DISCUSSION PAPER SERIES

DP13468

CREDIT AND INCOME

Manthos Delis, Fulvia Fringuellotti and Steven Ongena

FINANCIAL ECONOMICS

ISSN 0265-8003

CREDIT AND INCOME

Manthos Delis, Fulvia Fringuellotti and Steven Ongena

Discussion Paper DP13468
Published 20 January 2019
Submitted 12 January 2019

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

CREDIT AND INCOME

Abstract

Using a unique data set of business loan applications to a single bank from individuals who are majority owners of small firms, we study how bank credit origination or denial affects individuals' income. The bank cutoff rule based on the applicants' credit score creates a sharp discontinuity in the decision to originate loans or not. We show that loan origination increases recipients' income five years onward by more than 10% compared to denied applicants. The effect is more pronounced in rural and low-income areas. Our results suggest an important role for banks` credit decisions on the distribution of income.

JEL Classification: G21, D31, E24

Keywords: credit constraints, Income, Business loans, Income inequality, regression discontinuity design

Manthos Delis - m.delis@montpellier-bs.com
Montpellier Business School

Fulvia Fringuellotti - fulvia.fringuellotti@bf.uzh.ch
University of Zurich

Steven Ongena - steven.ongena@bf.uzh.ch
University of Zurich, Swiss Finance Institute, KU Leuven and CEPR and CEPR

Acknowledgements
We are grateful to participants of the CEPR's Endless Summer Conference on Financial Intermediation and Corporate Finance. We also thank seminar participants in the Essex Business School (University of Essex), Adam Smith Business School (University of Glasgow), Montpellier Business School, and the Athens University of Economics and Business. For comments, we thank Tobias Berg, Jerry Coakley, Yota Deli, Pedro Gete, Jens Hagendorff, Iftekhar Hasan, Olivier De Jonghe, Sotirios Kokas, Alexandros Kontonikas, Michael Lamla, Georgios Panos, Andrea Presbitero, Simon Price, and Farzad Saidi. Ongena acknowledges financial support from ERC ADG 2016 - GA 740272 lending.

Credit and Income

Manthos Delis
Montpellier Business School

Fulvia Fringuellotti
University of Zurich and *Swiss Finance Institute*

Steven Ongena
University of Zurich, Swiss Finance Institute, KU Leuven and *CEPR*

Delis: Montpellier Business School-Montpellier Research in Management, Email: m.delis@montpellier-bs.com; Fringuellotti; Department of Banking and Finance, University of Zurich, CH-8032, Zurich, Email: fulvia.fringuellotti@bf.uzh.ch; Ongena: Department of Banking and Finance, University of Zurich, CH-8032, Zurich, Email: steven.ongena@bf.uzh.ch. We are grateful to participants of the CEPR's Endless Summer Conference on Financial Intermediation and Corporate Finance. We also thank seminar participants in the Essex Business School (University of Essex), Adam Smith Business School (University of Glasgow), Montpellier Business School, and the Athens University of Economics and Business. For comments, we thank Tobias Berg, Jerry Coakley, Yota Deli, Pedro Gete, Jens Hagendorff, Iftekhar Hasan, Olivier De Jonghe, Sotirios Kokas, Alexandros Kontonikas, Michael Lamla, Georgios Panos, Andrea Presbitero, Simon Price, and Farzad Saidi. Ongena acknowledges financial support from ERC ADG 2016 - GA 740272 lending.

Credit and Income

Abstract

Using a unique data set of business loan applications to a single bank from individuals who are majority owners of small firms, we study how bank credit origination or denial affects individuals' income. The bank cutoff rule based on the applicants' credit score creates a sharp discontinuity in the decision to originate loans or not. We show that loan origination increases recipients' income five years onward by more than 10% compared to denied applicants. The effect is more pronounced in rural and low-income areas. Our results suggest an important role for banks` credit decisions on the distribution of income. (99 words)

Keywords: Credit constraints; Income; Business loans; Income inequality; Regression discontinuity design
JEL classification: G21; D31; E24; O15

I. Introduction

Over past decades, the gap between the rich and the poor has risen in most OECD countries (OECD, 2015), yielding a lively debate on the sources of this development and the proper measures to contain the problem. The role of finance is at the forefront of the relevant academic literature (e.g., Greenwood and Jovanovic, 1990; Galor and Zeira, 1993; Demirguc-Kunt and Levine, 2009; Beck et al., 2010). This study aims to identify and quantify how banks` credit decisions (credit origination or denial) affects applicants' future income. The findings have important implications for the relation between credit and individuals' income, and reflect on how credit origination or constraints affect the distribution of income.

Theoretically, asymmetric information between lenders and borrowers affects credit availability. Because the enforcement of loan contracts is imperfect, lenders often require borrowers to pledge collateral. Lenders also ration credit based on an expected probability of repayment. In general, credit expansion accompanies a relaxation of credit constraints, leading to more financing opportunities for the full spectrum of potential borrowers (including the poor) and a possible tightening of the income distribution (Banerjee and Newman, 1993; Galor and Zeira, 1993).

However, credit-constrained individuals often have less wealth, and their exclusion from credit can foster persistent income growth and inequality. More specifically, financial frictions in the form of informational asymmetry imply an important role for wealth (or capital) endowment in liquidity creation. The endowment represents a fixed cost for credit access. The relatively poor cannot always overcome it, irrespective of the quality of their investment ideas, due to adverse selection and moral hazard in the loan origination process. Thus, returns on capital can lead to high persistence in income growth only for those with substantial wealth (Piketty, 1997; Mookherjee and Ray, 2003; Demirguc-Kunt and Levine, 2009). Further, returns on investment usually increase with the amount of capital wealthier individuals employ, initiating a second-order effect due to economies of scale in larger projects (e.g., Evans and Jovanovic, 1989; Greenwood and Jovanovic, 1990).

The existence of a causal link between access to credit and income inequality presupposes that banks' credit decisions (positive or negative) and the associated access (or lack thereof) to credit have a direct effect on individuals' income. Take, for example, two individuals with approximately the same income and credit quality. One gets a new business loan approved; the other does not. If loan origination implies an increase in the income of the former relative to the latter, then credit affects the income distribution.

A simple plot between GDP per capita (or the Gini coefficient) and the ratio of private credit to GDP for 150 countries over 1960-2015, shows that income (income inequality) is strongly and positively (negatively) correlated to private credit from banks and other financial institutions over GDP (Figure 1). Of course, this relation cannot be interpreted as causal; it is confounded by reverse causality, meaning that income inequality may actually drive credit expansion (Kumhof and Ranciere, 2010; Rajan, 2010) and/or omitted-variable bias due to factors jointly affecting the distribution of income and the degree of financial depth, which are difficult to measure (e.g., the availability of new investment ideas).

[Insert Figure 1 here]

Our study provides the first empirical analysis of how access to credit affects individuals' income by comparing the future incomes of accepted applicants to those of rejected applicants. We identify this effect using a unique data set of business loan applications to a single European bank. Our focus is on loan applications from small and micro enterprises that are majority-owned by individuals who have an exclusive relationship with the bank (i.e., they do not obtain credit from other regulated commercial banks). For these applicants, the bank has information on the business owners' incomes and decides whether to grant the loans based on a credit score cutoff rule. Specifically, each applicant receives a credit score at the time of the loan application. Credit is granted to applicants whose credit scores are above the cutoff, and denied otherwise.

The uniqueness of our data lies in the available information on the majority owners, which encompasses income, wealth, and the credit scores assigned by the bank, as well as other applicant and firm characteristics. Importantly, the exclusivity of the relationship between the bank and the applicant means that most applicants (accepted and rejected) reapply for loans. This in turn means the bank maintains information on applicants' income after the original credit decision.

The availability of credit score and future applicants' income is crucial for our identification strategy because it allows us to exploit the cutoff rule as a source of exogenous variation in the credit decision. Our approach builds on the idea that individuals whose credit scores are around the cutoff are virtually the same in terms of credit quality, yielding a regression discontinuity design (RDD). This implies identification from comparing changes in the income of accepted and denied applicants, who prior to the bank's credit decision have similar credit scores (including similar incomes).

We show that a loan origination increases the recipient's income five years onward by more than 10% compared to denied applicants, regardless of whether we control for application probability. The economic interpretation of this finding is that marginally accepted applicants benefit from an approximately 10% increase in their incomes compared to marginally rejected applicants, thereby significantly widening income inequality between the two groups. This finding is robust to several re-specifications and is not affected by the mix of the control variables. Further, the RDD passes the tests for credit score manipulation, and the control variables are continuous around the cutoff.

We further relate our finding to income inequality by calculating inequality measures (Gini coefficients and Theil indices) for the loan applicants around the cutoff. We show that the Gini and Theil indices increase (wider income distribution) for the sample of individuals five years after the credit decision compared to the year of the credit decision. Using the same inequality measures, we also document tighter income distribution among accepted applicants and wider income distribution among rejected applicants. These findings are consistent with the theory of a negative nexus between finance and inequality when there is access to credit (Greenwood and Jovanovic, 1990).

