

WORKING PAPER

Behavioral and Economic Determinants of Loan Repayment Delays

AUTHOR(S)

Anvar Jamolov

Feruza Kudratova

Sardor Agzamov

DATE

November 2026

ISSUE CODE

CBU / 2026 / WP-01

www.cbu.uz

Disclaimer: All rights reserved. The views expressed in this paper are solely those of the authors and do not necessarily reflect the official position of the Central Bank of Uzbekistan. The Central Bank of Uzbekistan assumes no responsibility for the contents of the paper. Any reproduction of these materials is permitted only with the express consent of the authors.

Behavioral and Economic Determinants of Loan Repayment Delays

Anvar Jamolov* Feruza Kudratova† Sardor Agzamov‡

November 2026

Abstract

We study what drives household loan delinquency using Uzbekistan's credit registry, a panel of 233 million contract–month observations on 16 million loans held by 7.2 million borrowers between January 2023 and March 2026. We estimate logit models of monthly delinquency with bank, region, and year fixed effects, for all loans together and for each of seven loan types separately. Delinquency is lower among higher-income borrowers and on contracts with a co-borrower, falls with age over most of the working life, and is higher for mainly preferential terms and online origination and higher for larger loans, male borrowers also.

JEL classification: G21, G51, D14, G28, O16

Keywords: loan delinquency; preferential lending; household credit risk; product heterogeneity; digital lending; emerging markets

*Research and Analysis Department, Central Bank of Uzbekistan. ajamolov@cbu.uz
†Research and Analysis Department, Central Bank of Uzbekistan. f.qudratova@cbu.uz
‡Statistics and Balance of Payment Department, Central Bank of Uzbekistan. sagzamov@cbu.uz

1 Introduction

Household credit has expanded enormously in many emerging economies over the past decade. It's provided a route to finance for households that previously had no access to credit, but it has also raised issues around repayment capacity and the accumulation of credit risk. Loan delinquency (i.e. the first point at which credit risk is visible) is the key indicator of that risk. The drivers of delinquency are obviously of concern to lenders who are pricing and screening loans, to supervisors who are monitoring the level of risk in the banking system and to policymakers who are looking to expand financial inclusion without necessarily compromising stability.

Studies examining borrower repayment have looked at various borrower, loan, risk sharing and preferential or state directed credit characteristics (e.g. gender of borrower (D'Espallier et al., 2011; Agarwal et al., 2018), borrower age (Agarwal et al., 2009; Debbaut et al., 2016), loan size and interest rate (Adams et al., 2009; Kim et al., 2017), presence of guarantees and joint liability (Sharma and Zeller, 1997; Giné and Karlan, 2014; Bachas et al., 2021), preferential or state directed credit programs (Khwaja and Mian, 2005; Kanz, 2016)). However, much of the research has been conducted at the program, product or institution level (for example (Godquin, 2004) for the woman borrower and (Shahriar et al., 2020) for repayment of woman borrowers across MFI's) with findings for many of the key determinants sometimes contradictory and generally not provided for different types of credit. For example there is evidence that the repayment of borrowers of older age increases (in a monotonic fashion) (Debbaut et al., 2016; Li and Wu, 2024) while there is also evidence of a U-shaped relationship between borrower age and default (Agarwal et al., 2009). Also there is evidence that preferential credit increases default in some environments (Khwaja and Mian, 2005; Looney and Yannelis, 2015) while it is found to decrease it in others (Ding et al., 2011; Hembre et al., 2021). The reason for the differential effect of a particular variable may well be that the variable in question is interacted with the type of credit.

We test the explanation using loan level data from the credit registry of the Central Bank of the Republic of Uzbekistan on all bank loans to individuals. The market of

household credit has been witnessing rapid growth in the volume of household credit products that include a wide range of products, and a considerable part of loans is given at subsidized interest rates in the framework of government programs since 2017. The panel covers 16 million loan contracts made by 7.2 million borrowers from January 2023 to March 2026 and seven types of loans, including mortgage credit, micro credit, micro loans, auto credit, credit card, education credit, and consumer credit. We use logit models to study the monthly probability of loan delinquency on the pooled portfolio of loans and for each loan type separately, depending on borrower income, age, and gender; loan amount, interest rate, maturity, and number of loans; co-borrowing, preferential terms, and origination channel, with fixed effects of bank, region, and year. The fact that all product-level models are estimated on the basis of the same credit registry, the same time period, and the same institutional context allows us to conclude that any observed differences in loan types.

The primary finding is that the determinants of delinquency are mostly product-specific, meaning that the coefficients from pooled regression models might be biased because of heterogeneity across portfolios. In the pooled portfolio, the probability of being in delinquent status is smaller for the clients with higher income, for older clients and when the loan has co-borrower, while it is bigger for men, larger loan amount, preferential loans and loans which were applied via the Internet. Several of the mentioned effects depend on product types. For instance, the effect of the loan size on the probability of delinquent status is positive in the case of auto, micro and consumer lending, but negative in the case of mortgage and education lending. Male customers are more likely to be in delinquent status in the case of microloans and credit card lending, but less likely to be in this status in the case of auto and education credit lending. Preferential terms increase the probability of delinquency for mortgage and auto loans, but reduce it for education credit lending. Furthermore, the coefficient of the number of loans is opposite in sign to the one of the pooled regression model, as it increases the probability. The fact that there is co-borrowing is always connected with reduced delinquency where it exists, while online lending is always correlated with increased delinquency where the effect is statistically

significant.

This paper has three major contributions to make. First, it estimates similar models across seven products in a single credit market, showing that the relationship between borrower and contract characteristics and repayment outcomes are directly influenced by the type of credit. This finding is in line with the conflicting results obtained in previous studies. The gender reversal in the products suggests that the gender gap in repayments depends on the context, loan purpose and selection rather than on a fixed gender trait (Godquin, 2004; Reboul et al., 2021). The difference in the age profile of products is also a good reason for previous studies finding both monotonic and U-shaped trends: (Agarwal et al., 2009; Debbaut et al., 2016). To our knowledge, this paper is the first to compare preferential loans and market-priced loans in a systematic way, in terms of the various types of household credit products, in the same registry. The notion of preferential terms carries positive connotations in some cases and negative in others, depending on the product as borrower selection becomes the dominant criteria in some programs, and loan design in others (Khwaja and Mian, 2005; Hembre et al., 2021). Third, it builds on the existing literature on risk-sharing in group lending and government guarantees (Sharma and Zeller, 1997; Giné and Karlan, 2014; Bachas et al., 2021) to individual risk-sharing in formal bank lending. It also discovers that in an emerging market loan originations that are online are correlated to higher delinquency rates than loan originations conducted in branches. This finding is in contrast to what has been found in the United States, as the rate of default of mortgages from technology-based lenders that process applications online is either equal to or lower than that of traditional lenders (Fuster et al., 2019). In general, this paper adds to the scarce evidence on household credit risk in Central Asia, based on near-universal registry data, rather than on data from a single lender's portfolio.

2 Literature Review

This section synthesizes evidence on the characteristics included in the model, such as borrower gender and age, loan size, maturity, pricing, multiple borrowing, co-borrowing,

guaranties, and preferential lending. The influence of each characteristic differs across credit products and institutional contexts.

Gender and repayment. Substantial evidence indicates that women generally repay more reliably than men. D'Espallier et al. (2011) demonstrate that, among 350 microfinance institutions in 70 countries, a higher proportion of female clients is associated with lower portfolio risk, fewer write-offs, and fewer provisions. Agarwal et al. (2018) combine personal bankruptcy records with motor-vehicle accident data in Singapore and find that women's odds of personal bankruptcy are only 28 percent of men's, after controlling for demographic characteristics, housing type, and cultural and spatial fixed effects. Shahriar et al. (2020) identify a micro-level mechanism using a randomized trust game and loan-repayment game among 894 rural Bangladeshi subjects, showing that approximately one-third of the gender gap in repayment is attributable to differences in innate trustworthiness between men and women. However, this advantage is not universal. Godquin (2004) finds no female repayment advantage in Bangladesh once group composition, loan rationing, and social ties are controlled for.

Chen et al. (2018) demonstrate that loans funded by female investors on a Chinese peer-to-peer platform default more frequently, although this effect is limited to less-educated and lower-income investors. Fonseca et al. (2012) report that intra-couple decision-making is sensitive to the relative education level of both spouses. In a field experiment involving 17,545 FinTech borrowers, Bao and Huang (2026) observe that women are more responsive to social-pressure reminders, whereas men are more responsive to financial-incentive reminders.

Gender also affects the supply side of lending. An analysis of 375,438 loans from a Malian bank by Deiß et al. (2025) indicates that women are more likely to experience minor delays but are less likely to have severe delays exceeding 30 days. Reboul et al. (2021) find that women hold higher absolute debt, while men have higher earnings. Women borrow less relative to income and more frequently borrow to repay other loans.

Age and repayment. The evidence regarding age and repayment is mixed. Debbaut et al. (2016) find that credit card default rates decline broadly monotonically with age

in the United States, whereas Agarwal et al. (2009) document a U-shaped relationship, with optimal financial decisions occurring in middle age. Using 84,000 Chinese loan records, Li and Wu (2024) report a steady decline in default risk with age, robust to probit and propensity-score-matching specifications. In Mali, older borrowers are less likely to fall into short-term arrears but are not less likely to default seriously (Deiß et al., 2025). Livshits et al. (2010) offer a supply-side explanation for the increasing share of young defaulters: reduced lending costs, driven by computerized credit scoring, and lower bankruptcy filing costs have expanded credit access to riskier borrowers, particularly those without established credit histories.

Loan terms and delinquency. Loan size and maturity primarily influence delinquency through the repayment burden imposed on borrowers. Attanasio et al. (2008) report that demand for U.S. car loans is highly sensitive to loan maturity but largely insensitive to interest rates, except among high-income households. In addition to the previous findings, the results suggest that borrowers are restricted in their ability to borrow due to the monthly installment rather than the overall cost of borrowing. Consistent with the previous findings, (Ganong and Noel, 2020) demonstrated that reducing monthly payments through mortgage modifications during the Great Recession significantly decreased delinquency rates, whereas reducing the principal amount of mortgages with no change in monthly payments had little to no impact on delinquency rates. Consequently, longer loan terms may lower delinquency rates by reducing monthly payments; however, they also increase borrowers' exposure to adverse financial conditions.

