are often subject to significantly higher interest rates than others, further exacerbating their risk of default. Based on estimated functions, our research enables us to analyze hypothetical scenarios (counterfactual analysis). This analysis reveals that both factors contribute equally (50/50) to the overall risk gap for vulnerable individuals. However, for the highly vulnerable population, the impact of inherent risk factors plays a slightly more significant role, contributing 60% to the risk gap compared to 40% stemming from higher interest rates.

7 Conclusion

This study explored the risk gap in default probabilities between vulnerable populations and the rest of the population in the context of consumer loans. We employed Linear Probability Models (LPMs) with various specifications to understand the source of this disparity and investigate how specific loan features influence default probability. We combined data at the national identifier level of loans with receipt of social benefits to analyze default, vulnerability, and the role of interest rate.

Our findings reveal that vulnerable individuals exhibit a significantly higher default risk than non-vulnerable populations. Notably, this gap diminishes when considering characteristics of the individual's indebtedness and the influence of interest rates. We provided empirical evidence that interest rates are crucial, explaining approximately 30 percent of the default gap for vulnerable individuals and 40 percent for highly vulnerable individuals. In addition, we conducted a counterfactual exercise that yielded similar figures (40 percent and 50 percent, respectively).

Loan source also matters, with borrowers relying solely on lending firms exhibiting a higher likelihood of default than those using banks. However, this effect is partially mitigated when accounting for interest rates. Moreover, other individual indebtedness characteristics, such as income proxy, credit card ownership, and loan characteristics, influence default probability.

While this study provides valuable insights, it is crucial to acknowledge limitations. The estimated relationship between default risk and interest rates might not capture the behavior for rates below the observed range. Additionally, due to data limitations, the study utilizes a proxy for the individual's interest rate.

These findings highlight the importance of considering individual characteristics and loan features when assessing default risk. They also emphasize the role of interest rates in perpetuating the risk gap for vulnerable populations. Addressing these issues could involve promoting financial inclusion initiatives, implementing regulations for fair interest rate practices, and encouraging financial literacy programs. Nevertheless, the feasibility of a market solution for the highly vulnerable group should be thoroughly studied. Lump-sum transfers involving fixed monthly amounts may not adequately respond to seasonal needs or unforeseen shocks. One possible approach is to provide this population group with affordable access to credit and cash transfers.

Future research directions include exploring alternative methodologies to estimate the impact of interest rates on default risk with more granular data, investigating the causal mechanisms behind the observed relationship between loan source and default probability, and analyzing the effectiveness of different policy interventions to reduce the risk gap for vulnerable populations.

By understanding the factors influencing default risk for vulnerable individuals and addressing the limitations of this study, future research can inform policy and practices to promote financial inclusion and reduce inequalities in the consumer credit market.

References

Arnabal, R., S. Taroco, C. Dassatti, V. Landaberry, and J. Ponce (2023). Endeudamiento de las personas físicas en Uruguay. Documento de trabajo del Banco Central del Uruguay 007-2023.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics 143*(1), 30–56.

Becher, S. I., Y. Feldman, and O. Lobel (2023). *Poor Consumer(s) Law: The Case of High-Cost Credit and Payday Loans*, pp. 384–414. Cambridge Law Handbooks. Cambridge University Press.

Bhuta, N., P. M. Skiba, and J. Tobacman (2015). Payday loan choices and consequences. *Journal of Money, Credit and Banking 47*(2/3), 223–259.

Bhutta, N., A. Hizmo, and D. Ringo (2022). How much does racial bias affect mortgage lending? evidence from human and algorithmic credit decisions. Finance and Economics Discussion Series 2022-067. Washington: Board of Governors of the Federal Reserve System.

Brune, L., E. Chyn, and J. Kerwin (2021). Pay me later: Savings constraints and the demand for deferred payments. *American Economic Review 111*(7), 2179–2212.

Butler, A. W., E. J. Mayer, and J. P. Weston (2022). Racial Disparities in the Auto Loan Market. *The Review of Financial Studies 36*(1), 1–41.

Camara, S. (2022). TANK meets Diaz-Alejandro: Household heterogeneity, non-homothetic preferences & policy design. (2201.02916).

Castellanos, S., D. Jiménez-Hernández, A. Mahajan, and A. Seira (2020). Expanding financial access via credit cards: Evidence from Mexico. Mimeo.

Chatterjee, S., D. Corbae, K. Dempsey, and J.-V. Ríos-Rull (2023). A quantitative theory of the credit score. *Econometrica 91*(5), 1803–1840.

DINEM (2011). Informe MIDES: Evaluación y seguimiento de programas 2009-2010. tech. rep.

Dobbie, W., A. Liberman, D. Paravisini, and V. Pathania (2021). Measuring Bias in Consumer Lending. *The Review of Economic Studies 88*(6), 2799–2832.

Dobridge, C. (2016). For better and for worse? effects of access to high-cost consumer credit. Finance and Economics Discussion Series 2016.056. Washington: Board of Governors of the Federal Reserve System.

Durkin, T. A. and G. Elliehausen (2014). Consumer Lending. *Oxford University Press*.