We also examine the heterogeneity of our findings in interesting subsamples reflecting additional aspects of how credit affects income and its distribution. We first document stronger effects in low-income regions compared to high-income regions. This suggests that a bank's credit decision is even more important for an applicant's future income in low-income regions, thus potentially affecting income distribution within and across regions. Second, we use the Great Recession to examine how an economic crisis and associated credit crunch affect the credit-income relation. The identified effect is somewhat stronger during the crisis period, in line with the premise that a credit crunch causes more harm to people with lower credit scores.

From an empirical viewpoint, our study relates to the literature that looks broadly at how financial development and/or credit constraints affect income distribution by relying on aggregate (at the country or regional level) measures of inequality (mostly the Gini index) and financial development. This body of literature provides mixed results. Clarke et al. (2006), Beck et al. (2010), Kappel (2010), Hamori and Hashiguchi (2012), Delis et al. (2014), and Naceur and Zhang

(2016), for example, document a negative relation between financial development and income inequality, consistent with the idea that credit expansion implies relaxed credit constraints. Denk and Cournède (2015), Jauch and Watzka (2016), and de Haan and Sturm (2017), however, point instead to a negative relation, suggesting that financial development improves access to credit only for the rich. Our paper also relates to several studies on financial development and inequality (for a thorough review, see Demirguc-Kunt and Levine, 2009). We contribute to this literature documenting the effect of credit origination on income and income inequality at the individual, micro level.

Another strand of related recent literature examines how credit constraints affect economic outcomes using data on loan applications (such as ours). Berg (2018), for example, shows that credit denial has stronger negative real effects on low-liquidity firms, which need to increase cash holdings and dispose of other assets in response to a loan rejection. A broader body of literature documents how financial constraints affect the transmission of a credit shock due to changes in monetary policy (Gertler and Gilchrist, 1994; Kashyap and Stein, 2000; Jiménez et al. 2012) or bank stability (Gan, 2007; Duchin et al., 2010; Chodorow-Reich, 2014; Balduzzi, Brancati, and Schiantarelli, 2017; Cingano, Manaresi, and Sette, 2013; Bentolila et al., 2017; Acharya et al., Forthcoming; Popov and Rocholl, forthcoming).

The next section describes the data set and empirical identification, emphasizing the particular RDD. Section III presents the empirical results regarding how bank credit decisions affect loan applicants' income; it also links these effects to income distribution. Section IV concludes the paper.

II. Data and Empirical Identification

A. Loan Applications

We use a unique sample of loan applications to a single European bank. The bank provides credit to a wide array of small and large firms, as well as to consumers, households, and the public sector. We use only loan applications from small and micro enterprises that are majority-owned by specific individuals, for which the bank has important information for our analysis.[1] Specifically, we have information on whether the loan is originated or denied, as well as loan characteristics, firm characteristics, and applicant characteristics. For originated loans, loan characteristics include the amount, maturity, collateral, and other features (covenants, performance-pricing provisions). Firm characteristics encompass several accounting variables, such as assets and sales, profits, leverage, etc.

What makes this data unique is information on the applicant (the firm's majority owner). The applicant characteristics include income, assets (wealth), gender, education, relationship with the bank (an exclusive relationship or not), and the credit score assigned by the bank. For two reasons, we focus on loan applications from individuals who have exclusive relationships with their banks. First, the bank has income information for these applicants for several years before and after the loan origination. Second, these applicants are generally unable to obtain credit from another bank, especially if their applications are denied; moreover, they cannot access capital markets due the firm's small size.[2]

Each applicant is given a credit score at the time of the application, and this score is the decisive factor in loan origination. For comparative purposes, we normalize the credit score to be

[1] Using the European Commission's definition, a small enterprise has total assets less than €10 million; a micro enterprise less than €2 million in assets.
[2] We have information about this exclusivity from the bank. However, the firms can receive credit (obviously at higher rates) in the shadow-banking sector.

around the cutoff value of 0. The bank originates the loan if the credit score is higher than 0 and denies the loan otherwise. For very few applications (72 cases), this criterion does not hold. These exceptions are possibly due to data-entry mistakes and thus we disregard them in our analysis. We explicitly define the credit score along with all the variables used in our empirical analysis in Table 1 and provide summary statistics in Table 2.

[Insert Tables 1 & 2 about here]

Using this information, we generate a balanced panel data set, where applicants are the cross-sectional unit of the panel and years 2002-2016 are the time unit. For each applicant, we know his/her income and wealth over the full sample period, as well as for at least five years before and after the loan application. This means that the individuals in our sample do not necessarily apply for loans in some years. This sample also includes information for the rest of the applicant and firm characteristics defined in Table 1. This stringent cleansing process yields 234,420 observations corresponding to 15,628 individual applicants over 2002-2016.[3] In this panel, there are 61,863 loan applications (the sample in the majority of our empirical tests). We report summary statistics for the variables in Table 2.

The mean future income (respectively, in one year, three years, and five years) tends to rise over time for loan applicants. Banks accept (or partially accept) approximately 87% of loan applications and reject 13%. This rejection rate is a bit higher than the rejection rates reported in the European Commission/European Central Bank Survey on access to finance for enterprises (SAFE).[4] The reason is that some missing observations on variables in our empirical analysis correspond to individuals with strong bank ties (i.e., individuals for whom the bank already has

---

[3] The actual number of loan applications from small and micro enterprises, including business-loan applications from individuals who have nonexclusive relationships with the bank, as well as those from applicants for which we lack dynamic income information, is 513,525.
[4] See, https://ec.europa.eu/growth/content/survey-access-finance-enterprises-safe-was-published-today_ga.

information) who are usually not rejected. If anything, this biases our results in favor of denied applicants. However, our identification approach, based on individuals around the credit score cutoff, should mitigate such concern. After its transformation, the mean credit score is positive and equal to approximately 0.1. Average loan duration is roughly three years.

Summary statistics for our control variables show that the mean applicant has tertiary education and total wealth of €187,200 (see Table 2). The mean firm size (total assets) is €369,500, and mean firm leverage is 20.7%, which is comparable to European averages (e.g., Carvalho, 2017). Overall, the summary statistics show that our data set is consistent with the mean value of our variables at the European level.

B. Empirical Identification

Three important features of our data set are the availability of information about (i) originated and denied loans, (ii) the exclusivity of the relationship between loan applicants and banks, and (iii) applicants' income before and after the loan application. Based on these features, a standard identification method would compare the incomes of approved applicants (the treated group) with the incomes of rejected applicants (the control group) before and after the loan decision. Unfortunately, the treatment here is endogenous to several factors behind the bank's decision to grant the loan, making a differences-in-differences exercise far from optimal.

The fourth and most important feature of our data set for identification purposes is the availability of information on credit scores and the perfect correlation of the scores above the cutoff with loan origination.[5] This implies a sharp discontinuity in treatment as a function of credit score.[6]

[5] This is after dropping the 72 exceptions due to data entry errors.
[6] Berg (2018) exploits a similar type of discontinuity to investigate how loan rejection affects firms' cash holdings.

Therefore, we rely on a sharp RDD using credit score as the assignment (also referred to as "the running" or "the forcing") variable (Imbens and Lemieux, 2008; Lee and Lemieux, 2010).

Assuming that the relation between access to credit and income is linear, the simplest form of the RDD is:

$$y_{i,t+n} = a_0 + a_1 D_{it} + a_2 (x_{it} - \bar{x}) + u_{it}. \quad (1)$$

In equation 1, y is applicant's i income in the n^{th} year ahead of the loan application, which takes place in year t. D is a binary variable that equals 1 if the credit score is above the cutoff and zero otherwise, which determines whether the loan is granted. Thus, a_1 is the treatment effect. Also, $x_{it} - \bar{x}$ is the distance between the cutoff and applicant i's credit score given at the time of the loan application.

The distribution of applicant's income depicted in Figure 2 exhibits a regular shape. The main assumption for the validity of this model, similar to any other RDD, is that applicants cannot precisely manipulate their credit scores. If applicants, even while having some influence, are unable to manipulate their credit scores precisely, the variation in treatment around the cutoff provides a randomized experiment. The lack of precise manipulation is the most compelling requirement of the RDD vis-à-vis other identification methods, such as differences-in-differences or instrumental variables (Lee and Lemieux, 2010).

[Insert Figure 2 about here]

Theoretically, precise manipulation is unlikely, as loans officers' prudent behavior should prevent applicants from having exact information on their credit scores. We demonstrate, through a specific statistical test, that this is also unlikely from an econometric viewpoint. Specifically, we test for manipulation of the assignment variable around the cutoff. Self-selection or nonrandom sorting of applicants would entail a discontinuous change in the distribution of the credit score.

Figure 2 shows that the probability density of the credit score does not jump around the cutoff. In line with the graphical evidence, the formal test of Cattaneo et al. (2018) confirms there is no statistical evidence of manipulation of the forcing variable (see Table 3 and Figure 3).