In U.S. subprime automobile lending, larger loan amounts create greater risk of default for each individual borrower; however, those applicants that pose the greatest risk to default are the ones that apply for the largest loan amounts (Adams et al., 2009; Einav et al., 2012). Kim et al. (2017) found that larger high interest rate private loans made in Korea were predictors of delinquency on formal sector debts and concluded that they reflected deterioration in financial position rather than moral hazard. Dobbie and Skiba (2013) distinguished between the two effects, finding that eligibility criteria that allow borrowers to obtain larger loan amounts reduce default rates in payday lending, but

borrowers who independently obtain larger loan amounts default at higher rates.

Delinquency may be viewed as a continuous process rather than an isolated occurrence. In their study using data from Korean credit cards, Kim et al. (2018) found that the factors that influence when individuals first become delinquent differ significantly from those that lead to higher delinquency levels. For example, younger employees and individuals age 45 or older were found to have a lower probability of becoming delinquent; however, if they did become delinquent, then they were more likely to develop serious delinquency problems. A similar result was found in a UK-based study by Leow and Crook (2014), which reported that employed cardholders were unlikely to experience arrears but were less likely to resolve these issues once they became delinquent.

Multiple borrowing and credit information. The number of loans a borrower holds may indicate either risk or reliability. When competing microfinance lenders entered Ugandan markets, repayment and savings at the largest incumbent lender declined, which the authors attribute to clients borrowing from multiple lenders simultaneously (McIntosh et al., 2005). Theoretical research suggests that competition between lenders can increase such double-dipping and the resulting default (Guha and Roy Chowdhury, 2013). However, multiple borrowing does not always indicate financial distress: among Ghanaian microborrowers, over-indebtedness is linked to adverse shocks, low investment returns, and weak debt literacy, but not to the number of loans held (Schicks, 2014).

A borrower's continued ability to obtain credit may effectively discipline them. In South Africa, a randomly assigned promise of lower-cost subsequent loans for borrowers who make their payments on time resulted in a 13-21% decrease in borrower defaults (Karlan and Zinman, 2009). At the international level, sharing information among lenders has been shown to increase lending activity and decrease borrower default (Jappelli and Pagano, 2002). Adoption of a credit bureau by a Guatemalan micro-lending organization resulted in increased lending activity without resulting in an increase in borrower defaults. Once borrowers realized that the bureau existed, they demonstrated a slight improvement in repayment behavior (de Janvry et al., 2010).

Co-borrowing and guarantees. Guarantees and joint liability may both reduce

delinquency by spreading liability and encouraging monitoring, and by affecting which borrowers are lent to. U.S. lenders lend to marginal borrowers when the share of guaranteed loans incentivizes them to do so (Bachas, Kim and Yannelis, 2021), but this implies guaranteed lending is directed at riskier borrowers. Studies on group lending indicate that the social aspect of shared responsibility primarily influences repayment. In a study of 128 microfinance groups in Bangladesh, group size, credit rationing, and social relationships predicted loan repayment Sharma and Zeller (1997). Additionally, results from two randomized control trials conducted in the Philippines showed that when formal joint responsibility for loans was removed from the experiment, the rate of loan defaults did not increase Giné and Karlan (2014). These findings suggest that social ties among group members, along with the ability to monitor one another's actions, ultimately lead to high rates of repayment, rather than formal responsibility for each other's loans.

Preferential and state-directed lending Preferential credit influences delinquency through both selection effects and borrower incentives. In the context of U.S. student lending, the doubling of default rates is primarily observed among new borrowers at for-profit and two-year institutions, with changes in borrower and school characteristics accounting for between half and two-thirds of this increase (Looney and Yannelis, 2015). Mueller and Yannelis (2019) claim roughly 30% of the increase comes from borrower composition and a similar percentage from house price decline. When lenders cannot perfectly observe risk, this implies that loan terms determine which market composition the lender is offering (Stiglitz and Weiss, 1981). Loans allocated on non-commercial criteria may thus attract less qualified borrowers. In Pakistan, politically connected firms borrow 45 percent more and default 50 percent more, but only at government banks; private banks do not display this bias (Khwaja and Mian, 2005). These programs may also undermine repayment discipline. For example, India's 2008 debt relief reduced beneficiaries' concern about the reputational cost of default (Kanz, 2016). However, repayment worsened only within the beneficiaries not experiencing financial distress and improved among those who were (Mukherjee, Subramanian and Tantri, 2018).

Preferential terms do not uniformly result in higher delinquency rates. Mortgages

originated through U.S. state housing finance agencies are approximately 30 percent less likely to default or enter foreclosure than comparable conventional loans, partly due to enhanced servicing and homebuyer education (Hembre, Moulton and Record, 2021). Similarly, community-reinvestment mortgages exhibit significantly lower default rates than subprime loans extended to comparable borrowers (Ding, Quercia, Li and Ratcliffe, 2011). Consequently, the effect of preferential terms varies across products, depending on the rigor of borrower selection and the specific features of program design.

Macroeconomic factors. Delinquency rates are affected by the economic cycle. An analysis of approximately 700,000 U.S. credit card accounts indicates that a one-percentage-point increase in county unemployment leads to a 2 to 3 percentage point rise in delinquency, even after controlling for state and time effects Agarwal and Liu (2003). Personal bankruptcies also rise during economic downturns (Garrett and Wall, 2014). Macroeconomic variables improve credit-risk models: bank interest rates and unemployment significantly influence loss given default on UK credit cards (Bellotti and Crook, 2012), and the inclusion of macroeconomic and behavioural variables enhances the predictive accuracy of dynamic models of credit card default (Bellotti and Crook, 2013).

3 Model

The binary outcome of interest is whether a contract has overdue principal or not in each month. We thus use a logistic regression model to model the probability of delinquency, wherein the log-odds of the event are assumed to be linearly related to the regressors (Gujarati and Porter, 2009; Wooldridge, 2020). Consumer credit risk is often studied in academics and industry using the account–month panel of credit card data, as pioneered by Gross and Souleles (2002), who model the probability of delinquency and bankruptcy, and by Agarwal and Liu (2003) who demonstrate that macroeconomic conditions influence delinquency probabilities. There are several further developments of the discrete-choice structure, namely, Bellotti and Crook (2013) and Crook and Bellotti (2010) fit DT survival models with time-varying behavioral and macroeconomic variables, Kim et al. (2018) and Leow and Crook (2014) develop models of delinquency state transition processes and also compare different machine learning classification algorithms with a discrete-choice baseline model and Li and Wu (2024) do so too. Similar to the U.S., the same specification dominates the volume of loan-level studies of bank and registry data outside the U.S., such as Jiménez et al. (2014) in Spain and Deiß, Ölkers and Mußhoff (2025) in Mali, for which they consider a commercial bank's 375,438 loans, estimate logit models of repayment performance while controlling for borrower and loan characteristics, agency and year fixed effects, and heteroskedasticity-robust standard errors. Following this literature, we model the level of delinquency risk (as opposed to the timing of default) rather than the characteristics of a borrower, loan, and lender that are associated with repayment difficulties at any given month, our research question.

Let $\mathbf{x}_{it}$ collect the borrower, loan, and contract characteristics defined in Table 1, together with the bank, region, and year indicators, and let $b(i)$, $r(i)$, and $y(t)$ denote the bank that issued contract i, the region in which it was issued, and the calendar year of reporting month t. Writing $P_{it} = \Pr\big(D_{it} = 1 \mid \mathbf{x}_{it}\big)$ for the conditional probability that the delinquency indicator of equation (5) equals one, the pooled model for all loan types is

$$P_{it} = \Lambda(\eta_{it}) = \frac{\exp(\eta_{it})}{1 + \exp(\eta_{it})}, \qquad \text{or equivalently} \qquad \ln\left(\frac{P_{it}}{1 - P_{it}}\right) = \eta_{it}, \tag{1}$$

where $\Lambda(\cdot)$ is the logistic cumulative distribution function and η_{it} is the linear index

$$\begin{aligned}\eta_{it} = \alpha &+ \beta_1 \text{Income}_{it} + \beta_2 \text{Age}_{it} + \beta_3 \text{AgeSq}_{it} + \beta_4 \text{LoanAmount}_i + \beta_5 \text{Rate}_i \\ &+ \beta_6 \text{Maturity}_i + \beta_7 N_{it} + \gamma_1 \text{Gender}_i + \gamma_2 \text{CoBorrower}_i + \gamma_3 \text{Preferential}_i \\ &+ \gamma_4 \text{Online}_i + \mu_{b(i)} + \lambda_{r(i)} + \tau_{y(t)}.\end{aligned} \tag{2}$$

The quadratic age term is the squared age in the estimation sample,

$$\text{AgeSq}_{it} = \text{Age}^2, \tag{3}$$

and enters every specification in this form. Income and loan amount are strongly right-skewed and therefore enter in natural logarithms, as in equations (6) and (7), so that Income_{it} and LoanAmount_i denote the logarithm of the borrower's average monthly income and of the contract amount, respectively. N_{it} is the number of contracts the borrower has signed up to the reporting date.

This heterogeneity is typically captured by bank, region and time effects in loan-level studies of credit registers, and is treated in this way in our study. The bank effects μ_b reflect differences in the underwriting practices, collection practices, and composition of banks' lending portfolios; the region effects λ_r reflect the persistent regional differences in economic growth, labour markets, and payment culture; and the year effects τ_y reflect aggregate shocks common to all borrowers in a year, such as changes in monetary policy, inflation, and the business cycle, which the literature identifies as important drivers of delinquency (Agarwal and Liu, 2003; Bellotti and Crook, 2013).

The coefficients capture the effect on the log odds of being delinquent (Gujarati and Porter, 2009; Wooldridge, 2020). Their signs and magnitudes provide meaningful information, taking into account the other predictors, as well as bank, region, and time dummies. However, the resulting change in the probability depends on the borrower's baseline risk through the factor $p(1-p)$. In most of the relevant literature, the odds ratio $\exp(\beta)$ is referred to as such, as this is constant for all borrowers (e.g. Deiß, Ölkers and Mußhoff, 2025).

There has been some contradictory evidence regarding the shape of the age-default relation, some papers showing the risk to be monotonically decreasing with age (Debbaut et al., 2016; Li and Wu, 2024), while others reporting that the relation has a U-shape (Agarwal et al., 2009). To accommodate non-linearity without committing to one or the other specification, we control for age along with AgeSq_{it}, the squared age, as given by equation (3).