Elliehausen, G. and E. Lawrence (2009). An analysis of consumers' use of payday loans. George Washington University, Financial Services Research Program, Monograph No. 41.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance 77*(1), 5–47.

Gardner, J. (2022). *The Future of High-Cost Credit. Rethinking Payday Lending.* Bloomsbury Publishing.

Ghent, A. C., R. Hernández-Murillo, and M. T. Owyang (2014). Differences in subprime loan pricing across races and neighborhoods. *Regional Science and Urban Economics 48*(C), 199–215.

Hanson, A., Z. Hawley, H. Martin, and B. Liu (2016). Discrimination in mortgage lending: Evidence from a correspondence experiment. *Journal of Urban Economics 92*(C), 48–65.

Kaplan, G., G. L. Violante, and J. Weidner (2014). The Wealthy Hand-to-Mouth. *Brookings Papers on Economic Activity 45*(1 (Spring), 77–153.

Lagomarsino, A. (2021). Do cash assistance programs create welfare traps? Mimeo.

Olivieri, C., G. Sanroman, and R. Quagliotti (2022). Debit and credit card holdings: effects of the Uruguayan Financial Inclusion Law. Working paper No. 04/22, dECON.

Spector, M. (2008). Taming the beast: Payday loans, regulatory efforts, and unintended consequences. DePaul Law Review, Vol. 57, No. 4, SMU Dedman School of Law Legal Studies Research Paper No. 212.

Tables and Figures

Table 1: Persons indebted with consumer credit

	#	Has bank debt (%)	Has lending firm debt (%)	Has credit card (%)	Has defaulted at least once (%)
Persons	2,596,077	80.4	55.9	80.5	49.9

Notes: This table shows some descriptives for the sample of persons in the Credit Bureau data. Default means credit ranking 4 or 5 (90 days or more delinquent).

Table 2: Vulnerable households

	#	Vulnerable (%)	Highly vulnerable (%)	Set in capital city (%)	Have minors (%)	Average members (#)
Households	292,580	25.4	37.7	33.8	81.1	3.2

Notes: This table shows some descriptives for the sample of vulnerable households. The vulnerable group includes people who are eligible for the AFAM program of cash transfers (but not for the TUS program). The highly vulnerable group includes people who are eligible for the TUS program of cash transfers.

Table 3: Vulnerable persons

	Vulnerable (Elegible AFAM)	*Highly vulnerable (Elegible TUS)*
# Persons	155,351	289,266
Females (%)	59.5	54.2
Mean age	37.7	35.5
Higher level of education		
Incomplete elementary school (%)	11.3	22.8
Complete elementary school (%)	48.9	55.7
Incomplete secondary school (%)	25.3	11.4
Complete secondary school (%)	14.5	10.1
Occupation		
Unemployed (%)	14.6	16.0
Salaried employee (%)	24.9	14.0
Self employed (%)	23.5	22.9
Pensioner (%)	5.7	5.2
Inactive (%)	27.3	35.1

Notes: This table shows some descriptives for the sample of vulnerable persons. The vulnerable group includes people who are eligible for the AFAM program of cash transfers (but not for the TUS program). The highly vulnerable group includes people who are eligible for the TUS program of cash transfers.

Table 4: Indebtedness of vulnerable persons

	Vulnerable (AFAM)	*Highly vulnerable (TUS)*
Has had consumer debt (%)	77.5	70.9
Has had credit card (%)	62.1	52.6
Has had bank debt (%)	54.4	42.3
Has had lending firm debt (%)	54.5	51.5
Has defaulted at least once (%)	75.4	82.8

Notes: This table shows some indebtedness descriptives for vulnerable persons. The vulnerable group includes people eligible for the AFAM program of cash transfers (but not for the TUS program). The highly vulnerable group includes people eligible for the TUS cash transfer program.

Table 5: Borrower-institution spells

	Vulnerable (AFAM)	*Highly vulnerable (TUS)*	*Non-vulnerable*
# Persons	263,739	362,834	4,770,635
Loan amount (avg)	20,278	16,963	30,242
Contingency when taking the loan (avg)	12,193	7,882	50,428
# institutions when taking the debt (avg)	3	2	3
Defaulted loan (%)	45.9	59.3	20.7
Lending firm loan (%)	61.6	66.5	44.7
Bank loan (%)	38.4	33.5	55.3

Notes: This table shows some descriptives for the spells sample. A spell is a continuous relationship between a debtor and an institution. The vulnerable group includes people eligible for the AFAM program of cash transfers (but not for the TUS program). The highly vulnerable group includes people eligible for the TUS program. Non-vulnerable includes all spells not on the MIDES data. All monetary figures are in December 2021 pesos, 1 US dollar = 44.3 pesos.