[Insert Table 3 & Figure 3 about here]

III. Empirical Results

A. Parametric Model

We first consider estimating equation (1) with a parametric model (OLS). We use clustered standard errors at the individual level to ensure robust inference. To allow for a differential effect on the two sides of the cutoff, we include the interaction $D_{it}(x_{it} - \bar{x})$, so that equation (1) becomes:

$$y_{i,t+n} = a_0 + a_1 D_{it} + a_2(x_{it} - \bar{x}) + a_3 D_{it}(x_{it} - \bar{x}) + u_{it}. \quad (2)$$

The coefficient of interest is a_1, which is the coefficient of the acceptance dummy *Granted*, which captures the treatment effect.

Table 4 reports the results. Specifications 1-3 use as a dependent variable the applicants' income one year ahead, three years ahead, and five years ahead of the loan application. We find a positive and statistically significant coefficient on *Granted* in all three specifications. The magnitude of this coefficient suggests a 5.1% increase in the incomes of approved applicants one year ahead of loan origination (column 1), a 7.3% increase three years ahead (column 2), and a 7% increase five years ahead (column 3). Also, the coefficient of the interaction between *Granted* and *Credit Score* is negative and statistically significant three and five years after loan origination, confirming our prior differential effect on the two sides of the cutoff.

[Insert Table 4 about here]

In specifications 4-6, we introduce the set of loan, firm, and applicant controls variables. Loan controls include the requested amount (*Loan amount*) and loan maturity (*Maturity*). Firm variables include total assets (*Firm size*) and leverage ratio (*Leverage*). Applicant controls include degree of education (*Education*) and income one year before the application (*Income t-1*). We provide thorough definitions for these variables in Table 1.

Indeed, the results are similar to those in the first three columns and, if anything, slightly strengthen. Being approved for a loan implies an increase in applicant income by 5.4% one year after of the loan decision (column 4), by 7.5% three years after (column 5), and by 7.2% five years after (column 6). Looking at the covariates, most are not statistically significant. This is not surprising, as many of them concur in determining the credit score. Nevertheless, we find a positive and statistically significant coefficient for *Income t-1*, suggesting persistence in the outcome variable. *Leverage* has a positive and significant coefficient, but it is largely collinear with the credit score.[7] We also find a positive coefficient on *Maturity*, although it is significant only in column 4. These results remain unchanged if we add industry, loan type, and year fixed effects to our specifications (results in Table A1 of the Appendix).

B. Local Linear Regression

The linear model identifies the treatment effect placing equal weight on all information available in the sample. This suggests a potential bias, as it treats observations far from the cutoff in the same way as observations close to the cutoff, and the treatment effect is estimated using two groups

[7] Our analysis focuses on firms able to raise external funds only by borrowing from the bank under study. In our specifications, we control for the leverage ratio observed in the year of the loan decision. The cutoff rule implies that applicants whose credit scores are above the cutoff are approved for a loan. As a consequence, leverage ratios increase in the year of the loan origination (see Figure 5). This explains why our covariate is to a large extent collinear with the credit score.

of individuals that might not be comparable. To handle this issue, we use a local linear regression (for a general description, see Imbens and Lemieux, 2008, and Calonico et al., 2014). The main advantage of this approach is the assignment of higher weights as we move closer to the cutoff (using a kernel smoother). We determine the optimal bandwidth using the approach in Calonico et al. (2014), and for efficient estimation we mainly base our inference on the local-quadratic bias-correction in Calonico et al. (2018).

Table 5 reports the estimates of the average treatment effect for our set of local linear regressions.[8] For each specification, we report the conventional RD estimates with conventional variance estimator (*Conventional*), the bias-corrected RD estimates with conventional variance estimator (*Bias-corrected*), and the bias-corrected RD estimates with robust variance estimator (*Robust*).

Regardless of whether we include (in columns 1-3) or do not include (in columns 4-6) the set of controls, we find that granting a loan has a positive and significant effect on an applicant's future income. Relying on *Robust* estimates for inference, we find an income increase of approximately 6% among approved applicants one year or three years after of the loan origination, and an increase of approximately 11% five years ahead.

Overall, the coefficients of *Granted* are comparable to those in the corresponding regressions of Table 4. The magnitudes of the effect are somewhat higher than those of the OLS regressions, especially considering the effect five years ahead. Given the small discrepancy in the results between the parametric and nonparametric RDD and the advantages of the nonparametric

[8] The average treatment effect here is the counterpart of the coefficient of the acceptance dummy in equation (2). It is nonparametrically identified as $\tau_{RD} \lim_{x \to \bar{x}^+} \mathbb{E}[y_{it}|x_{it} = x] - \lim_{x \to \bar{x}^-} \mathbb{E}[y_{it}|x_{it} = x]$.

RDD highlighted in the literature, we consider this method as our benchmark and we use it in most of our sensitivity tests (unless not applicable).

[Insert Table 5 about here]

An additional merit of the nonparametric RDD is the graphical inspection of the relation between access to credit and income that takes into account any potential nonlinearity. Figure 4 depicts applicants' income five years after the loan decision against the credit score (the figure is from column 3 of Table 5 and the effective observations used by the local linear regression). The figure shows a clear upward shift in applicants' income around the cutoff. This indicates that the treatment (loan origination) entails a sharp discontinuity in the outcome variable (income), corroborating our methodological approach.

[Insert Figure 4 about here]

On the use of control variables, a key assumption of the RDD is that the expectation of the outcome variable conditional on the assignment variable is continuous. This requires that the relation between the covariates and the credit score is smooth around the cutoff. A graphical inspection confirms that this condition is fulfilled (Figure 5). This means that our baseline model in equation (2) is well specified and, using the controls, will not significantly affect our main result.

[Insert Figure 5 about here]

Despite the advantage of focusing on observations close to the cutoff, the nonparametric approach does not necessarily represent the ideal functional form of the RDD. In light of that, Lee and Lemieux (2010) suggest relying on different bandwidth-selection methods to test if the results are stable across different specifications. Table A2 of the Appendix shows that the results remain unchanged when using the mean-squared error (MSE) or the common coverage error (CER) bandwidth selector. Also, Figure 6 shows that the significance of *Conventional* in model (3) is robust to different windows around the cutoff where (small-sample) inference is conducted.[9]

[Insert Figure 6 about here]

Overall, our analysis shows that credit decisions have real effects on income. Consider two applicants: the first has a credit score slightly above the cutoff; the second has a credit score slightly below the cutoff. At the time of the loan application, the credit quality of these two individuals is virtually the same. However, the cutoff rule implies that credit is granted only to the former. The increase in income experienced after loan origination documents a causal link between access to credit and income.

C. Robustness Tests

In principle, wealthier individuals should be able to maintain higher incomes over time through higher investment. Accordingly, part of the macro inequality literature highlights the role of initial GDP per capita and suggests controlling for some sort of historical (or initial) wealth conditions when estimating models of inequality (e.g., Li et al., 1998). To this end, we use individual wealth in the first year before the loan application in which this information is available (*Initial wealth;* see Table 1).

As with the rest of the control variables, we show in Figure A1 that *Initial wealth* is continuous around the cutoff. Of course, adding this variable to our covariates entails a substantial drop in the number of observations in the sample. This is the reason we leave this exercise as a robustness test. The nonparametric results in Table 6 show that including initial wealth does not yield significantly different results. If anything, the treatment effect is slightly stronger, with the

[9] Inference in Table 5 is based, instead, on large-sample approximations (Calonico et al., 2014).

only exception of the three-year horizon from the loan decision. We obtain similar patterns when using the parametric RDD (Table A3 of the Appendix).

[Insert Table 6 about here]

So far, our framework does not explicitly model the probability to apply for a loan in a specific year. Given that our sample is a balanced panel of bank customers with exclusive relationships and these customers appear in the panel irrespective of whether they apply for a loan in a given year, we can model the probability of receiving a loan application, and examine its effect in our baseline model. Econometrically, this implies limiting a form of selection bias in the estimation of the treatment effect.

Specifically, we use a parametric two-stage selection model as in e.g. Heckman (1976), Dass and Massa (2011), and Jimenez et al. (2014). In the first stage, we estimate the probability that a bank customer applies for a loan in a specific year (probit model). The right-hand side variables in the first stage are those in columns 4-6 of Tables 4 and 5, excluding the credit score (which is unknown to the applicant) and including *Gender*.[10] In the second stage, we run a similar regression to the one implied in equation (2), in which we use the predicted instantaneous probability of applying for a loan (hazard rate) from the first stage as an additional control variable.

Table 7 reports the estimation results. The first-stage results show that income and wealth positively and strongly affect the hazard rate of a loan application, in line with the premise that wealthier individuals are more likely to apply for credit. The same holds for larger and more leveraged firms. Interestingly, we also find that male applicants are 1.6% more likely to apply for credit than female applicants are. The second-stage results are fully in line with Table 4, even though the probability of loan application enters with a highly significant and positive coefficient.

[10] We find that *Gender* is significantly correlated with the probability to apply for a loan but does not explain income in the baseline specifications.

To account for selection of loan applicants, we prefer the standard parametric model because it is standard in the applied economics/finance literature, whereas the nonparametric models are quite rare in this respect. However, we do experiment with a semiparametric model, where we save the parametric first-stage prediction and include it in the nonparametric second stage. Again, the results, reported in Table 8, are consistent with those of Table 5.