The retail lending products vary in their purpose, collateral, pricing, distribution, and borrower targets, hence the delinquency determinants may be different from one type to another. We use this difference in estimating the model separately for seven loan types: mortgage credit, micro credit, micro loans, auto credit, credit cards, education credit, and consumer credit. For product k, the model is estimated on the subsample of contracts of that type:

$$\Pr\Big(D_{it} = 1 \mid \mathbf{x}_{it}^{k}, k(i) = k\Big) = \Lambda\Big(\alpha^{k} + \mathbf{c}_{it}'\boldsymbol{\beta}^{k} + \mathbf{d}_{i}^{k\prime}\boldsymbol{\gamma}^{k} + \mu_{b(i)}^{k} + \lambda_{r(i)}^{k} + \tau_{y(t)}^{k}\Big). \tag{4}$$

Every coefficient and every fixed effect carries the superscript k, so the effect of each characteristic, as well as the bank, region, and year effects, is allowed to differ freely across products.

4 Data

Data source. We use loan-level supervisory data from the credit registry of the Central Bank of the Republic of Uzbekistan. Commercial banks are required to submit all retail credit contracts for registration with the registry and to submit monthly updates of their status, so the information refer to all bank credit contracts, and not a sample. Individual records are one contract per one reporting month, and include the terms of the contract (amount, interest rate, signing date of the contract, end date, type of credit, channel of disbursement, type of collateral), the status of repayment (outstanding and overdue balances), the bank or region where the contract was signed, and borrower characteristics (date of birth, gender, average monthly income over the previous 12 months).

Each loan has a unique contract registration number throughout its life, and each individual has a personal identification number that connects all their contracts, even if they are at other banks. The unit of observation is the contract–month. This is an unbalanced panel of observations starting from January 2023 to March 2026 with 232,893,178 observations across 16,269,308 contracts held by 7,217,039 borrowers, with the number of observations per contract varying across the panel.

Variables. Table 1 gives the source field and construction rule for each variable. We index contracts by i, borrowers by j, and reporting months by t; time-varying variables are measured at t, not at origination.

The dependent variable indicates whether there is an overdue principal balance at time t.

$$D_{it} = \mathbf{1}[\text{Overdue amount}_{it} > 0]. \tag{5}$$

Since it is recomputed monthly, the contract may move in and out of delinquency, and the concept captures even the first difficulty with repayment rather than the extreme case of default.

The age variable is expressed in years as of the reporting date, while the gender variable is a dummy variable in which 1 represents males. The total number of loans indicates the total number of credit contracts that the borrower has entered into with any bank up to reporting period t, which includes the current credit contract, and it takes into account both active and repaid credit contracts. It is an increasing variable because it increases at each reporting period with the number of new loans taken by the borrower.

Income and loan amount are strongly right-skewed and enter in natural logarithms:

$$\text{Income}_{it} = \ln(\text{Monthly income}_{it}), \tag{6}$$

$$\text{LoanAmount}_{i} = \ln(\text{Contract amount}_{i}), \tag{7}$$

so observations with zero reported income are treated as missing. The interest rate is the annual rate paid by the borrower, in percent, and maturity is the number of months between the signing and end dates. Loans under subsidised or state-supported

programmes are identified by their below-market pricing. Therefore, preferential loans can be determined when the interest rate charged on the loan by the borrower is less than 18%. This will be achieved using a dummy variable which will be assigned the value 1 if the interest rate charged by the borrower is less than 18%.

The remaining variables are dummies. Online equals one if the loan was disbursed online and zero if it was disbursed at a bank branch. Co-borrower equals one if a co-borrower identifier is registered on the contract.

Sample cleaning. We drop observations with a maturity below zero or above 240 months, monthly income above 100 million soum, a contract amount above 1 billion soum, or a negative interest rate, all of which indicate data-entry errors. Additionally, only 60.4% of the observations have non-missing income, so we only include these in the sample. After cleaning, the data set consists of 140,630,546 observations, from which we have made empirical analysis.

Descriptive statistics. Table 2 summarises the variables. The average borrower is 36.1 years old and reports a monthly income of 6 million soum, well above the median of 3.7 million; loan amounts are similarly skewed (mean 28.2 million, median 10.0 million), which motivates the log transformations. The average interest rate is 28.3 percent, the average maturity 43.1 months, and the average borrower has 3.1 loans. Men account for 58.7 percent of observations, preferential loans for 25.2 percent, and online loans for 48.1 percent; only 2.9 percent of contracts have a co-borrower. The delinquency rate is 13.2 percent.

The delinquency patterns of borrowers and loans based on borrower characteristics and loan characteristics are presented in Figures 1 through 6. Delinquency rates are calculated by dividing the total number of contract-months where a borrower was behind in their payments on the principal portion of their mortgage by the total number of contract-months across all loans during the period from January 2023 to March 2026.

The delinquency rate is highest among younger borrowers (Figure 1a): 17.7 percent of borrowers ages 25 to 30 are delinquent, a rate that decreases steadily with age. Men are more likely to be delinquent than women at all ages until around 55, when the trend

Table 1: Construction of variables

Variable	Registry field(s)	Construction rule
Dependent variable		
Delinquency	Overdue loan amount	1 if overdue amount > 0 in month t; 0 otherwise
Borrower characteristics		
Age	Date of birth; reporting date	(Reporting date − date of birth) in days / 365
Gender	Borrower gender	1 = male (code 1); 0 = female (code 2)
Number of loans	Borrower ID; contract registration number; signing date	Number of the borrower's contracts signed on or before the reporting date
ln(Income)	Income	Log of average monthly income over the previous 12 months (soum)
Loan characteristics		
ln(Loan amount)	Contract amount	Log of contract amount (soum)
Interest rate	Interest rate paid by borrower	Annual rate, %
Maturity	Signing date; end date	(End date − signing date) in days / 30; kept if between 0 and 240 months
Preferential	Interest rate	1 if rate $< 18\%$; 0 if rate $\geq 18\%$
Online	Disbursement mode	1 if disbursed online; 0 if offline
Co-borrower	Co-borrower ID	1 if a co-borrower ID is registered; 0 otherwise

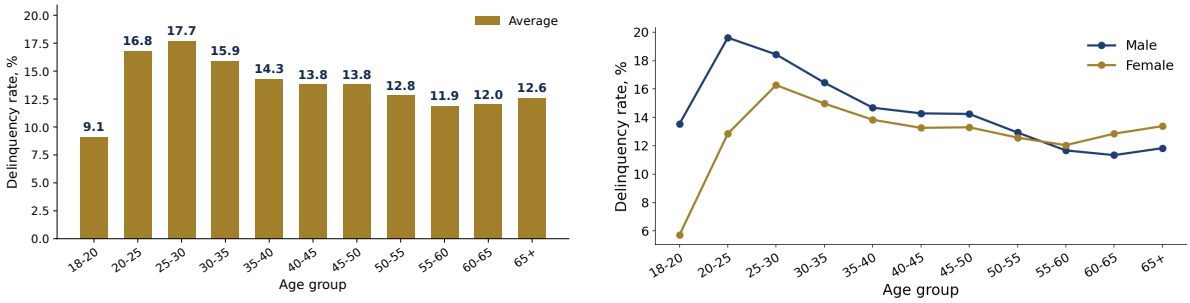
Notes: Time-varying variables are measured at the monthly reporting date t. Bank, region and year fixed effects are constructed from the reporting bank code, the region of issuance and the reporting date.

Table 2: Descriptive statistics

Variable	Mean	SD	Median	Min	Max
Panel A. Continuous variables					
Borrower characteristics					
Income (mln soum)	6.02	7.21	4.02	0.3	100.00
Age (years)	36.09	11.16	34.40	18	98.3
Number of loans	3.07	2.93	2.00	1	18
Loan characteristics					
Loan amount (mln soum)	28.19	52.99	10.00	0.1	1,000.00
Interest rate (%)	28.26	14.79	28.00	0.00	80.00
Maturity (months)	43.15	40.83	36.40	1	240.0
Panel B. Binary variables					
Variable					Share equal to one (%)
Delinquency					13.20
Male					58.67
Co-borrower					2.94
Preferential					25.2
Online origination					48.07

Notes: The statistics cover 140,630,546 contract–month observations from January 2023 to March 2026. Money amounts are in millions of soum, whereas income and loan amount are logged in the analysis (equations (6) and (7)). Variables are defined in Table 1.

Figure 1: Delinquency and borrower age

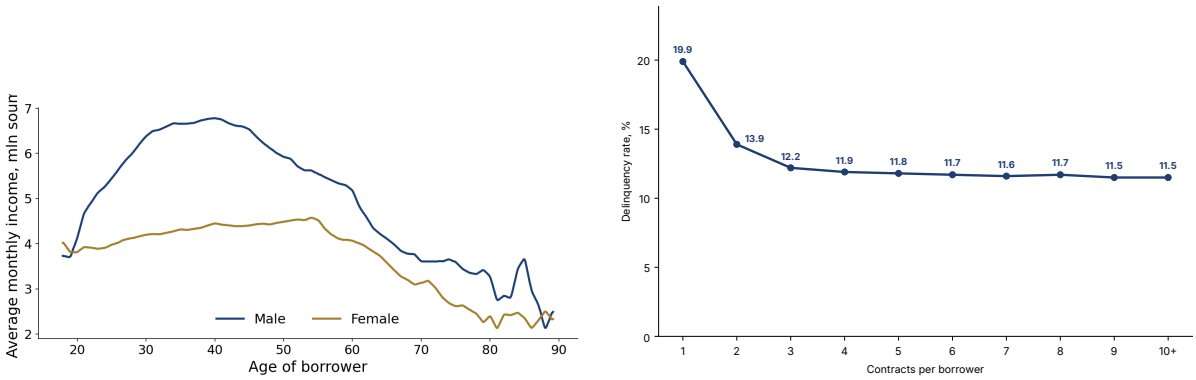


(a) Delinquency rate by age group **(b)** Delinquency rate by age group and gender

Note: Share of contract–months with overdue principal, in percent, by age group (panel a) and by age group and gender (panel b). All loan types, January 2023–March 2026.

changes (Figure 1b).