Table 6: Default and vulnerability: Linear Probability Models. Dependent variable: Default

	Without interest rate			With interest rate			With fixed effects		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
OLS	OLS	CV-Lasso	OLS	OLS	CV-Lasso	OLS	OLS	CV-Lasso	
Vulnerable (AFAM)	0.255***	0.174***	0.175***	0.219***	0.174***	0.171***	0.169***	0.140***	0.131***
	(0.001)	(0.001)	(0.001)	(0.006)	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)
Only banks		-0.047***	-0.046***		0.000	-0.000		0.025***	0.026***
		(0.000)	(0.001)		(0.003)	(0.003)		(0.000)	(0.001)
Only lending firms		0.178***	0.175***		0.099***	0.099***		0.014***	0.012***
		(0.001)	(0.001)		(0.004)	(0.004)		(0.001)	(0.001)
Contingency (in logs)		-0.014***	-0.014***		-0.013***	-0.013***		-0.009***	-0.009***
		(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)
Credit card		0.084***	0.084***		0.061***	0.061***		0.033***	0.033***
		(0.000)	(0.000)		(0.005)	(0.005)		(0.000)	(0.000)
First loan		0.027***	0.026***		0.026***	0.025***		0.041***	0.040***
		(0.001)	(0.001)		(0.004)	(0.004)		(0.001)	(0.001)
New Institution		0.137***	0.137***		0.134***	0.134***		0.094***	0.095***
		(0.000)	(0.000)		(0.007)	(0.007)		(0.000)	(0.000)
# of Institutions		0.039***	0.039***		0.033***	0.033***		0.023***	0.023***
		(0.000)	(0.000)		(0.001)	(0.001)		(0.000)	(0.000)
Observations	4,993,309	4,993,309	4,993,309	4,993,309	4,993,309	4,993,309	4,993,311	4,993,311	4,993,311
Rsquare	0.0196	0.1188	0.1207	0.0819	0.1323	0.1332	0.1449	0.1742	0.1749
AUPRC	0.2783	0.4444	0.4475	0.3659	0.4588	0.4608	0.4662	0.5037	0.5049
AUROC	0.6052	0.7292	0.7298	0.6918	0.7392	0.7394	0.7410	0.7585	0.7586
Brier	0.2179	0.2028	0.2025	0.2085	0.2007	0.2006	0.1993	0.1948	0.1947

Notes: This table presents results from three sets of models. The first set excludes interest rates, the second incorporates a third-degree polynomial of a proxy for interest rates, and the third introduces firm and time-fixed effects to control interest rate heterogeneity. The first and second sets also include a squared time trend and a constant. In Columns 4 to 6, clustered standard errors are reported, with each cluster defined as the interaction of the period with the type of lender, reflecting the aggregation level of the proxy for the interest rate. CV-Lasso figures represent marginal effects and correspond to the estimation of the selected specification using the entire sample. In Column (9), CV-Lasso uses OLS residuals of the outcome and covariates on fixed-effects." ** $p < 0.05$, *** $p < 0.01$

Table 7: Default and highly vulnerability: Linear Probability Models. Dependent variable: Default

	Without interest rate			With interest rate			With fixed effects		
	OLS	OLS	CV-Lasso	OLS	OLS	CV-Lasso	OLS	OLS	CV-Lasso
Highly vulnerable (TUS)	0.389***	0.276***	0.269***	0.342***	0.278***	0.265***	0.261***	0.218***	0.197***
	(0.001)	(0.001)	(0.001)	(0.008)	(0.006)	(0.004)	(0.001)	(0.001)	(0.001)
Only banks		-0.048***	-0.046***		0.000	0.001		0.025***	0.025***
		(0.000)	(0.001)		(0.003)	(0.003)		(0.000)	(0.001)
Only lending firms		0.185***	0.177***		0.103***	0.099***		0.018***	0.014***
		(0.001)	(0.001)		(0.004)	(0.004)		(0.001)	(0.001)
Contingency (in logs)		-0.014***	-0.014***		-0.012***	-0.012***		-0.009***	-0.009***
		(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)
Credit card		0.088***	0.088***		0.064***	0.064***		0.034***	0.033***
		(0.000)	(0.000)		(0.005)	(0.004)		(0.000)	(0.000)
First loan		0.032***	0.027***		0.031***	0.026***		0.044***	0.041***
		(0.001)	(0.001)		(0.004)	(0.004)		(0.001)	(0.001)
New Institution		0.136***	0.138***		0.133***	0.133***		0.094***	0.096***
		(0.000)	(0.000)		(0.007)	(0.007)		(0.000)	(0.000)
# of Institutions		0.039***	0.039***		0.033***	0.033***		0.023***	0.022***
		(0.000)	(0.000)		(0.001)	(0.001)		(0.000)	(0.000)
Observations	5,092,040	5,092,040	5,092,040	5,092,040	5,092,040	5,092,040	5,092,042	5,092,042	5,092,042
Rsquare	0.0561	0.1533	0.1574	0.1191	0.1667	0.1698	0.1804	0.2079	0.2091
AUPRC	0.3473	0.5062	0.5134	0.4397	0.5193	0.5265	0.5292	0.5625	0.5640
AUROC	0.6012	0.7237	0.7252	0.6876	0.7333	0.7346	0.7366	0.7539	0.7544
Brier	0.2208	0.2051	0.2044	0.2107	0.2029	0.2024	0.2011	0.1965	0.1964