[Insert Tables 7 & 8 about here]

D. Reflection on Income Inequality

A natural implication of our key finding is that income distribution changes. Specifically, we expect that a bank's credit decision increases income inequality between groups of individuals who have similar characteristics (individuals around the cutoff) but receive different credit decisions (accept vs. reject). It is difficult to extend this implication to the full array of income distribution, because most people (and certainly the rich) are granted loans. However, we can construct inequality measures around the cutoff for individual income at the time of loan application (t) and five years ahead (t+5). As our sample around the cutoff, we use individuals with credit scores less than the absolute value of 0.1.[11]

Panel A of Table 9 reports the results for the Gini coefficient and the Theil index. Both the indices increase from time t to time t+5, reflecting higher income inequality. The effect is economically large and equivalent to that identified in Table 5. Specifically, the Gini coefficient increases by approximately 9% and the Theil index increases by approximately 10%, indicating considerably higher income inequality after the bank credit decisions for the sample of individuals close to the cutoff.

[11] Alternatively, we use the effective observations left and right of the cutoff produced by the local linear regression in column 6 of Table 5. The results are very similar.

[Insert Table 9 about here]

In Panel B of Table 9, we construct equivalent Gini and Theil indices for accepted and rejected applicants. The indices show that for accepted applicants, the Gini and Theil indices are significantly lower, whereas for the rejected applicants they are higher. This is consistent with the premise that positive credit decisions allow individuals close to the cutoff to increase their incomes, thereby tightening the income distribution among accepted individuals. In contrast, negative credit decisions are consistent with widening income distribution among rejected individuals, who are the relatively poor.

We conduct two more tests to reflect how credit decisions affect income distribution. The first concerns the role of applicant location based on regional income, distinguishing between low-income regions and high-income regions. In Table 10, we replicate the analysis in columns 4-6 of Table 5, separating our full sample into low-income and high-income regions based on median income. We expect that the income elasticity to credit decisions is higher in low-income regions, where credit constraints should also be relatively higher.[12]

The results show that this is indeed the case. We find that five years after a bank's credit decision, accepted applicants have 12% higher incomes than rejected applicants in low-income regions. The equivalent effect in the high-income regions is 9%, indicating that the incomes of individuals in high-income regions are less affected by credit decisions compared to low-income regions (where credit constraints are higher). The 3% difference is already economically significant, but we expect it to be considerably stronger in countries with severe regional inequalities and credit constraints.

[Insert Table 10 about here]

[12] In our sample, the mean value of *Granted* in high-income regions is 0.880; it is 0.853 in the low-income regions.

As a final test, we consider the role of the Great Recession. During this period, Europe experienced sharp losses in household wealth and aggregate demand, substantial contraction of credit, and increased unemployment (e.g., IMF, 2009; ECB, 2016). In such context, entrepreneurs face riskier investment opportunities and lower profits. This yields increased dependence on bank credit, even for business survival and especially for small firms. If this prevails, loan origination has a stronger effect on applicant incomes during the crisis period, and a negative credit decision widens the income distribution.

To examine the role of the crisis in our results, we split the sample into the 2000-2008 and the 2009-2016 periods. We leave 2008 in the pre-crisis period because credit from banks in European countries was still rising that year. Similarly, we include the full period after the crisis because credit from banks to the private sector over GDP decreased in 2009-2016.[13]

Table 11 reports the results from the two samples. We find that three to five years after a bank's credit decision, access to credit has a stronger effect on applicant incomes during a crisis than in normal times. In particular, we find that approved applicants' incomes rose 9.6% five years ahead of the loan origination during 2000-2008 (column 6), versus 10.5% in the crisis and post-crisis periods (column 3). We conclude that, in the medium to long run, a loan origination has a stronger effect on applicant incomes during periods of higher credit constraints than in normal times.

[Insert Table 11 about here]

[13] See https://data.worldbank.org/indicator/FS.AST.PRVT.GD.ZS?locations=XC.

IV. Conclusions

Credit constraints potentially hinder income growth opportunities, especially for those with low incomes and a lack of collateral. Using data from business loan applications to a single European bank, we study and quantify how a bank's credit decisions (acceptance or rejection) affect individuals' future incomes. We look at loan applications from small and micro enterprises that are majority-owned by individuals for which we have detailed information on past and future income, the credit score assigned by the bank, and the exclusivity of relationship lending (among many other applicant and firm characteristics).

Our identification strategy comprises a regression discontinuity design, exploiting exogenous variation in the credit decision from the cutoff rule on the basis of credit score. Essentially, with this strategy we compare individuals with credit scores (and thus very similar characteristics guiding the credit decision) around the cutoff. We show that access to credit has a positive effect on individual income. Specifically, the income of accepted applicants is approximately 4% higher than the income of denied applicants one to three years ahead of the loan decision; this jumps to 10% five years ahead. This finding is robust to several re-specifications and robustness tests.

We also explore how income distribution changes with bank credit decisions. We first show that the Gini and Theil indices increase for individuals around the cutoff, reflecting increased income inequality within that sample. We also show that credit decisions have a somewhat stronger effect on applicants' future incomes in low-income regions (vs. high-income regions) and during the crisis and post-crisis period (vs. the pre-crisis period). These results highlight the heterogeneous effects credit availability has on applicants' future incomes due to increased credit constraints. They also highlight differential effects on income distribution.

Our findings have two key and interrelated economic implications. First, credit decisions strongly affect applicants' future income and its subsequent dynamics, altering lifetime income expectations and potentially applicants' economic decisions. Second, credit decisions exert substantial effects on income inequality between individuals who prior to the credit decision have similar credit scores. Importantly, these effects are more potent for applicants in low-income regions and during crisis and post-crisis periods.

Our findings suggest that an otherwise efficient credit decision affects income distribution and thus supports policy interventions aimed at increasing credit access. Relevant actions are those of the European Bank for Reconstruction and Development (EBRD) and the European Investment Bank (EIB), which selectively target credit-constrained individuals with good investment ideas.

Our findings also open up a discussion on whether central banks (via specialized institutions such as the EBRD and EIB) could direct a small part of the money-creation process to good investment ideas from loan applicants whom the banking system rejected due to a lack of credit history or collateral. We leave the thorough examination of the effects of these policies to future research.

References

Acharya, V., Eisert, T., Eufinger, C., Hirsch, C., Forthcoming. Real effects of the sovereign debt crisis in Europe: Evidence from syndicated loans. *Review of Financial Studies*.

Adams, R. B., Ferreira, D., 2009. Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics* 94, 291-309.

Balduzzi, P., Brancati, E., Schiantarelli, F., 2017. Financial markets, banks' cost of funding, and firms' decisions: lessons from two crises. *Journal of Financial Intermediation* 36, 1-15.

Banerjee, A. V., Newman, A. F., 1993. Occupational choice and the process of development. *Journal of Political Economy* 101 (2), 274-298.

Beck, T., Levine, T., Levkov, A., 2010. Big bad banks? The winners and losers from bank deregulation in the United States. *Journal of Finance* 65 (5), 1637-1667.

Becker, S. O., Egger, P. H., Von Ehrlich, M., 2013. Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy* 5 (4), 29-77.

Bentolila, S., Jansen, M., Jiminéz, G., Ruano, S., 2017. When credit dries up: Job losses in the great recession. *Journal of the European Economic Association* 16 (3), 650-695

Berg, T., 2018. Got rejected? Real effects of not getting a loan. *Review of Financial Studies* 31 (12), 4912-4957.

Calonico, S., Cattaneo, M. D., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* 82 (6), 2295-2326.

Calonico, S., Cattaneo, M. D., Farrell, M. H., 2018. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association* 0 (0), 1-13.

Carvalho, I., 2017. What is the impact of the access to external finance on the capital structure of SMEs in Europe? PhD Dissertation, Católica Lisbon School of Business & Economics.

Cattaneo, M. D., Titiunik, R., Vazquez-Bare, G., 2016. Inference in regression discontinuity design under local randomization. *Stata Journal* 16 (2), 331-367.

Cattaneo, M. D., Jansson, M., Ma, X., 2018. Manipulation testing based on density discontinuity. *Stata Journal* 18 (1), 234-261.

Charness, G., Gneezy, U., 2012. Strong evidence of gender differences in risk taking. *Journal of Economic Behavior and Organization* 83, 50-58.

Chodorow-Reich, G., 2014. The employment effects of credit market disruptions: Firm-level evidence from the 2008-2009 financial crisis. *Quarterly Journal of Economics* 129, 1–59.

Cingano, F., Manaresi, F., Sette., E., 2013. Does credit crunch investments down? *Review of Financial Studies* 29, 2737–73.

Clarke, G. R. G., Xu, L. C., Zou, H. F., 2006. Finance and income inequality: What do data tell us? *Southern Economic Journal* 72 (3), 578-596.

Dass, N., Massa, M., 2011. The impact of a strong bank-firm relationship on the borrowing firm. *Review of Financial Studies* 24, 1204-1260.