Figure 2: Borrower income and the number of contracts



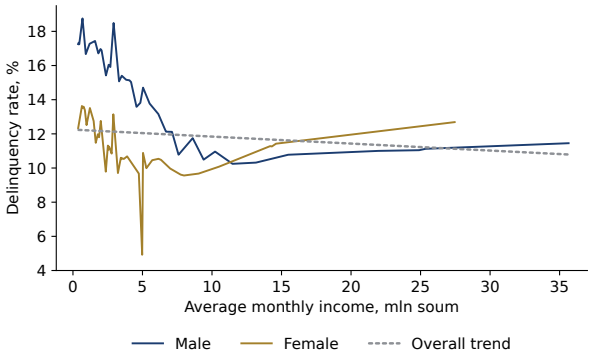
(a) Average monthly income by age and gender **(b)** Delinquency rate by number of contracts

Note: In panel (a), mean monthly income is shown in million soums for individual ages, per gender. Panel (b) shows the percentage of contract months in which the outstanding principal amount was late, per individual contract count. Borrowers who had 10 or more contracts are classified together. The period is January 2023-March 2026.

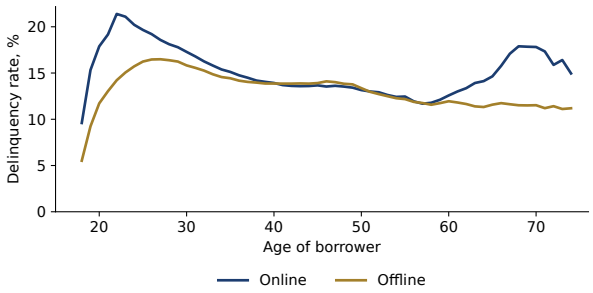
Income rises through the middle years of life and is generally higher for males (Figure 2a). By contrast, delinquency falls with the number of contracts a borrower holds (Figure 2b): 19.9 percent of the contract-months of borrowers with a single contract are delinquent, compared with 13.9 percent for borrowers with two contracts and about 12 percent for those with three or more, beyond which the rate is roughly flat.

For males, delinquency drops dramatically at the lower end of the income distribution (Figure 3a). The frequency of delinquency is lower for women with low income than for men with low income but begins to increase gradually after about 10 million soum. In

Figure 3: Delinquency by income, gender and lending channel



(a) Delinquency rate by income and gender

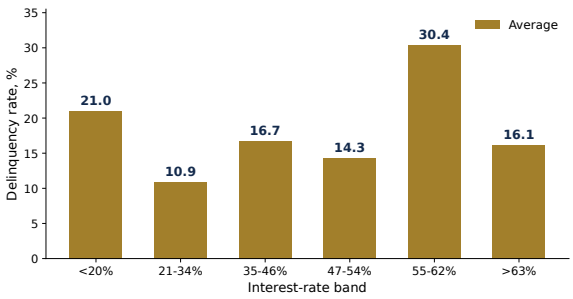


(b) Delinquency rate by age and lending channel

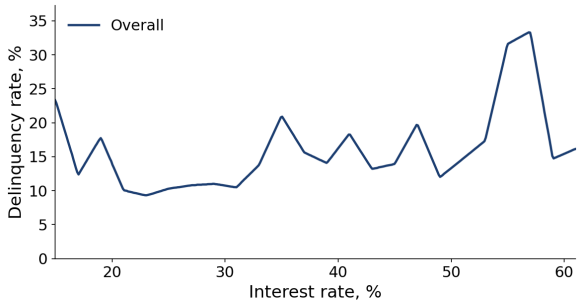
Note: Percentage share of the contract–months with overdue principal payments, as a percentage. Panel (a) plots the percentage share rate against the average monthly salary, in millions of soum, for both male and female individuals and the dotted line represents the linear trend of all loan borrowers. Panel (b) shows the percentage share rate by single year of age for online and bank loans.

Panel (b), we compare the lending channels (Figure 3b). Almost at all ages, online loans are more likely to be delinquent.

Figure 4: Loan pricing and delinquency



(a) Delinquency rate by interest-rate band



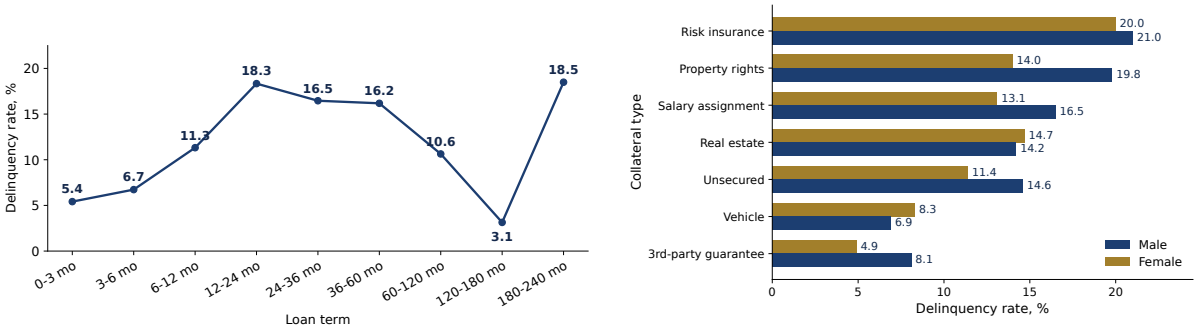
(b) Delinquency rate by interest rate

Note: Share of contract-months in arrears of the principal, per cent, by interest rate bands (panel a) and by annual interest rate (panel b). Both panels analyse for all loan types.

Loan pricing and delinquency is lowest on loans that charge 21-34 percent interest, and rises and falls on loans that charge more than 34 percent, but rises sharply to 30.4 percent on loans that charge 55-62 percent interest (Figures 4a and 4b).

The longer the maturity of the loan, the higher the delinquency rate, and loans with a term of 120–180 months have the lowest rate, although it is not zero, and then it increases with the loan term further on (Figure 5a). The loans secured by risk insurance or property rights have the highest delinquency rate (nearly 20 percent), while those secured

Figure 5: Loan maturity, collateral and delinquency



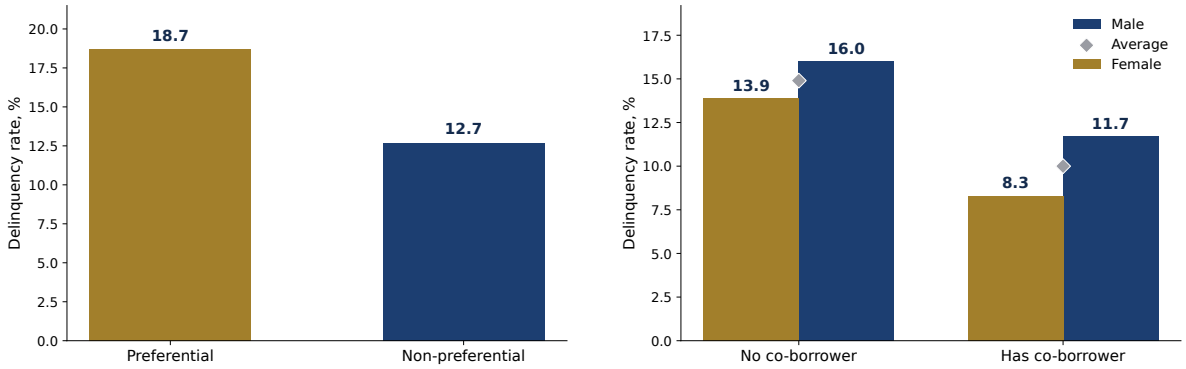
(a) Delinquency rate by loan term

(b) Delinquency rate by collateral type

Note: Share of contract–months with overdue principal, in percent. Panel (a) plots the rate by contractual loan term for all borrowers. Panel (b) plots the rate by collateral type, separately for men and women.

by vehicles or third-party guarantees have the lowest (Figure 5b).

Figure 6: Preferential pricing, co-borrowing and delinquency



(a) Contract-month delinquency by loan-pricing type

(b) Delinquency rate by co-borrower status

Note: Panel (a) plots the share of contract-month delinquency for preferential loans (interest rate below 18 per cent) and market-priced loans; percentages below the bars are shares of all contracts. Panel (b) plots the share of contract–months with overdue principal by co-borrower status and gender; diamonds mark the average for each group.

The preferential loans are significantly riskier than the market-priced loans (Figure 6a). Co-borrowers, on the other hand, have delinquency rates that are, on average, one-third lower, with an even greater decrease for women than men (Figure 6b).

5 Empirical Analysis

5.1 Pooled estimates

Results from the pooled logit regression of the model in equation (2), applied to all loans, are shown in Table A.1. All estimates control for fixed effects and use clustered standard errors at the borrower level. Column (1) contains coefficient estimates, Column (2) shows standardised coefficients for continuous variables, and Column (3) presents marginal effects on the probability of becoming delinquent, in percentage points, evaluated at the means. All variables of interest are statistically significant at the 1 per cent level. In the following discussion, marginal effects are used to describe the magnitude of the effects.

The higher the income, the lower the probability of becoming delinquent. However, the effect is relatively small. A 1 log-point increase in income reduces the probability of delinquency by 1.65 percentage points. This is consistent with the assumption that individuals with higher incomes have more capacity to maintain debt payments in the face of unexpected expenditures or a drop in income.

The relationship with age is non-linear. The linear term is negative (-0.029). At the average age, one extra year reduces the probability of delinquency by 0.15 percentage points. The probability of delinquency decreases with age and the rate of decrease becomes smaller for older borrowers, but it does not increase again. It is thus more similar to the monotonic decrease in Debbaut et al. (2016) and Li and Wu (2024) than to the U-shape in Agarwal et al. (2009) and agrees with Figure 1a.

With regard to loan characteristics, larger loan amounts and higher interest rates are positively correlated with delinquency. The probability of delinquency increases by 0.05 percentage points when the loan is 10 percent larger and increases by 0.09 percentage points when the interest rate is increased by one percentage point. Longer maturity loans have a somewhat lower probability of delinquency: each additional twelve months decreases the probability of delinquency by about 0.18 percentage points, presumably because monthly payments become smaller. Each additional loan in the borrower's credit portfolio increases the probability of delinquency by 0.12 percentage points. The number

of loans depends on the borrower's previous behavior, so we consider it as an association, not a causal effect.