Notes: This table presents results from three sets of models. The first set excludes interest rates, the second incorporates a third-degree polynomial of a proxy for interest rates, and the third introduces firm and time-fixed effects to control interest rate heterogeneity. The first and second sets also include a squared time trend and a constant. In Columns 4 to 6, clustered standard errors are reported, with each cluster defined as the interaction of the period with the type of lender, reflecting the aggregation level of the proxy for the interest rate. CV-Lasso figures represent marginal effects and correspond to the estimation of the selected specification using the entire sample. In Column (9), CV-Lasso uses OLS residuals of the outcome and covariates on fixed-effects. ** $p < 0.05$, *** $p < 0.01$

Table 8: Default among vulnerable people: Linear Probability Models. Dependent variable: Default

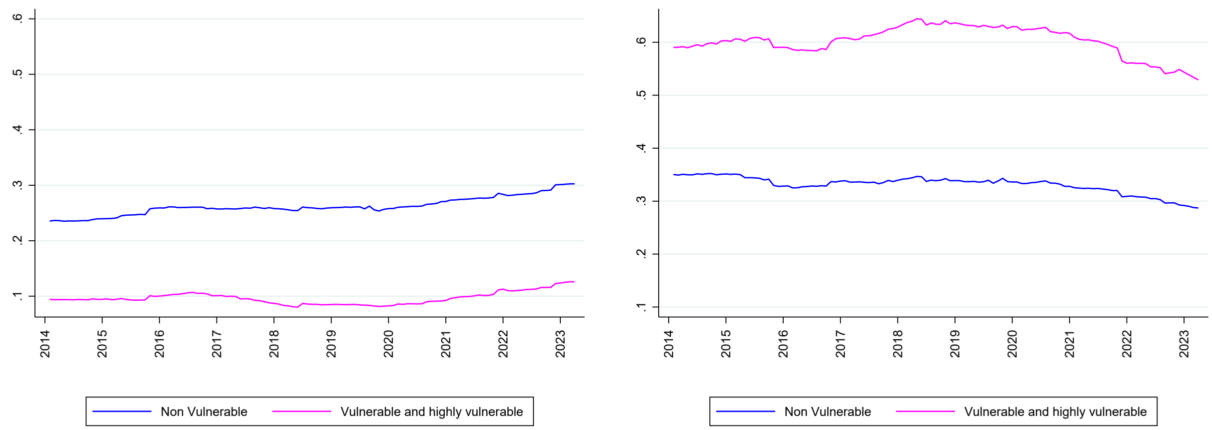
	Without interest rate			With a proxy of interest rate			With fixed effects		
	(1) OLS	(2) OLS	(3) CV-Lasso	(4) OLS	(5) OLS	(6) CV-Lasso	(7) OLS	(8) OLS	(9) CV-Lasso
Highly vulnerable (TUS)	0.135***	0.001	0.003	0.116***	0.001	0.002	0.090***	0.000	0.010***
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)
Deficiency index (ICC)		0.245***	0.240***		0.245***	0.243***		0.197***	0.163***
		(0.008)	(0.005)		(0.008)	(0.009)		(0.005)	(0.005)
Female		-0.007***	-0.008***		-0.005***	-0.006***		-0.015***	-0.015***
		(0.002)	(0.001)		(0.002)	(0.002)		(0.001)	(0.001)
Capital city		0.025***	0.025***		0.029***	0.029***		0.021***	0.021***
		(0.004)	(0.001)		(0.003)	(0.003)		(0.001)	(0.001)
Complete elementary school		0.024***	0.025***		0.024***	0.025***		0.018***	0.019***
		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)
Secondary school		-0.021***	-0.020***		-0.021***	-0.021***		-0.018***	-0.016***
		(0.002)	(0.002)		(0.002)	(0.002)		(0.002)	(0.002)
Unemployed		0.083***	0.084***		0.084***	0.084***		0.074***	0.075***
		(0.002)	(0.002)		(0.002)	(0.002)		(0.002)	(0.002)
Inactive		0.032***	0.032***		0.031***	0.031***		0.028***	0.028***
		(0.003)	(0.002)		(0.002)	(0.002)		(0.002)	(0.002)
Self employed		0.067***	0.067***		0.069***	0.068***		0.058***	0.058***
		(0.002)	(0.002)		(0.002)	(0.002)		(0.002)	(0.002)
Pensioner		0.017***	0.018***		0.007	0.009*		0.009***	0.010***
		(0.006)	(0.003)		(0.005)	(0.005)		(0.003)	(0.003)
Young		0.061***	0.063***		0.057***	0.059***		0.056***	0.057***
		(0.005)	(0.001)		(0.005)	(0.005)		(0.001)	(0.001)
Adult		-0.046***	-0.046***		-0.051***	-0.051***		-0.045***	-0.045***
		(0.002)	(0.002)		(0.002)	(0.002)		(0.001)	(0.001)
Mayor		-0.082***	-0.082***		-0.099***	-0.100***		-0.090***	-0.089***
		(0.004)	(0.003)		(0.004)	(0.004)		(0.003)	(0.003)
Observations	621,782	621,782	621,782	621,782	621,782	621,782	621,782	621,782	621,782
Rsquare	0.0215	0.1885	0.1898	0.1473	0.2122	0.2115	0.2310	0.2790	.