De Haan, J., Sturm, J. E., 2017. Finance and income inequality: A review and new evidence. *European Journal of Political Economy* 50, 171-195.

Delis, M. D., Hasan, I., Kazakis, P., 2014. Bank regulations and income inequality: Empirical evidence. *Review of Finance* 18, 1811-1846.

Denk, O., Cournède, B., 2015. Finance and income inequality in OECD countries. OECD Economic Department Working Papers No. 1224.

Duchin, R., Ozbas, O., Sensoy, B.A., 2010. Costly external finance, corporate investment, and the subprime mortgage crisis. *Journal of Financial Economics* 97, 418–35.

ECB, 2016. The household finance and consumption survey: Results from the second wave. ECB Statistics Paper No 18. December 2016. Household Finance and Consumption Network.

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

Eckel, C. C., Grossman, P. J., 2008. Men, women and risk aversion: Experimental evidence. Handbook of Experimental Economics Results, Volume I, Chapter 113, Elsevier.

Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. *Review of Economic Studies* 60 (1), 35-52.

Gan, J., 2007. The real effects of asset market bubbles: Loan- and firm-level evidence of a lending channel. *Review of Financial Studies* 20, 1941–73.

Gertler, M., Gilchrist., S., 1994. Monetary policy, business cycles, and the behavior of small manufacturing firms. *Quarterly Journal of Economics* 109, 309–40.

Greenwood, J., Jovanovic, B., 1990. Financial development, growth, and the distribution of income. *Journal of Political Economy* 98 (5), 1076-1107.

Hamori, S., Haschiguchi, Y., 2012. The effect of financial deepening on inequality: Some international evidence. *Journal of Asian Economics* 23, 353-359.

Heckman, J. J., 1976. The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement* 5, 475–492.

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

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

IMF, 2009. World economic outlook. Crisis and recovery. April 2009. World Economic and Financial Surveys.

Iyer, R., Peydró, J. L., da-Rocha-Lopes, S., Schoar, A., 2014. Interbank liquidity crunch and the firm credit crunch: evidence from the 2007-2009 crisis. *Review of Financial Studies* 27 (1), 347-372.

Jauch, S., Watzka, S., 2016. Financial development and income inequality: A panel data approach. *Empirical Economics* 51, 291-314.

Jianakoplos, N. A., Bernasek, A., 1998. Are women more risk averse? *Economic Inquiry* 36, 620-630.

Jiménez, G., Ongena, S., Peydró, J. L., Saurina, J., 2012. Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications. *American Economic Review* 102, 2301–6.

Jiménez, G., Ongena, S., Peydró, J. L., Saurina, J., 2014. Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking? *Econometrica* 82 (2), 463-505.

Kappel, V., (2010). The effects of financial development on income inequality and poverty. CER-ETH Working Paper 10/127.

Kashyap, A., Stein., J., 2000. What do a million observations on banks say about the transmission of monetary policy. *American Economic Review* 90, 407–28.

Kruse, J. B., Thompson, M. A., 2003. Valuing low probability risk: Survey and experimental evidence. *Journal of Economic Behavior and Organization* 50, 495-505.

Kumhof, M., rancière, R., 2010. Inequality, leverage and crises. IMF Working Paper 10/268.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. *Journal of Economic Literature* 48, 281-355.

Li, H., Squire, L., & Zou, H. F., 1998. Explaining international and intertemporal variations in income inequality. Economic Journal 108, 26-43.

Moore, E., Eckel, C. C., 2003. Measuring ambiguity aversion. Working Paper.

Naceur, S. B., Zhang, R., 2016. Financial development, inequality and poverty: Some international evidence. IMF Working Paper 16/32.

OECD, 2015. In it together: Why less inequality benefits all. Overview of inequality trends, key findings and policy directions. OECD Publishing.

Popov, A., Rocholl., J., Forthcoming. Financing constraints, employment, and labor compensation. *Journal of Financial Intermediation*.

Rajan, R. G., 1992. Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance* 47 (4), 1367-1400.

Rajan, R. G., 2010. Fault lines. How hidden fractures still threaten the world economy. Princeton University Press.

Schubert, R., Brown, M., Gysler, M., Brachinger, H. W., 1999. Gender and economic transactions. AEA Papers and Proceedings 89 (2), 381-385.

Sharpe, S., 1990. Asymmetric information, bank lending and implicit contracts: A stylized model of consumer relationships. *Journal of Finance* 45 (4), 1069-1087.

Solt, Frederick. 2016. "The Standardized World Income Inequality Database." *Social Science Quarterly* 97. SWIID Version 6.2, March 2018.

Table 1
Data and variable definitions

Variable	Description
A. Dimension of the data	
Individuals	Loan applicants who have an exclusive relationship with the bank and are majority owners (own more than 50%) of a firm. These borrowers apply to the bank for one or more business loans during the period 2002-2016 and the loan is either originated or denied. Due to the exclusive relationship, the bank holds information on the individuals' income and wealth even outside the year of loan application.
Year	The years covering the period 2002-2016.
B. Dependent variable	
Income	The euro amount of individuals' total annual income (in log).
C. Explanatory Variables: Running variable and cutoff	
Credit score	The credit score of the applicant, as calculated by the bank. We normalize this variable to take values around the cutoff of 0. The bank originates the loan if the credit score is higher than 0 and denies the loan otherwise.
Granted	A dummy variable equal to 1 if the loan is originated (Credit score>640) and 0 otherwise (Credit score<640).
D. Other covariates	
Education	An ordinal variable ranging between 0 and 5 if the individual completed the following education. 0: No secondary; 1: Secondary; 2: Post-secondary, non-tertiary; 3: Tertiary; 4: MSc, PhD or MBA.
Firm size	Total firm's assets (in log).
Firm leverage	The ratio of firm's total debt to total assets.
Gender	A dummy variable equal to 1 if the applicant is a male and 0 otherwise.
Loan amount	Log of the loan facility amount in thousands of euros.
Maturity	Loan duration in months.
Wealth	The euro amount of individuals' total wealth, as estimated by the bank (in log).
Initial wealth	Individuals' wealth in the first year before the loan application in which this information is available (one to five years before).

Table 2
Summary statistics

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

	Obs.	Mean	St. dev.	Min.	Max.
Income	61,863	11.01	0.376	9.852	12.29
Income t-1	57,682	10.58	0.406	9.804	12.62
Income t+1	57,766	11.10	0.388	9.866	12.58
Income t+3	49,514	11.14	0.373	9.987	12.57
Income t+5	41,391	11.16	0.363	10.04	12.62
Granted	61,863	0.867	0.498	0	1
Credit score	61,863	0.103	1.205	-2.921	2.100
Education	61,863	2.975	1.018	0	5
Firm size	61,863	12.821	0.806	2.500	12.03
Firm leverage	61,863	0.207	0.0249	0.143	0.917
Loan amount	61,863	2.323	0.845	0.679	7.480
Maturity	61,863	34.35	10.14	7	103
Wealth	61,863	12.14	0.556	8.564	14.05
Initial wealth	40,953	12.09	0.406	7.952	14.20

Table 3
Manipulation test

The table reports results from the manipulation testing procedure using the local polynomial density estimator proposed by Cattaneo et al. (2018). To perform this test, we rely on the local quadratic estimator with cubic bias-correction and triangular kernel. We report the conventional test statistic (Conventional) and the robust bias-corrected statistic (Robust) along with the corresponding p-value. The null hypothesis consists in no manipulation.

	T-stat	P-value
Conventional	1.5861	0.1127
Robust	1.2064	0.2277

Table 4
Results from parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is OLS on the RDD model of equation (2). Specifications (1) to (3) do not include any covariate besides the treatment and assignment variables. More covariates are included in specifications (4) to (6). The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Granted	0.0512***	0.0730***	0.0699***	0.0536***	0.0754***	0.0718***
	(0.0062)	(0.0064)	(0.0069)	(0.0063)	(0.0066)	(0.0072)
Credit score	-0.0015	0.0060	0.0120***	-0.0056	0.0027	0.0084*
	(0.0038)	(0.0039)	(0.0042)	(0.0039)	(0.0041)	(0.0044)
Granted x Credit score	-0.0013	-0.0122**	-0.0216***	0.0026	-0.0087	-0.0168***
	(0.0052)	(0.0053)	(0.0057)	(0.0053)	(0.0056)	(0.0060)
Income t-1				0.0958***	0.0653***	0.0452***
				(0.0041)	(0.0043)	(0.0045)
Education				0.0023	-0.0017	0.0004
				(0.0016)	(0.0017)	(0.0019)
Firm size				-0.0004	0.0030	-0.0015
				(0.0021)	(0.0022)	(0.0024)
Firm leverage				0.1872***	0.2877***	0.2435***
				(0.0672)	(0.0745)	(0.0778)
Loan amount				-0.0008	-0.0023	-0.0014
				(0.0020)	(0.0021)	(0.0023)
Maturity				0.0004**	0.0001	0.0002
				(0.0002)	(0.0002)	(0.0002)
Constant	11.0740***	11.1044***	11.1301***	9.9753***	10.3098***	10.5980***
	(0.0045)	(0.0047)	(0.0051)	(0.0517)	(0.0535)	(0.0558)
Observations	57,766	49,514	41,391	53,585	45,333	37,210
Adjusted R-squared	0.004	0.010	0.010	0.015	0.015	0.013
Clustering	Individual	Individual	Individual	Individual	Individual	Individual