The effects of the binary variables are considerably larger. Men are 2.6 percentage points more likely to be delinquent than women, which agrees with earlier studies showing better repayment among women (D'Espallier et al., 2011; Shahriar et al., 2020). Loans with a co-borrower are 2.1 percentage points less likely to be delinquent. Preferential loans are 6.0 percentage points more likely to be delinquent than loans at market rates, and loans issued online are 4.4 percentage points more likely to be delinquent than loans issued in a branch.

The coefficients in column (1) are measured in different units (years of age, months of maturity, percentage points of interest, log points of income), so their sizes cannot be compared directly. Column (2) makes the continuous variables comparable by expressing each of them in standard deviations. On this scale, evaluated at the mean age, age and income have effects of similar size (about -0.16 and -0.15). They are followed by the interest rate (0.13), maturity (-0.08) and loan size (0.08).

5.2 Heterogeneity across loan types

Evaluating equation (4) individually for each type of product highlights how much is hidden in the combined estimates. The coefficients, standardised coefficients, and marginal effects are presented in Tables A.2, A.3, and A.4. All models include fixed effects by bank, region, and year and robust standard errors by borrowing identity. Some variables have similar signs for all types of products: the income variable has a negative effect on delinquency in six of the seven loan types, having a co-borrower decreases delinquency in all product types where there is co-borrowing, and online loans increase delinquency in all product types where the effect is significant. The remaining variables vary in sign across products.

Delinquency declines with higher income for almost all types of loans, but the association varies greatly by type of product (Figure 7, panel a).

Similarly, the association between age and delinquency is product-dependent. For

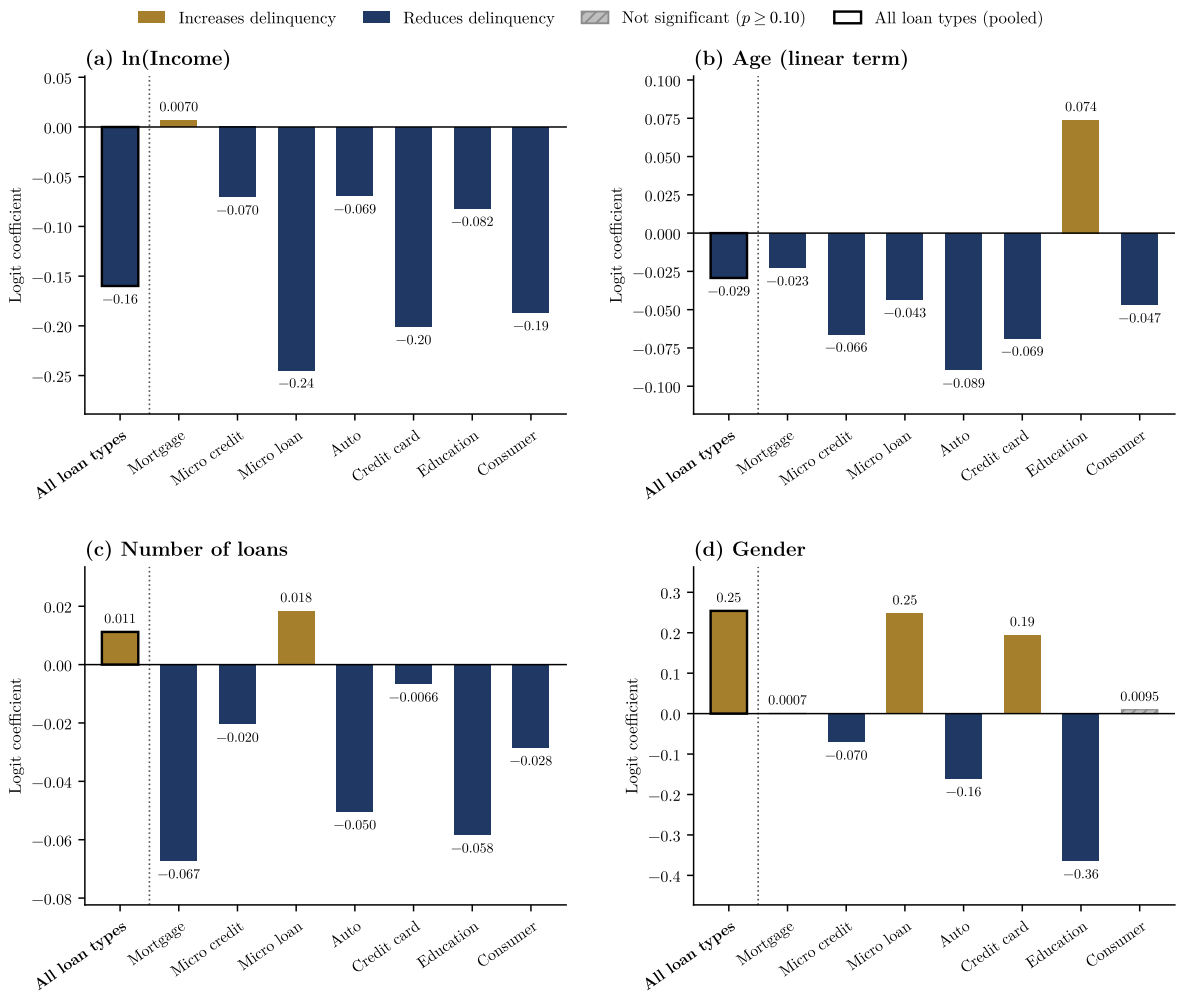


Figure 7: Borrower characteristics and delinquency: logit coefficients by loan type

Note: Bars represent logit coefficients of the log-odds of delinquency, panel (d), where gender = 1 is male. The first bars shown in the panels (in bold) are the pooled estimate from regression (1) of Table A.1, whereas other bars are the product-specific estimates from Table A.2. The gold bars signify positive coefficient significance at the 10 per cent level and navy bars signify negative coefficient significance at the 10 per cent level, whereas hatched grey bars represent non-significance.

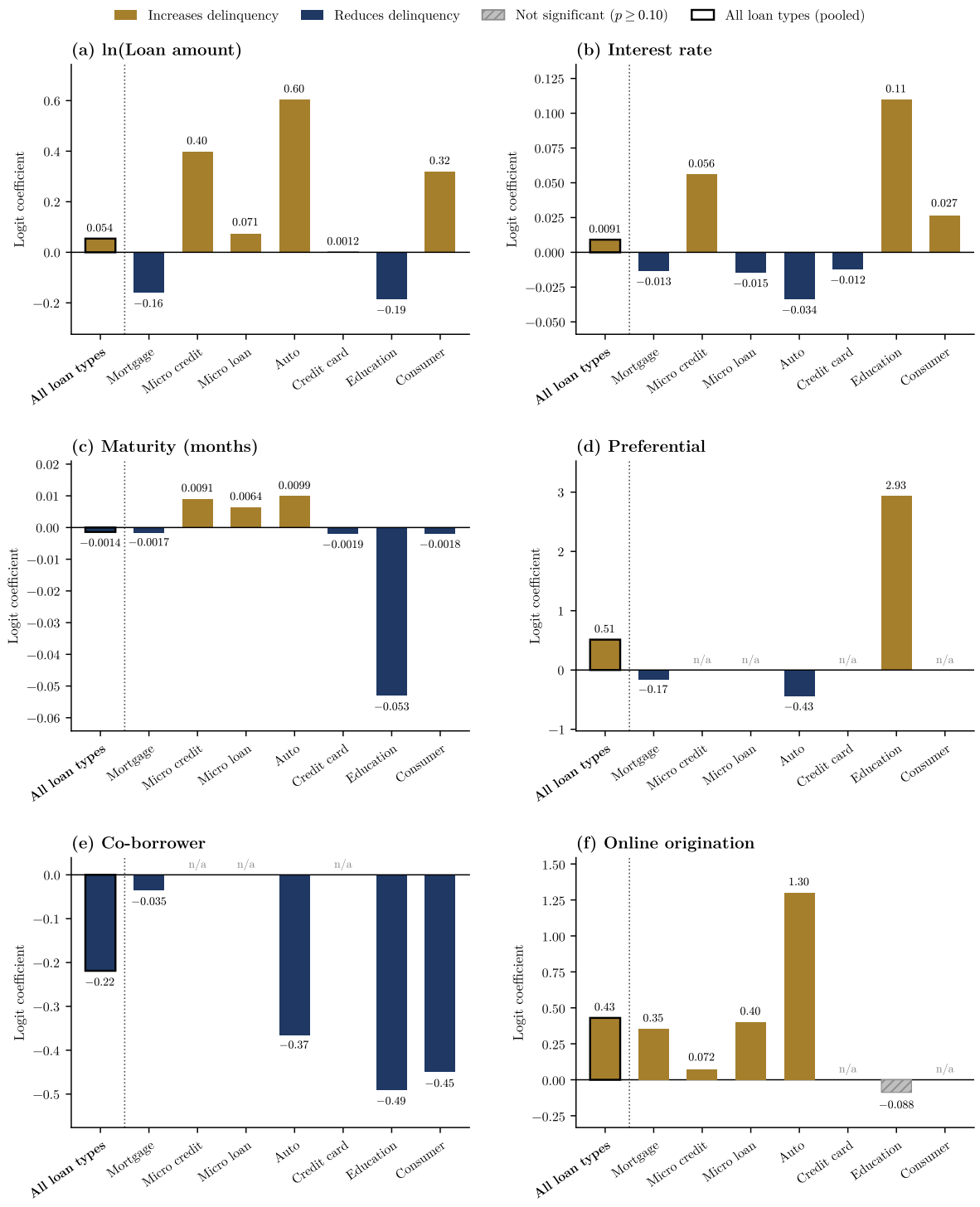


Figure 8: Loan characteristics and delinquency: logit coefficients by loan type

Note: "n/a" indicates that the regressor is not part of that product's specification because it does not vary within the product.

most products, the delinquency rate decreases with age in early and middle adulthood and increases among older borrowers, resulting in a U-shaped association. For some products, however, the delinquency rate either continues to decrease throughout the majority of ages or exhibits an inverted-U shape. All of these associations found in previous research occur in our dataset, but for different products. It seems that the association between age and delinquency is as dependent on the credit product as the age of the borrower itself.

Loan size and pricing have specific effects on each product as well. For most unsecured and depreciating-collateral products, loan size is positively correlated with delinquency because it increases the repayment burden and because there is no collateral or the collateral depreciates. This correlation applies to interest rates as well, where the higher rate is positively correlated with delinquency. The correlation in the other products is insignificant and negative.