Notes: This table presents results from three sets of models. The first set excludes interest rates, the second incorporates a third-degree polynomial of a proxy for interest rates, and the third introduces firm and time-fixed effects to control interest rate heterogeneity. The first and second sets also include a squared time trend and a constant. In Columns 4 to 6, clustered standard errors are reported, with each cluster defined as the interaction of the period with the type of lender, reflecting the aggregation level of the proxy for the interest rate. CV-Lasso figures represent marginal effects and correspond to the estimation of the selected specification using the entire sample. In Column (9), CV-Lasso uses OLS residuals of the outcome and covariates on fixed-effects. The sample is vulnerable individuals. Omitted categories are: incomplete elementary school, salaried employees, and young adults. Indebtedness characteristics are also included but not reported in the table. ** $p < 0.05$, *** $p < 0.01$

Figure 1: Proportion of borrowers by type of institution

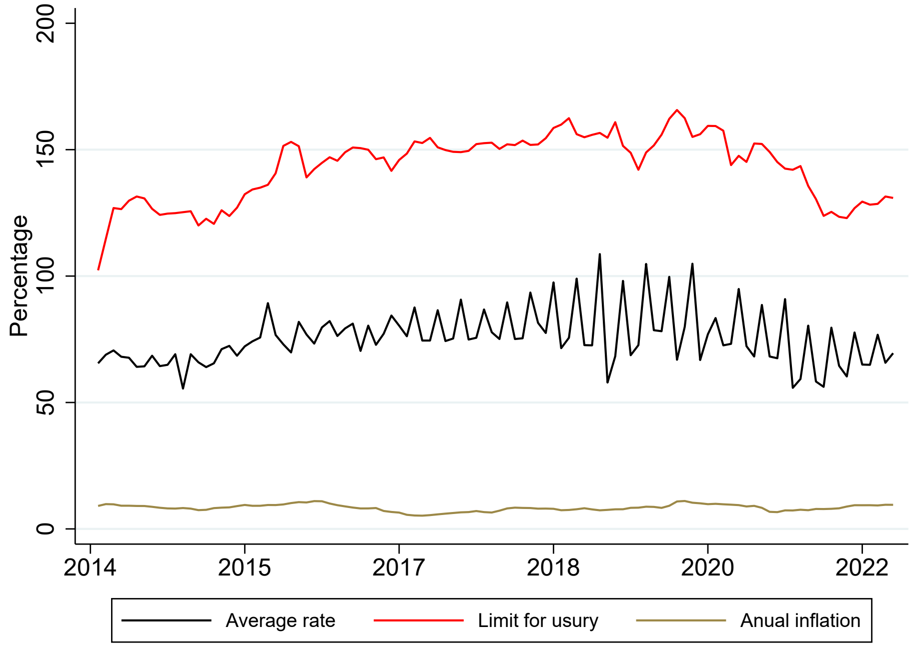
(a) Only banks

(b) Only lending firms



Notes: This figure shows the proportion of borrowers who borrow only from banks (panel a) and only from lending firms (panel b), separate into non-vulnerable and vulnerable groups. The vulnerable group includes all people eligible for the AFAM or the TUS program.