Table 5
Results from non-parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. Specifications (1) to (3) do not include any covariate besides the assignment variable (Credit score). More covariates are included in specifications (4) to (6). The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5	(4) Income t+1	(5) Income t+3	(6) Income t+5
Conventional	0.0599***	0.0605***	0.107***	0.0623***	0.0605***	0.105***
	(0.0127)	(0.0134)	(0.0166)	(0.0126)	(0.0146)	(0.0170)
Bias-corrected	0.0632***	0.0572***	0.113***	0.0649***	0.0564***	0.112***
	(0.0127)	(0.0134)	(0.0166)	(0.0126)	(0.0146)	(0.0170)
Robust	0.0632***	0.0572***	0.113***	0.0649***	0.0564***	0.112***
	(0.0150)	(0.0159)	(0.0188)	(0.0150)	(0.0172)	(0.0194)
Obs.	57,766	49,514	41,391	53,585	45,333	37,210
Eff. obs. left of cutoff	8,731	7,510	4,487	8,274	6,171	4,061
Eff. obs. right of cutoff	9,186	7,855	4,686	8,670	6,398	4,232
BW estimate	61.37	61.30	44.03	62.61	54.76	44.08
BW bias	98.59	97.00	79.73	97.82	88.67	79.28

Table 6
Controlling for "initial" wealth

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. The table essentially replicates columns (3) to (6) of Table 5, the difference being the inclusion of Wealth t-5 as a control variable. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5
Conventional	0.0646***	0.0491***	0.112***
	(0.0148)	(0.0171)	(0.0227)
Bias-corrected	0.0681***	0.0450***	0.121***
	(0.0148)	(0.0171)	(0.0227)
Robust	0.0681***	0.0450**	0.121***
	(0.0175)	(0.0202)	(0.0260)
Obs.	36,856	28,604	20,481
Eff. obs. left of cutoff	5,312	4,238	2,207
Eff. obs. right of cutoff	5,572	4,386	2,295
BW estimate	57.92	58.91	42.43
BW bias	91.65	94.75	74.35

Table 7
Controlling for the probability of loan application in the parametric RDD

The table reports coefficients and standard errors (in parentheses) from a two-stage treatment effects model estimated with maximum likelihood. The first stage models the probability that individuals apply for a loan in a given year (probit model). The second stage is equivalent to equation (2), but including the fitted value of *Instantaneous probability of loan application* (i.e., the hazard rate) obtained in the first stage. The dependent variable is given in the first row of the table and all variables are defined in Table 1. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5
Granted	0.0468***	0.0738***	0.0751***
	(0.0078)	(0.0082)	(0.0091)
Credit score	-0.0061	0.0053	0.0058
	(0.0048)	(0.0051)	(0.0057)
Granted x Credit score	0.0027	-0.0100	-0.0137*
	(0.0065)	(0.0070)	(0.0076)
Instantaneous probability of loan application	0.0158*** (0.0026)	0.0271*** (0.0027)	0.0190*** (0.0030)
		First-stage results Probability of application	
Income		0.060***	
		(0.0086)	
Wealth		0.0065***	
		(0.0016)	
Education		0.0010	
		(0.0018)	
Firm size		0.004***	
		(0.0014)	
Firm leverage		0.273**	
		(0.1190)	
Gender		0.016**	
		(0.0073)	
Observations	34,448	28,662	23,049
Controls as in Table 4	Yes	Yes	Yes
Clustering	Individual	Individual	Individual

Table 8
Controlling for the probability of loan application in the non-parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. The table essentially replicates the analysis of columns 4-6 of Table 5, the difference being the inclusion of *Instantaneous probability of loan application* (i.e., the hazard rate) obtained in the first stage as a control variable in a non-parametric estimation of equation (2). Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	(1) Income t+1	(2) Income t+3	(3) Income t+5
Conventional	0.0388***	0.0449**	0.0934***
	(0.0142)	(0.0180)	(0.0207)
Bias-corrected	0.0415***	0.0445**	0.100***
	(0.0142)	(0.0180)	(0.0207)
Robust	0.0415**	0.0445**	0.100***
	(0.0169)	(0.0217)	(0.0241)
Obs.	34,448	28,662	23,049
Eff. obs. left of cutoff	6,136	3,909	2,738
Eff. obs. right of cutoff	6,400	4,031	2,852
BW estimate	71.87	54.16	47.40
BW bias	112.34	82.27	79.21

Table 9
Inequality measures

Panel A reports the Gini coefficient and the Theil index for individuals' income at time t and time t+5 around the cutoff (credit score < |0.1|). Panel B compares the equivalent Gini coefficients and Theil indices for the samples of granted and non-granted loans.

	Income t	Income t+5
Panel A. Inequality measures around the cutoff		
Gini coefficient	0.207	0.226
Theil index	0.067	0.074
Panel B. Inequality measures for accepted vs. denied applicants		
Credit is granted		
Gini coefficient	0.224	0.200
Theil index	0.080	0.065
Credit is denied		
Gini coefficient	0.193	0.214
Theil index	0.058	0.073

Table 10
Heterogeneity due to applicants' location

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. All specifications include the control variables as in Table 8. The first three and the last three specifications distinguish lower and higher income regions based on our sample's median. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The robust variance estimator is obtained according to Calonico at el. (2014).

	Low income			High income		
	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Robust	0.0642**	0.0710***	0.1203***	0.0605***	0.0597**	0.0926***
	(0.0279)	(0.0230)	(0.0380)	(0.0191)	(0.0182)	(0.0263)
Obs.	28,883	24,757	20,696	28,883	24,757	20,695
Eff. obs. left of cutoff	4,220	3,412	2,311	4,113	3,347	2,290
Eff. obs. right of cutoff	4,355	3,504	2,384	4,160	3,416	2,297
BW estimate	58.60	56.28	43.28	55.69	55.11	41.18
BW bias	94.30	88.25	75.61	92.50	88.26	72.16

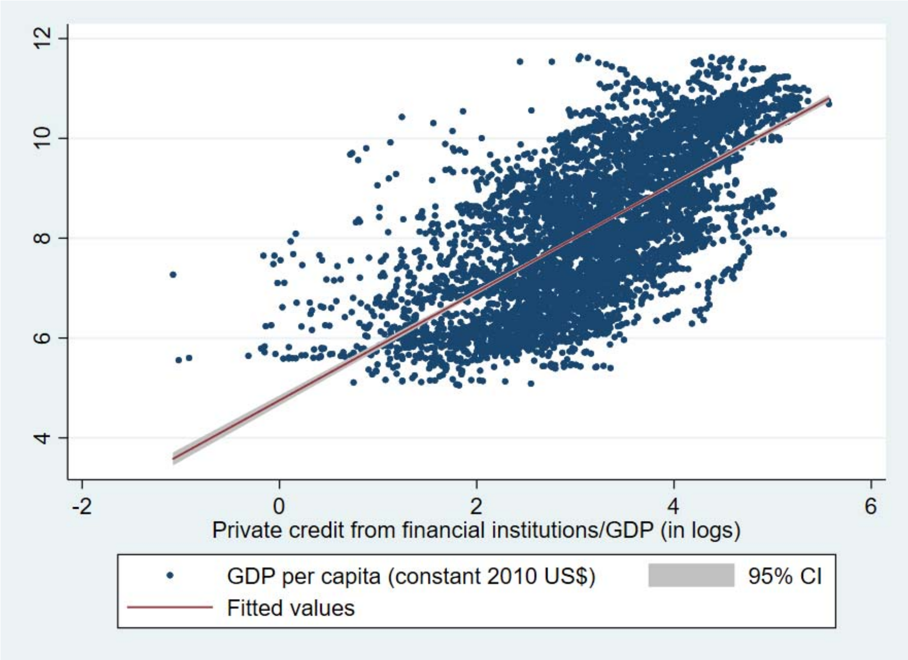
Table 11
Pre-post crisis

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator, the bias-corrected RD estimates with conventional variance estimator, and the bias-corrected RD estimates with robust variance estimator. All specifications include the control variables as in specifications 4-6 in Table 4. Specifications 1 to 3 are estimated using loan applications for the years 2009-2016 and specifications 4-6 using loan applications for 2000-2008 The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias. The bandwidth selection procedure is the one proposed by Calonico et al. (2014). The bias-corrected RD estimator and the robust variance estimator are obtained according to Calonico et al. (2018) and Calonico at el. (2014), respectively.