The maturity of the loan has a relatively small relationship with delinquency in most loan products, but it plays an important role in the case of education credit, where long-term maturities are highly correlated with low levels of delinquency. This is because longer education loans reduce the payment amount and postpone repayment to a period when the borrower's earnings are at their minimum. In several cases of short-term and depreciating collateral loan products, however, long-term maturities increase delinquency levels.

Gender is the least stable determinant of delinquency among the products. Men have a higher likelihood of delinquency for some credit products but a lower likelihood for others, and in some products, there is little difference between the two genders. The varying findings with respect to gender and repayment performance in the literature might thus be partly attributed to the different products analysed. On the other hand, co-borrowing is negatively associated with delinquency across all products.

There is a relationship between preferential terms and lower levels of delinquency in those products where the comparison with the market-priced loan can be made more directly. The only product that stands out here is education credit because of its higher probability of delinquency. It is most reasonable to interpret the direction of the effect

observed for education credits through selection. The eligibility for special educational loans is determined by certain criteria like family income or social class, hence the indicator mainly determines a new set of borrowers instead of different conditions being provided to the same borrower.

Online origination is the most consistent risk factor in the data: wherever the channel effect is statistically significant, loans originated online are more likely to become delinquent than those originated in branches. The effect varies considerably in size.

5.3 Robustness checks

Estimation on a 10 percent subsample. With roughly 140 million contract–months, almost any coefficient clears conventional significance thresholds, so we re-estimate the pooled model of Table A.1 on a random 10 percent subsample. The draw is at the borrower level, so a selected borrower enters with his or her full monthly history and the panel structure the clustered standard errors rely on is preserved; nothing else changes. Table A.5 shows every coefficient keeps its sign and its significance at the 1 percent level.

Interaction effects The baseline model assumes that the effect of each variable is the same for all borrowers, whatever their other characteristics. We relax this assumption by adding two interactions, one at a time: age with gender (Table A.6) and log income with log loan amount (Table A.7). Both interactions are significant at the 1 percent level, but they change the main results very little. The largest change in a marginal effect is for the number of loans, at 8 percent in the first specification and 16 percent in the second. No other marginal effect changes by more than 9 percent. The results in Section 5.1 therefore hold when these interactions are allowed. When reading these tables, note that the coefficients in columns (1) and (2) on the interacted variables show their effect when the other variable equals zero. For the log variables this lies far outside the data, so these coefficients cannot be compared with the baseline, and the marginal effects in column (3) should be used instead.

The age–gender interaction is negative (-0.0074), which means that delinquency falls

faster with age for men than for women. The higher risk of male borrowers is therefore concentrated among younger borrowers. The odds ratio of men relative to women falls from 1.47 at age 20 to 1.27 at age 40 and 1.09 at age 60, and the gap disappears at about age 72. The income–loan interaction is positive, so income protects less as loan size increases. A one log-point increase in income lowers the probability of delinquency by 1.99 percentage points for a loan one log point below the mean, by 1.59 points at the mean loan size, and by 1.19 points for a loan one log point above the mean.

6 Conclusion

Income and age are the most important determinants of delinquency among continuous variables. Higher income is associated with lower delinquency in six of seven products, especially in unsecured loans, where payments depend on the borrower's current income. Age has different patterns across products. This variable is U-shaped in five products, decreases across most of the age range in micro loans and the pooled portfolio, and is inverted-U shaped in education credit. Both the U-shape and the decreasing age pattern found in previous studies appear in our dataset, but in different products. Contract characteristics are as important as the borrower's features. Having a co-borrower is associated with lower delinquency in all products where co-borrowers are involved, and online origination is associated with higher delinquency in all products. Most importantly, some determinants switch signs across product categories. The loan amount, the interest rate, gender and the number of loans are all inversely correlated with one another in different segments, and hence the pooled estimates do not apply to any particular product type.

References

Adams, William, Liran Einav, and Jonathan Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, 2009, *99* (1), 49–84.

Agarwal, Sumit and Chunlin Liu, "Determinants of credit card delinquency and bankruptcy: Macroeconomic factors," *Journal of Economics and Finance*, 2003, *27* (1), 75–84.

Agarwal, Sumit, Jia He, Tien Foo Sing, and Jian Zhang, "Gender gap in personal bankruptcy risks: Empirical evidence from Singapore," *Review of Finance*, 2018, *22* (2), 813–847.

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson, "The age of reason: Financial decisions over the life cycle and implications for regulation," *Brookings Papers on Economic Activity*, 2009, (2). Fall 2009.

Attanasio, Orazio P., Pinelopi Koujianou Goldberg, and Ekaterini Kyriazidou, "Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data on Car Loans," *International Economic Review*, 2008, *49* (2), 401–436.

Bachas, Natalie, Olivia S. Kim, and Constantine Yannelis, "Loan guarantees and credit supply," *Journal of Financial Economics*, 2021, *139* (3), 872–894.

Bao, L. and D. Huang, "Gender differences in reactions to loan collection mechanisms: A large-scale natural field experiment," *Journal of Financial Intermediation*, 2026. Forthcoming.

Bellotti, Tony and Jonathan Crook, "Loss given default models incorporating macroeconomic variables for credit cards," *International Journal of Forecasting*, 2012, *28* (1), 171–182.

Bellotti, Tony and Jonathan Crook, "Forecasting and stress testing credit card default using dynamic models," *International Journal of Forecasting*, 2013, *29* (4), 563–574.

Chen, Jia, Jiajun Jiang, and Yu-Jane Liu, "Financial literacy and gender difference in loan performance," *Journal of Empirical Finance*, 2018, *48*, 307–320.

Crook, Jonathan and Tony Bellotti, "Time varying and dynamic models for default risk in consumer loans," *Journal of the Royal Statistical Society: Series A*, 2010, *173* (2), 283–305.

de Janvry, Alain, Craig McIntosh, and Elisabeth Sadoulet, "The Supply- and Demand-Side Impacts of Credit Market Information," *Journal of Development Economics*, 2010, *93* (2), 173–188.

Debbaut, Peter, Andra C. Ghent, and Marianna Kudlyak, "The CARD Act and young borrowers: The effects and the affected," *Journal of Money, Credit and Banking*, 2016, *48* (7), 1495–1513.

Deiß, Johanna, Tim Ölkers, and Oliver Mußhoff, "Gender and credit: Inequalities in loan size and credit repayment performance? Evidence from Mali," *Journal of Rural Studies*, 2025, *120*, 103888.

D'Espallier, Bert, Isabelle Guérin, and Roy Mersland, "Women and repayment in microfinance: A global analysis," *World Development*, 2011, *39* (5), 758–772.

Ding, Lei, Roberto G. Quercia, Wei Li, and Janneke Ratcliffe, "Risky Borrowers or Risky Mortgages: Disaggregating Effects Using Propensity Score Models," *Journal of Real Estate Research*, 2011, *33* (2), 245–278.

Dobbie, Will and Paige Marta Skiba, "Information Asymmetries in Consumer Credit Markets: Evidence from Payday Lending," *American Economic Journal: Applied Economics*, 2013, *5* (4), 256–282.

Einav, Liran, Mark Jenkins, and Jonathan Levin, "Contract Pricing in Consumer Credit Markets," *Econometrica*, 2012, *80* (4), 1387–1432.

Fonseca, Raquel, Kathleen J. Mullen, Gema Zamarro, and Julie Zissimopoulos, "What explains the gender gap in financial literacy? The role of household decision-making," *Journal of Consumer Affairs*, 2012, *46* (1), 90–106.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, "The role of technology in mortgage lending," *The Review of Financial Studies*, 2019, *32* (5), 1854–1899.

Ganong, Peter and Pascal Noel, "Liquidity versus Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession," *American Economic Review*, 2020, *110* (10), 3100–3138.

Garrett, Thomas A. and Howard J. Wall, "Personal-Bankruptcy Cycles," *Macroeconomic Dynamics*, 2014, *18* (7), 1488–1507.

Giné, Xavier and Dean S. Karlan, "Group versus individual liability: Short and long term evidence from Philippine microcredit lending groups," *Journal of Development Economics*, 2014, *107*, 65–83.

Godquin, Marie, "Microfinance repayment performance in Bangladesh: How to improve the allocation of loans by MFIs," *World Development*, 2004, *32* (11), 1909–1926.

Gross, David B. and Nicholas S. Souleles, "An empirical analysis of personal bankruptcy and delinquency," *The Review of Financial Studies*, 2002, *15* (1), 319–347.

Guha, Brishti and Prabal Roy Chowdhury, "Micro-Finance Competition: Motivated Micro-Lenders, Double-Dipping and Default," *Journal of Development Economics*, 2013, *105*, 86–102.

Gujarati, Damodar N. and Dawn C. Porter, *Basic Econometrics*, 5 ed., New York: McGraw-Hill Irwin, 2009.

Hembre, Erik, Stephanie Moulton, and Matthew Record, "Low-Income Homeownership and the Role of State Subsidies: A Comparative Analysis of Mortgage Outcomes," *Journal of Policy Analysis and Management*, 2021, *40* (1), 78–106.

Jappelli, Tullio and Marco Pagano, "Information Sharing, Lending and Defaults: Cross-Country Evidence," *Journal of Banking & Finance*, 2002, *26* (10), 2017–2045.

Jiménez, Gabriel, Steven Ongena, José-Luis Peydró, and Jesús Saurina, "Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?," *Econometrica*, 2014, *82* (2), 463–505.

Kanz, Martin, "What Does Debt Relief Do for Development? Evidence from India's Bailout for Rural Households," *American Economic Journal: Applied Economics*, 2016, *8* (4), 66–99.

Karlan, Dean and Jonathan Zinman, "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," *Econometrica*, 2009, *77* (6), 1993–2008.

Khwaja, Asim Ijaz and Atif Mian, "Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market," *Quarterly Journal of Economics*, 2005, *120* (4), 1371–1411.

Kim, Hyeongjun, Hoon Cho, and Doojin Ryu, "An empirical study on credit card loan delinquency," *Economic Systems*, 2018, *42* (3), 437–449.

Kim, Jung-Wook, Seungyeon Won, and Jung-In Kim, "Additional Credit for Liquidity-Constrained Individuals: High-Interest Consumer Credit in Korea," *Emerging Markets Finance and Trade*, 2017, *53* (1), 109–127.

Leow, Mindy and Jonathan Crook, "Intensity models and transition probabilities for credit card loan delinquencies," *European Journal of Operational Research*, 2014, *236* (2), 685–694.