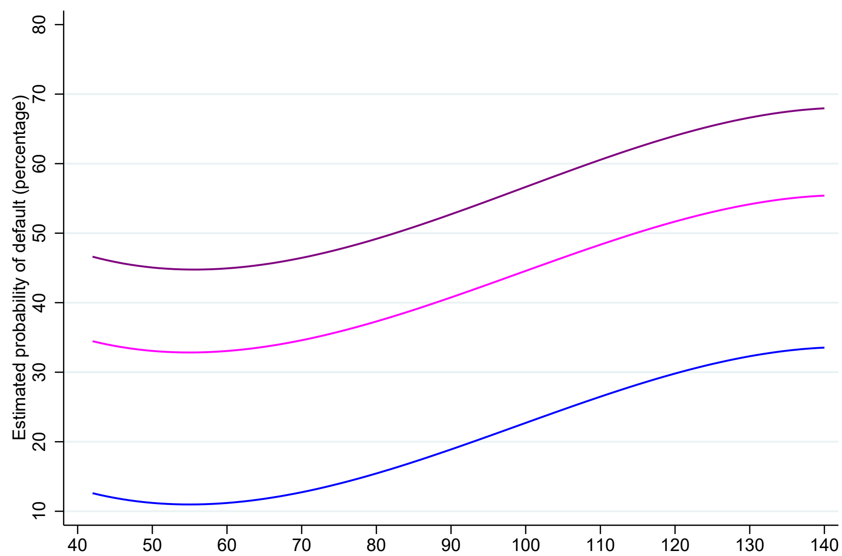
Figure 2: Interest rates



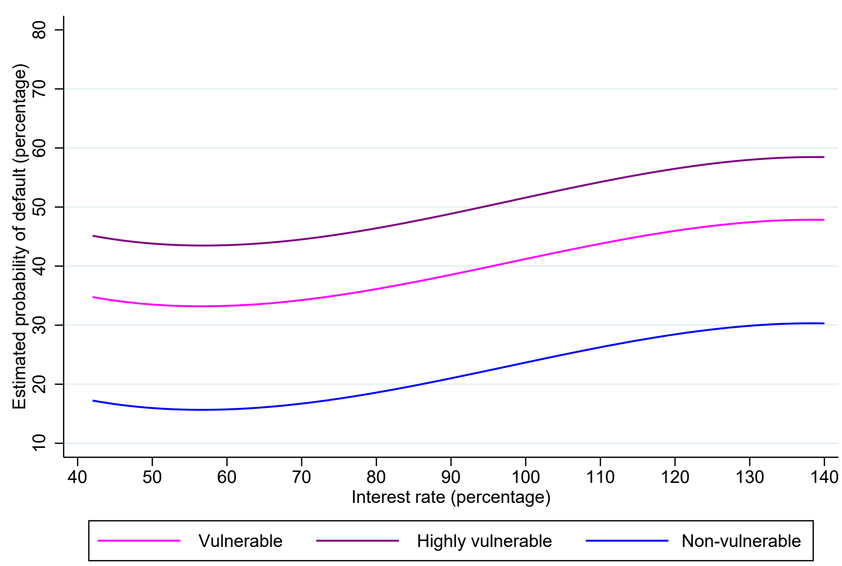
Notes: This figure shows the average interest rate, the limit for usury, and the annual inflation from 2014 to 2022 reported by the Central Bank.

Figure 3: Estimated probability of default across interest rates

(a) OLS without controls

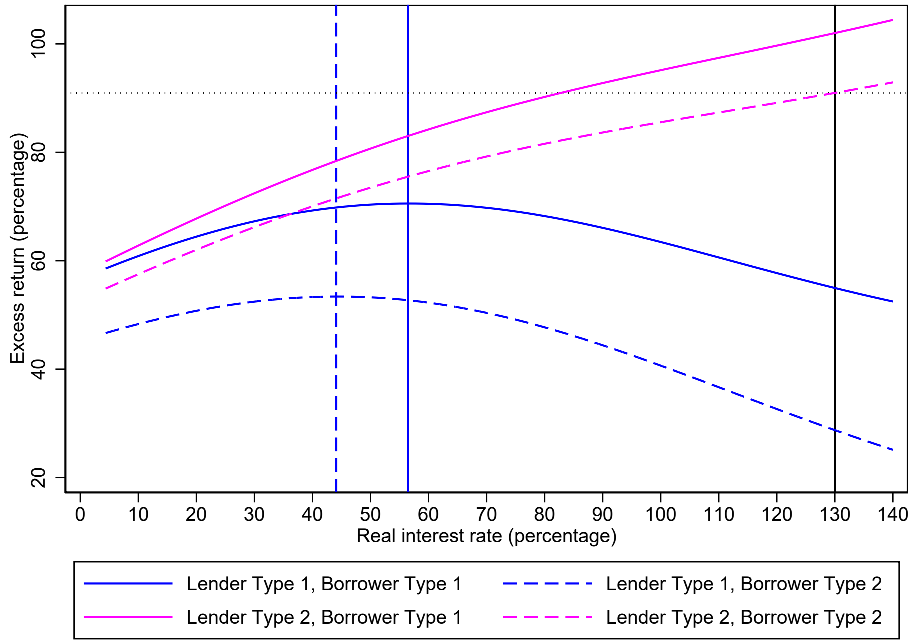


(b) OLS with controls



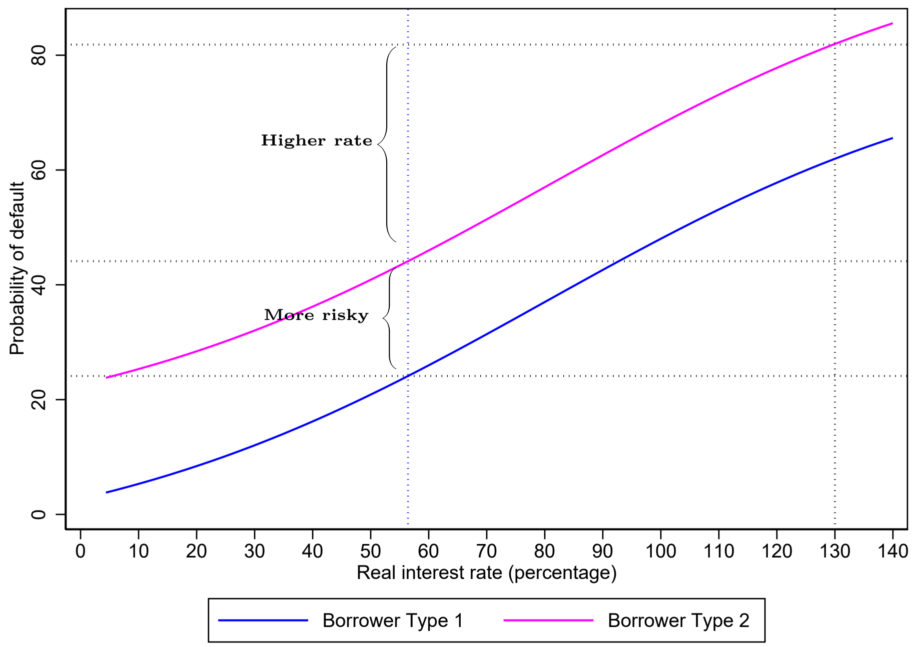
Notes: This table shows the estimated probability of default at each interest rate value on the sample, by three groups: non-vulnerable, vulnerable (eligible for AFAM cash transfer), and highly vulnerable (eligible for TUS cash transfer). The estimation uses a third-order polynomial for the interest rate. Panel (a) shows the estimation for the OLS without controls. Panel (b) shows the estimation for the OLS with controls. To calculate the estimated probability of default these controls are set at their mean.