	Crisis and post-crisis			Pre-crisis		
	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Conventional	0.0552***	0.0722***	0.105***	0.0627***	0.0456**	0.0964***
	(0.0211)	(0.0215)	(0.0201)	(0.0144)	(0.0178)	(0.0249)
Bias-corrected	0.0610***	0.0700***	0.112***	0.0639***	0.0395**	0.104***
	(0.0211)	(0.0215)	(0.0201)	(0.0144)	(0.0178)	(0.0249)
Robust	0.0610**	0.0700***	0.112***	0.0639***	0.0395*	0.104***
	(0.0249)	(0.0258)	(0.0229)	(0.0172)	(0.0207)	(0.0291)
Obs.	20,850	20,850	20,850	32,735	24,483	16,360
Eff. obs. left of cutoff	3,509	2,977	2,992	5,613	3,886	1,778
Eff. obs. right of cutoff	3,657	3,099	3,110	5,876	4,040	1,874
BW estimate	68.69	58.09	58.34	69.29	63.39	43.29
BW bias	109.90	87.97	103.87	106.17	108.54	72.05

Figure 1
Income and income inequality against credit

The first graph depicts GDP per capita (in constant 2010 US$) against the ratio of private credit to GDP (x-axis). The second graph depicts the Gini index against the ratio of private credit to GDP (x-axis). We report individual values, as well as fitted values using a linear regression model. The estimated slopes of the linear regressions are 1.087 and -0.077, respectively, and are statistically significant at the 1% level. Data on the Gini index are from the Standardized World Income Inequality Database (SWIID); data on credit and GDP per capita are from the World Development Indicators.



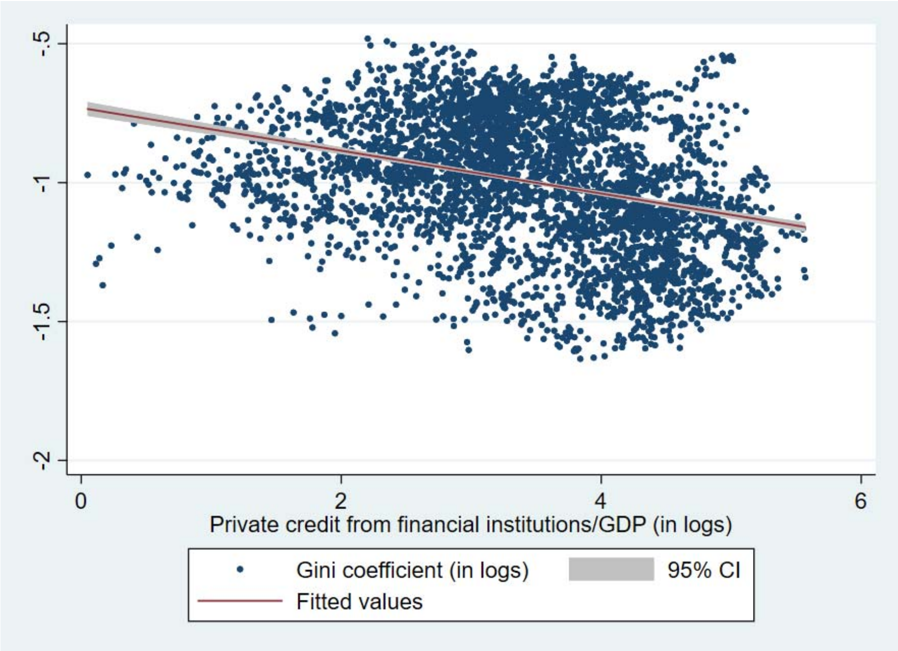
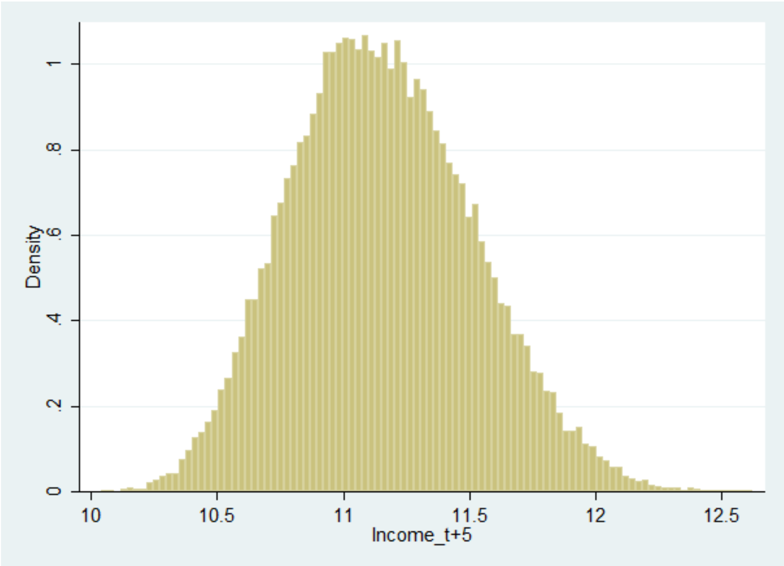


Figure 2
Densities of outcome and assignment variables

The figures report the probability densities for the outcome variable Income t+5 (top) and the assignment variable Credit score (bottom).



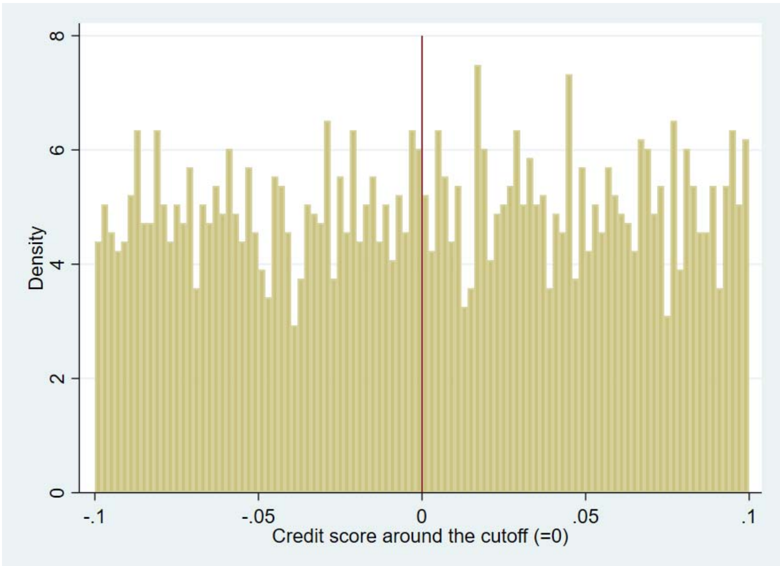


Figure 3
Manipulation test

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

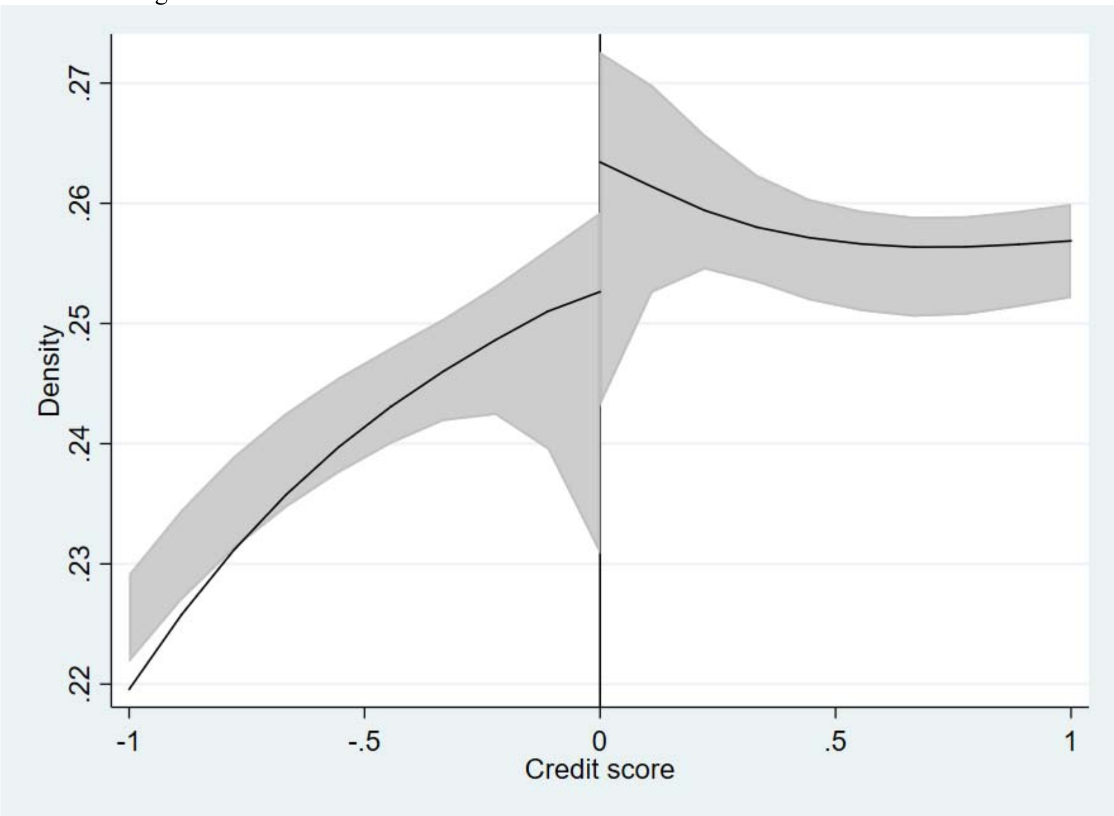


Figure 4

Applicants' income around the cutoff

The figure depicts applicants' *Income* five years ahead the loan decision (y-axis) against the Credit score (x-axis). The result is obtained from the non-parametric RDD. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of applicants' income below and above the cutoff.