Li, Huan and Weixing Wu, "Loan default predictability with explainable machine learning," *Finance Research Letters*, 2024, *60*, 104867.

Livshits, Igor, James MacGee, and Michèle Tertilt, "Accounting for the rise in consumer bankruptcies," *American Economic Journal: Macroeconomics*, 2010, *2* (2), 165–193.

Looney, Adam and Constantine Yannelis, "A Crisis in Student Loans? How Changes in the Characteristics of Borrowers and in the Institutions They Attended Contributed to Rising Loan Defaults," *Brookings Papers on Economic Activity*, 2015, *2015* (2), 1–89.

McIntosh, Craig, Alain de Janvry, and Elisabeth Sadoulet, "How Rising Competition Among Microfinance Institutions Affects Incumbent Lenders," *Economic Journal*, 2005, *115* (506), 987–1004.

Mueller, Holger M. and Constantine Yannelis, "The Rise in Student Loan Defaults," *Journal of Financial Economics*, 2019, *131* (1), 1–19.

Mukherjee, Saptarshi, Krishnamurthy Subramanian, and Prasanna Tantri, "Borrowers' Distress and Debt Relief: Evidence from a Natural Experiment," *Journal of Law and Economics*, 2018, *61* (4), 607–635.

Reboul, Elena, Isabelle Guérin, and Christophe Jalil Nordman, "The gender of debt and credit: Insights from rural Tamil Nadu," *World Development*, 2021, *142*, 105363.

Schicks, Jessica, "Over-Indebtedness in Microfinance – An Empirical Analysis of Related Factors on the Borrower Level," *World Development*, 2014, *54*, 301–324.

Shahriar, A. Z. M., L. A. Unda, and Q. Alam, "Gender differences in the repayment of microcredit: The mediating role of trustworthiness," *Journal of Banking and Finance*, 2020, *110*, 105685.

Sharma, Manohar and Manfred Zeller, "Repayment performance in group-based credit programs in Bangladesh: An empirical analysis," *World Development*, 1997, *25* (10), 1731–1742.

Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 1981, *71* (3), 393–410.

Wooldridge, Jeffrey M., *Introductory Econometrics: A Modern Approach*, 7 ed., Boston, MA: Cengage Learning, 2020.

A Appendix

Table A.1: Logit estimates of delinquency, full sample

	Coefficient	Standardised coefficient	Marginal effect (pp)
ln(Income)	−0.1599***	−0.1491***	−1.6522***
	(0.0011)	(0.0010)	(0.0112)
Age	−0.0293***	−0.3257***	−0.1507***
	(0.0006)	(0.0066)	(0.0011)
Age^2	0.0002***	0.1803***	—
	(0.000007)	(0.0067)	
ln(Loan amount)	0.0542***	0.0804***	0.5603***
	(0.0009)	(0.0014)	(0.0096)
Interest rate	0.0091***	0.1305***	0.0941***
	(0.0001)	(0.0021)	(0.0015)
Maturity (months)	−0.0014***	−0.0834***	−0.0148***
	(0.000032)	(0.0018)	(0.0003)
Number of loans	0.0112***	0.0372***	0.1155***
	(0.0005)	(0.0016)	(0.0050)
Gender (male=1)	0.2540***	0.2540***	2.5711***
	(0.0023)	(0.0023)	(0.0230)
Co-borrower	−0.2189***	−0.2189***	−2.0961***
	(0.0058)	(0.0058)	(0.0511)
Preferential	0.5122***	0.5122***	5.9615***
	(0.0059)	(0.0059)	(0.0758)
Online/Offline	0.4301***	0.4301***	4.4268***
	(0.0026)	(0.0026)	(0.0270)
Bank, region and year FE	Yes	Yes	Yes
Observations	140,630,546	140,630,546	140,630,546
Pseudo R^2	0.05	0.05	

Notes: Pooled logit estimates over all seven loan types, with standard errors clustered by borrower in parentheses. The dependent variable equals one if the contract has overdue principal in the reporting month. Column (1) reports raw coefficients; column (2) rescales the continuous regressors to units of one standard deviation, leaving the binary regressors unchanged; column (3) reports marginal effects on the probability of delinquency in percentage points, evaluated at the means of the regressors, with delta-method standard errors. The Age row of column (3) gives the total effect of one additional year, combining the linear and squared terms, so the squared term has no separate marginal effect. Marginal effects for the bank, region and year indicators are not reported, because a zero-to-one change in one indicator of a mutually exclusive set with the others held at their sample means is not a meaningful contrast. ***, **, and * denote significance at the 1, 5, and 10 percent levels.

Table A.2: Logit estimates of delinquency by loan type

	Mortgage credit	Micro credit	Micro loan	Auto credit	Credit card	Education credit	Consumer credit
ln(Income)	0.0070*	−0.0704***	−0.2444***	−0.0692***	−0.2009***	−0.0819***	−0.1867***
	(0.0038)	(0.0030)	(0.0015)	(0.0054)	(0.0027)	(0.0141)	(0.0117)
Age	−0.0227***	−0.0660***	−0.0434***	−0.0894***	−0.0686***	0.0739***	−0.0469***
	(0.0021)	(0.0025)	(0.0007)	(0.0025)	(0.0014)	(0.0141)	(0.0043)
Age^2	0.0002***	0.0006***	0.0003***	0.0008***	0.0007***	−0.0009***	0.0004***
	(0.00002)	(0.00003)	(0.00001)	(0.00003)	(0.00002)	(0.00021)	(0.00005)
ln(Loan amount)	−0.1594***	0.3989***	0.0714***	0.6044***	0.0012	−0.1853***	0.3201***
	(0.0069)	(0.0065)	(0.0012)	(0.0091)	(0.0013)	(0.0244)	(0.0159)
Interest rate	−0.0132***	0.0561***	−0.0145***	−0.0337***	−0.0121***	0.1099***	0.0266***
	(0.0008)	(0.0009)	(0.0001)	(0.0013)	(0.0001)	(0.0029)	(0.0011)
Maturity (months)	−0.0017***	0.0091***	0.0064***	0.0099***	−0.0019***	−0.0531***	−0.0018**
	(0.0001)	(0.0004)	(0.0001)	(0.0003)	(0.0001)	(0.0008)	(0.0009)
Number of loans	−0.0673***	−0.0204***	0.0184***	−0.0504***	−0.0066***	−0.0584***	−0.0284***
	(0.0017)	(0.0019)	(0.0005)	(0.0026)	(0.0008)	(0.0059)	(0.0047)
Gender (male=1)	0.0007	−0.0704***	0.2478***	−0.1607***	0.1944***	−0.3642***	0.0095
	(0.0061)	(0.0090)	(0.0029)	(0.0092)	(0.0052)	(0.0396)	(0.0161)
Co-borrower	−0.0351***	—	—	−0.3655***	—	−0.4901***	−0.4498***
	(0.0070)			(0.0218)		(0.0383)	(0.0453)
Preferential	−0.1677***	—	—	−0.4346***	—	2.9325***	—
	(0.0147)			(0.0269)		(0.1066)	
Online/Offline	0.3535***	0.0716*	0.4010***	1.3001***	—	−0.0878	—
	(0.0659)	(0.0380)	(0.0030)	(0.1088)		(0.1888)	
Bank, Region, and Year Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Logit coefficients, with standard errors clustered by borrower in parentheses, estimated separately for each loan type. The dependent variable equals one if the contract has overdue principal in the reporting month. Every model includes bank, region and year fixed effects. A dash indicates that the regressor is not part of that model's specification, because the corresponding feature does not vary within that product. ***, **, and * denote significance at the 1, 5, and 10 percent levels.

Table A.3: Standardised logit estimates of delinquency by loan type

	Mortgage credit	Micro credit	Micro loan	Auto credit	Credit card	Education credit	Consumer credit
ln(Income)	0.0069*	−0.0982***	−0.2134***	−0.0566***	−0.1818***	−0.0751***	−0.1700***
	(0.0037)	(0.0041)	(0.0013)	(0.0044)	(0.0024)	(0.0130)	(0.0106)
Age	−0.2070***	−0.7217***	−0.4863***	−0.9529***	−0.7411***	0.5623***	−0.5940***
	(0.0195)	(0.0272)	(0.0083)	(0.0265)	(0.0150)	(0.1070)	(0.0540)
Age^2	0.1857***	0.5437***	0.2693***	0.7776***	0.6123***	−0.4496***	0.5150***
	(0.0192)	(0.0278)	(0.0085)	(0.0271)	(0.0152)	(0.1036)	(0.0539)
ln(Loan amount)	−0.0875***	0.5783***	0.0777***	0.3135***	0.0018	−0.0953***	0.2329***
	(0.0038)	(0.0094)	(0.0013)	(0.0047)	(0.0019)	(0.0125)	(0.0115)
Interest rate	−0.0712***	0.7287***	−0.1606***	−0.1907***	−0.2672***	0.6979***	0.3060***
	(0.0045)	(0.0123)	(0.0014)	(0.0072)	(0.0031)	(0.0182)	(0.0121)
Maturity (months)	−0.0712***	0.1345***	0.0727***	0.1472***	−0.0421***	−1.0261***	−0.0285**
	(0.0041)	(0.0065)	(0.0013)	(0.0045)	(0.0021)	(0.0160)	(0.0140)
Number of loans	−0.1262***	−0.0505***	0.0657***	−0.0935***	−0.0210***	−0.1147***	−0.0590***
	(0.0032)	(0.0048)	(0.0018)	(0.0048)	(0.0025)	(0.0116)	(0.0098)
Gender (male=1)	0.0007	−0.0704***	0.2478***	−0.1607***	0.1944***	−0.3642***	0.0095
	(0.0061)	(0.0090)	(0.0029)	(0.0092)	(0.0052)	(0.0396)	(0.0161)
Co-borrower	−0.0351***	—	—	−0.3655***	—	−0.4901***	−0.4498***
	(0.0070)			(0.0218)		(0.0383)	(0.0453)
Preferential	−0.1677***	—	—	−0.4346***	—	2.9325***	—
	(0.0147)			(0.0269)		(0.1066)	
Online/Offline	0.3535***	0.0716*	0.4010***	1.3001***	—	−0.0878	—
	(0.0659)	(0.0380)	(0.0030)	(0.1088)		(0.1888)	
Bank, Region, and Year Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Same models as in Table A.2. Continuous regressors are rescaled to units of one standard deviation computed within each loan type, so a standardised coefficient measures how much of the within-product variation in delinquency that regressor accounts for; it is not comparable across products as a measure of economic magnitude. The binary regressors are not rescaled and are therefore not comparable with the rescaled continuous coefficients. Every model includes bank, region and year fixed effects. ***, **, and * denote significance at the 1, 5, and 10 percent levels.