Figure 4: Excess return over the risk-free rate,
(Fictitious parameters)



Note: This figure plots the excess return of Lenders 1 and 2 as a function of interest rate for each borrower, using fictitious data. The blue vertical lines show the optimal rate for Lender 1, the solid (dashed) line corresponds to Borrower 1 (2). The black vertical line is at the limit of usury.

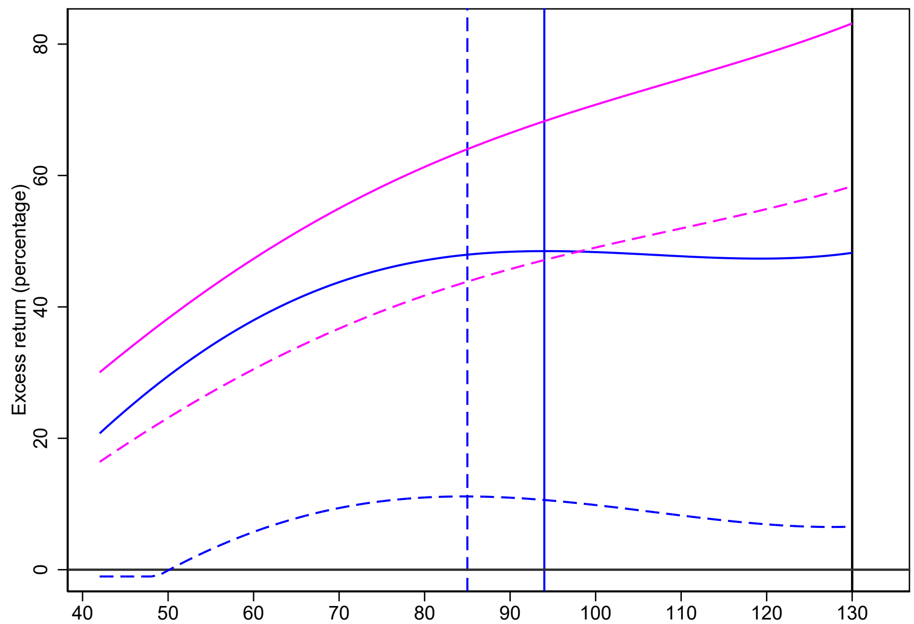
Figure 5: Probability of default and the risk gap
(Fictitious parameters)



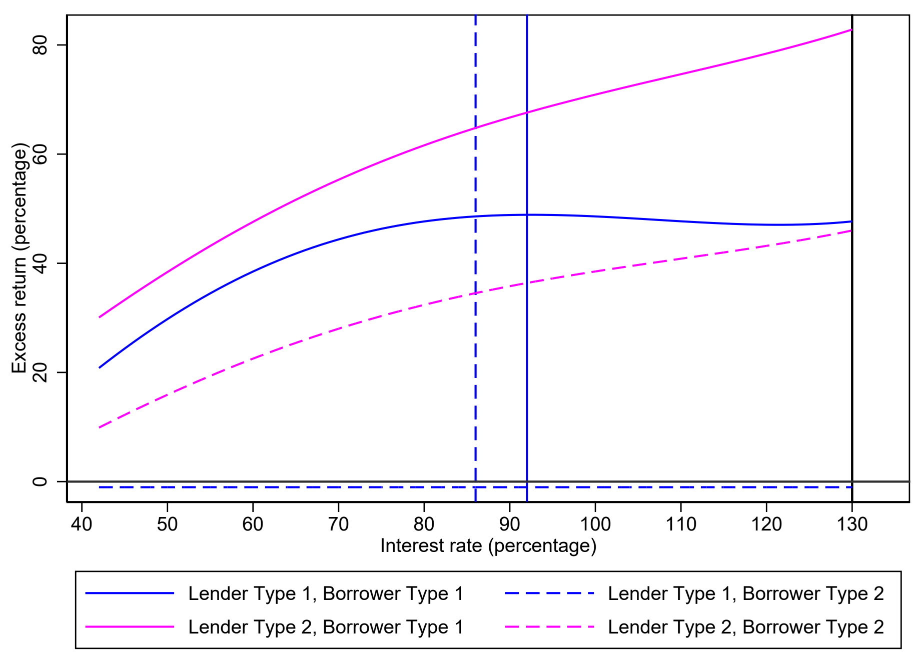
Notes: This figure plots each borrower's hypothetical probability of default, using fictitious data. The vertical lines show the optimal rate for Lender 1 when facing Borrower 1 (which resembles the observed average interest rates of banks in our sample (58) and the limit of usury (130).