Table A.4: Marginal effects at the means, full sample and by loan type (percentage points)

	Full sample	Mortgage credit	Micro credit	Micro loan	Auto credit	Credit card	Education credit	Consumer credit
ln(Income)	−1.652***	0.072*	−0.510***	−2.514***	−0.381***	−2.491***	−0.034***	−1.626***
Age	−0.151***	−0.042***	−0.147***	−0.223***	−0.118***	−0.211***	0.010***	−0.091***
ln(Loan amount)	0.560***	−1.643***	2.888***	0.734***	3.326***	0.015	−0.077***	2.788***
Interest rate	0.094***	−0.136***	0.406***	−0.149***	−0.185***	−0.150***	0.046***	0.232***
Maturity (months)	−0.015***	−0.017***	0.066***	0.066***	0.054***	−0.024***	−0.022***	−0.016**
Number of loans	0.115***	−0.693***	−0.148***	0.189***	−0.277***	−0.082***	−0.024***	−0.248***
Gender (male=1)	2.571***	0.007	−0.515***	2.497***	−0.914***	2.367***	−0.139***	0.083
Co-borrower	−2.096***	−0.360***	—	—	−1.753***	—	−0.183***	−3.364***
Preferential	5.962***	−1.815***	—	—	−2.054***	—	0.402***	—
Online/Offline	4.427***	4.165***	0.523*	3.933***	12.703***	—	−0.035	—

Notes: Marginal effects on the probability of delinquency, in percentage points, evaluated at the means of the regressors. The first column reports the pooled model estimated over all seven loan types; the remaining columns report models estimated separately by loan type. Every model includes bank, region and year fixed effects. For a continuous regressor, the effect is per one unit of that variable; for a binary regressor it is the discrete change from zero to one with every other regressor held at its mean. Income and loan amount enter in logarithms, so a 10 per cent increase corresponds to $\ln(1.1) \approx 0.095$ of the reported effect. The Age row gives the total effect of one additional year, combining the linear and squared terms; the squared term has no interpretation on its own and is therefore omitted. A dash indicates that the regressor is not part of that model's specification, because the corresponding feature does not vary within that product. Marginal effects are evaluated at each model's own mean vector, the predicted delinquency probability at which they are computed differs across columns, and an identical log-odds coefficient therefore implies a different change in absolute risk from one product to the next. Significance is assessed on delta-method standard errors clustered by borrower: ***, **, and * denote significance at the 1, 5, and 10 per cent levels.

Table A.5: Logit estimates of delinquency, 10 percent borrower subsample as a robustness check

	Coefficient	Standardised coefficient	Marginal effect (pp)
ln(Income)	−0.1583***	−0.1479***	−1.6283***
	(0.0034)	(0.0032)	(0.0352)
Age	−0.0307***	−0.3411***	−0.1495***
	(0.0019)	(0.0208)	(0.0035)
Age^2	0.000213***	0.1977***	—
	(0.000023)	(0.0209)	
ln(Loan amount)	0.0509***	0.0757***	0.5239***
	(0.0029)	(0.0044)	(0.0302)
Interest rate	0.0079***	0.1129***	0.0809***
	(0.0005)	(0.0067)	(0.0048)
Maturity (months)	−0.001210***	−0.0707***	−0.0124***
	(0.000100)	(0.0059)	(0.0010)
Number of loans	0.0098***	0.0324***	0.1003***
	(0.0015)	(0.0051)	(0.0158)
Gender(male=1)	0.2539***	0.2539***	2.5571***
	(0.0074)	(0.0074)	(0.0728)
Co-borrower	−0.2160***	−0.2160***	−2.0607***
	(0.0182)	(0.0182)	(0.1608)
Preferential	0.4650***	0.4650***	5.3267***
	(0.0187)	(0.0187)	(0.2365)
Online/Offline	0.4344***	0.4344***	4.4502***
	(0.0083)	(0.0083)	(0.0844)
Bank, region and year FE	Yes	Yes	Yes
Observations	14,057,278	14,057,278	14,057,278
Pseudo R^2	0.05	0.05	

Notes: Robustness check. Pooled logit estimates over all seven loan types, estimated on a random 10 percent subsample of borrowers. Each selected borrower enters with all of his or her monthly observations. The specification is identical to Table A.1. Standard errors clustered by borrower are in parentheses. The dependent variable equals one if the contract has overdue principal in the reporting month. Column (1) reports raw coefficients; column (2) rescales the continuous regressors to units of one standard deviation, leaving the binary regressors unchanged; column (3) reports marginal effects on the probability of delinquency in percentage points, evaluated at the means of the regressors, with delta-method standard errors. The Age row of column (3) gives the total effect of one additional year, combining the linear and squared terms, so the squared term has no separate marginal effect. Bank, region and year fixed effects are included but not reported. ***, **, and * denote significance at the 1, 5, and 10 percent levels.

Table A.6: Logit estimates of delinquency with an age–gender interaction, full sample

	Coefficient	Standardised coefficient	Marginal effect (pp)
ln(Income)	−0.1597***	−0.1490***	−1.6173***
	(0.0011)	(0.0010)	(0.0111)
Age	−0.0229***	−0.2541***	−0.1457***
	(0.0006)	(0.0068)	(0.0010)
Age^2	0.000172***	0.1592***	—
	(0.000007)	(0.0067)	
ln(Loan amount)	0.0538***	0.0798***	0.5445***
	(0.0009)	(0.0014)	(0.0094)
Interest rate	0.0089***	0.1280***	0.0904***
	(0.0001)	(0.0021)	(0.0015)
Maturity (months)	−0.001405***	−0.0819***	−0.0142***
	(0.000032)	(0.0018)	(0.0003)
Number of loans	0.0105***	0.0351***	0.1067***
	(0.0005)	(0.0016)	(0.0049)
Gender(male=1)	0.5332***	0.5332***	2.4996***
	(0.0080)	(0.0080)	(0.0229)
Co-borrower	−0.2184***	−0.2184***	−2.0480***
	(0.0058)	(0.0058)	(0.0500)
Preferential	0.5105***	0.5105***	5.8251***
	(0.0059)	(0.0059)	(0.0746)
Online/Offline	0.4283***	0.4283***	4.3186***
	(0.0026)	(0.0026)	(0.0265)
Interaction term			
Age × Male	−0.0074***	−0.1467***	—
	(0.0002)	(0.0039)	
Bank, region and year FE	Yes	Yes	Yes
Observations	140,630,546	140,630,546	140,630,546
Pseudo R^2	0.05	0.05	

Notes: Pooled logit estimates over all seven loan types, with standard errors clustered by borrower in parentheses. The specification is identical to Table A.1 with a single interaction, age with gender, added. The dependent variable equals one if the contract has overdue principal in the reporting month. Column (1) reports raw coefficients; column (2) rescales the continuous regressors to units of one standard deviation, leaving the binary regressors unchanged; column (3) reports total marginal effects on the probability of delinquency in percentage points, evaluated at the means of the regressors, with delta-method standard errors, summing every term in which a variable appears. The Age and Male coefficients in columns (1) and (2) are the effect of each at the other's zero and are not interpretable as effects at the mean; the interaction is read through column (3) and through the conditional slopes discussed in the text. The Age^2 and interaction rows have no separate marginal effect of their own and are left blank. ***, **, and * denote significance at the 1, 5, and 10 percent levels.

Table A.7: Logit estimates of delinquency with an income–loan-size interaction, full sample

	Coefficient	Standardised coefficient	Marginal effect (pp)
ln(Income)	−0.8150***	−0.7602***	−1.5883***
	(0.0098)	(0.0091)	(0.0110)
Age	−0.0278***	−0.3087***	−0.1416***
	(0.0006)	(0.0066)	(0.0010)
Age^2	0.00018***	0.1671***	—
	(0.000007)	(0.0067)	
ln(Loan amount)	−0.5485***	−0.8139***	0.5317***
	(0.0090)	(0.0133)	(0.0092)
Interest rate	0.0094***	0.1353***	0.0946***
	(0.0001)	(0.0021)	(0.0015)
Maturity (months)	−0.001359***	−0.0793***	−0.0136***
	(0.000031)	(0.0018)	(0.0003)
Number of loans	0.0134***	0.0446***	0.1344***
	(0.0005)	(0.0016)	(0.0048)
Gender (male=1)	0.2564***	0.2564***	2.5162***
	(0.0023)	(0.0023)	(0.0225)
Co-borrower	−0.2168***	−0.2168***	−2.0138***
	(0.0058)	(0.0058)	(0.0494)
Preferential	0.5229***	0.5229***	5.9274***
	(0.0058)	(0.0058)	(0.0740)
Online/Offline	0.4373***	0.4373***	4.3663***
	(0.0026)	(0.0026)	(0.0263)
Interaction term			
ln(Income) × ln(Loan amount)	0.0399***	1.2319***	—
	(0.0006)	(0.0182)	
Bank, region and year FE	Yes	Yes	Yes
Observations	140,630,546	140,630,546	140,630,546
Pseudo R^2	0.05	0.05	

Notes: Pooled logit estimates over all seven loan types, with standard errors clustered by borrower in parentheses. The specification is identical to Table A.1 with a single interaction, log income with log loan amount, added. The dependent variable equals one if the contract has overdue principal in the reporting month. Column (1) reports raw coefficients; column (2) rescales the continuous regressors to units of one standard deviation, leaving the binary regressors unchanged; column (3) reports total marginal effects on the probability of delinquency in percentage points, evaluated at the means of the regressors, with delta-method standard errors, summing every term in which a variable appears. The ln(Income) and ln(Loan amount) coefficients in columns (1) and (2) are the effect of each when the other equals zero, which lies far outside the range of the data, and are not interpretable as effects at the mean; the interaction is read through column (3) and through the conditional slopes discussed in the text. The Age^2 and interaction rows have no separate marginal effect of their own and are left blank. ***, **, and * denote significance at the 1, 5, and 10 percent levels.