Figure 6: Excess return over the risk-free rate
(Estimated probability of default)

(a) Vulnerable (AFAM) and non-vulnerable



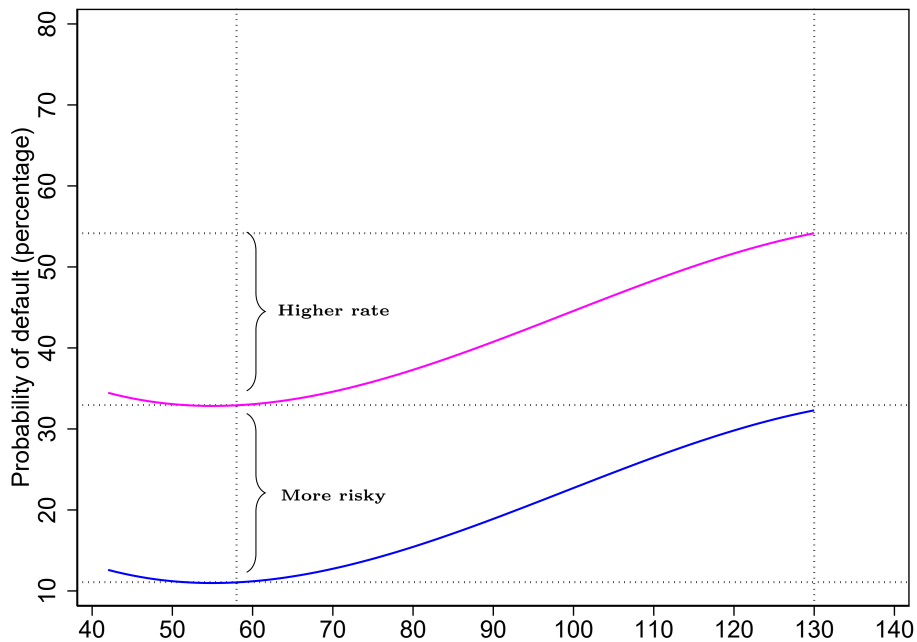
(b) Highly-vulnerable (TUS) and non-vulnerable



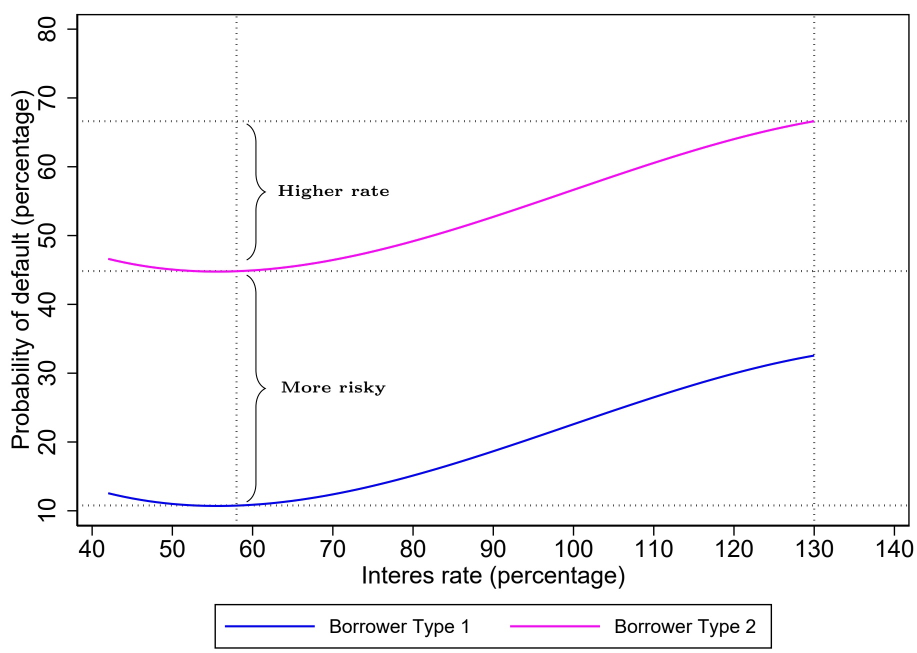
Notes: These figures plot the excess return of Lenders 1 and 2 as a function of interest rate for each borrower, using estimated parameters for the probability of default. Borrower 1 is the non-vulnerable group, and Borrower 2 is the vulnerable/highly vulnerable (AFAM/TUS) group. We use the model that includes the time trend and its square, a dummy variable for AFAM/TUS, and a third-order polynomial in the interest rate proxy, with no other covariates. The blue vertical lines show the optimal rate for Lender 1, the solid (dashed) line corresponds to Borrower 1 (2). The black vertical line is at the hypothetical limit of usury.

Figure 7: Probability of default and the risk gap
(Estimated parameters)

(a) Vulnerable (AFAM) and non-vulnerable



(b) Highly-vulnerable (TUS) and non-vulnerable



Notes: These figures plot the probability of default as a function of interest rate for each borrower, using estimated parameters for the probability of default. Borrower 1 is the non-vulnerable group, and Borrower 2 is the vulnerable/highly vulnerable (AFAM/TUS) group. We use the model that includes the time trend and its square, a dummy variable for TUS, and a third-order polynomial in the interest rate proxy, with no other covariates. The vertical dotted lines show hypothetical maximum rates set by the regulator (130%) and the average of the proxy for the interest rates of the banks in our sample (58%).