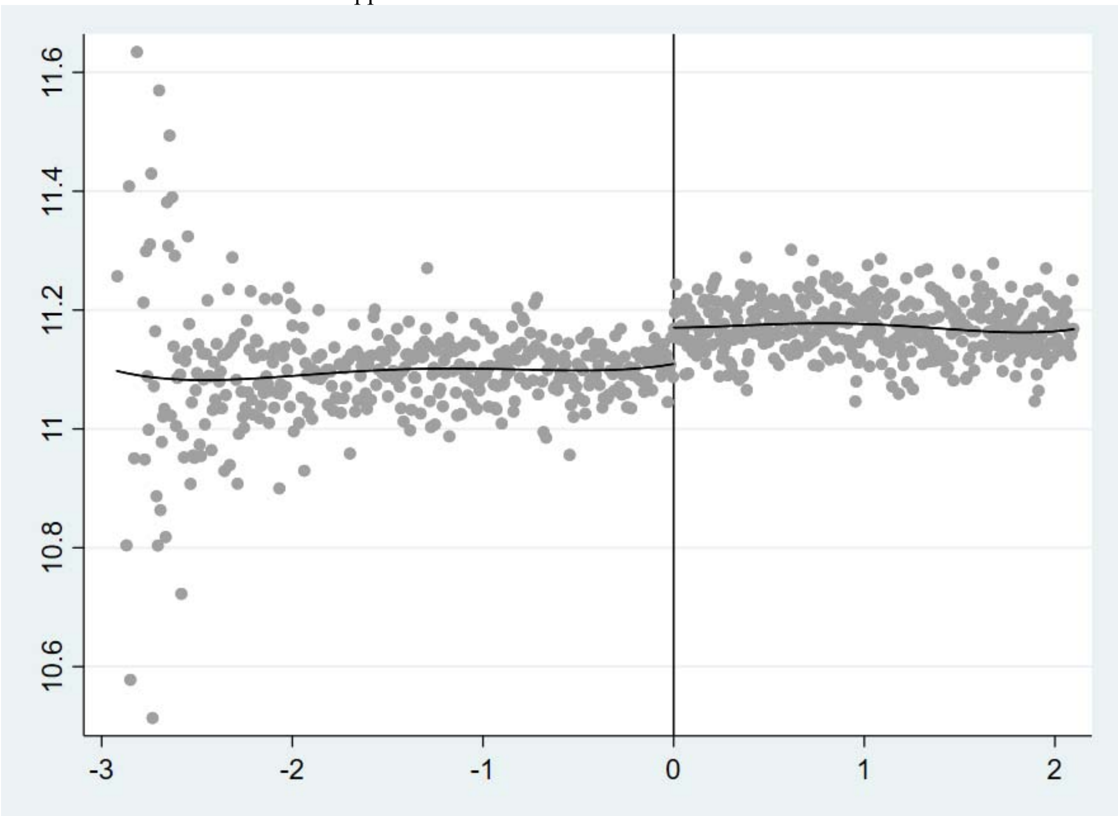
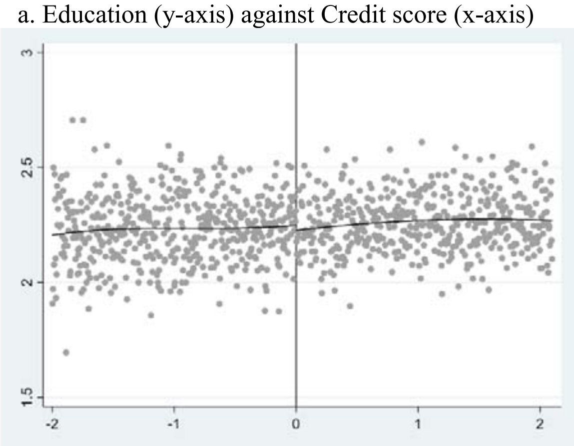


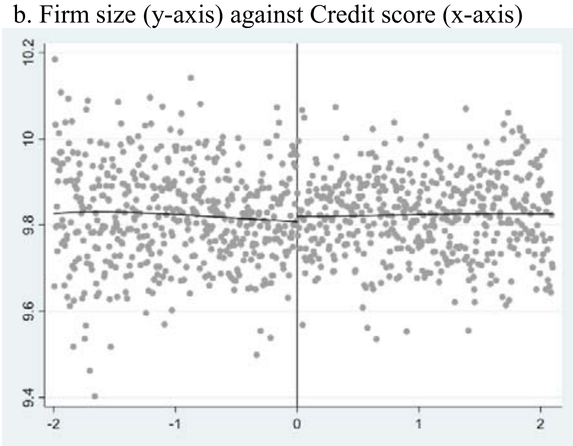
Figure 5
Covariates around the cutoff

The figure reports a plot for each control variable against the Credit score. The covariates include Education, Firm size, Firm leverage, Loan amount and Maturity. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of each covariate below and above the cutoff.

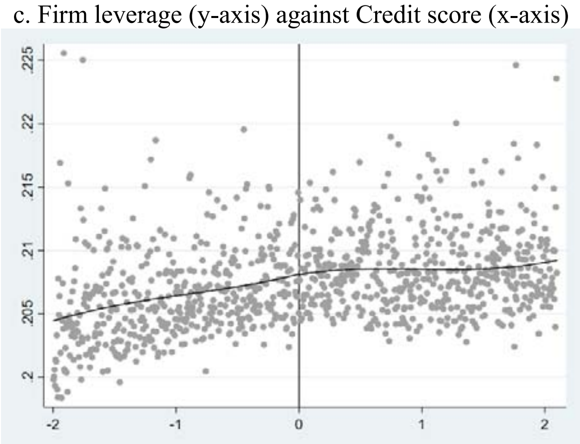
a. Education (y-axis) against Credit score (x-axis)



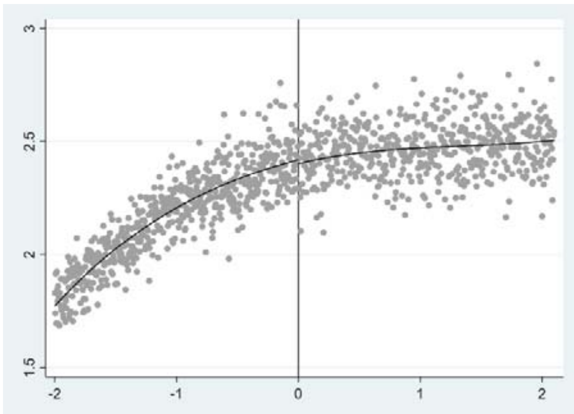
b. Firm size (y-axis) against Credit score (x-axis)



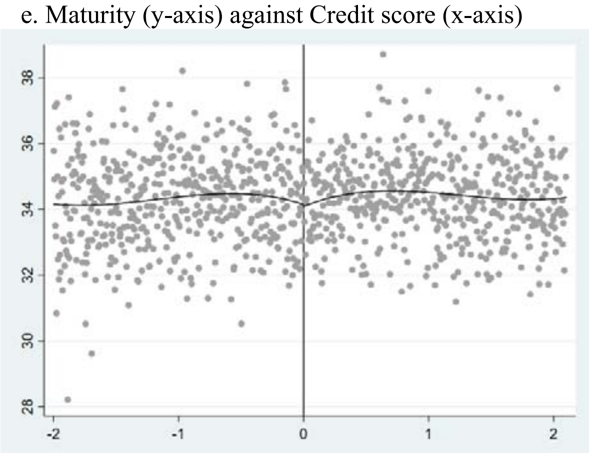
c. Firm leverage (y-axis) against Credit score (x-axis)



d. Loan amount (y-axis) against Credit score (x-axis)



e. Maturity (y-axis) against Credit score (x-axis)



f. Income t-1 (y-axis) against Credit score (x-axis)

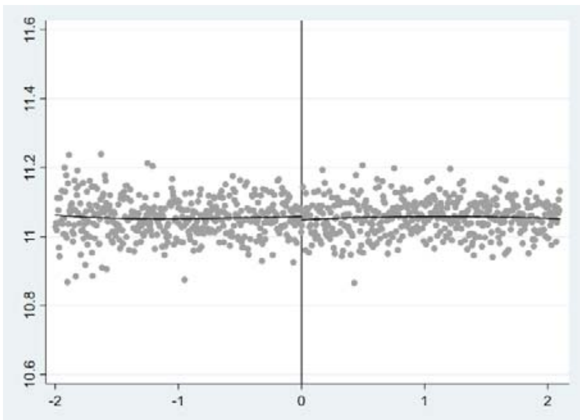
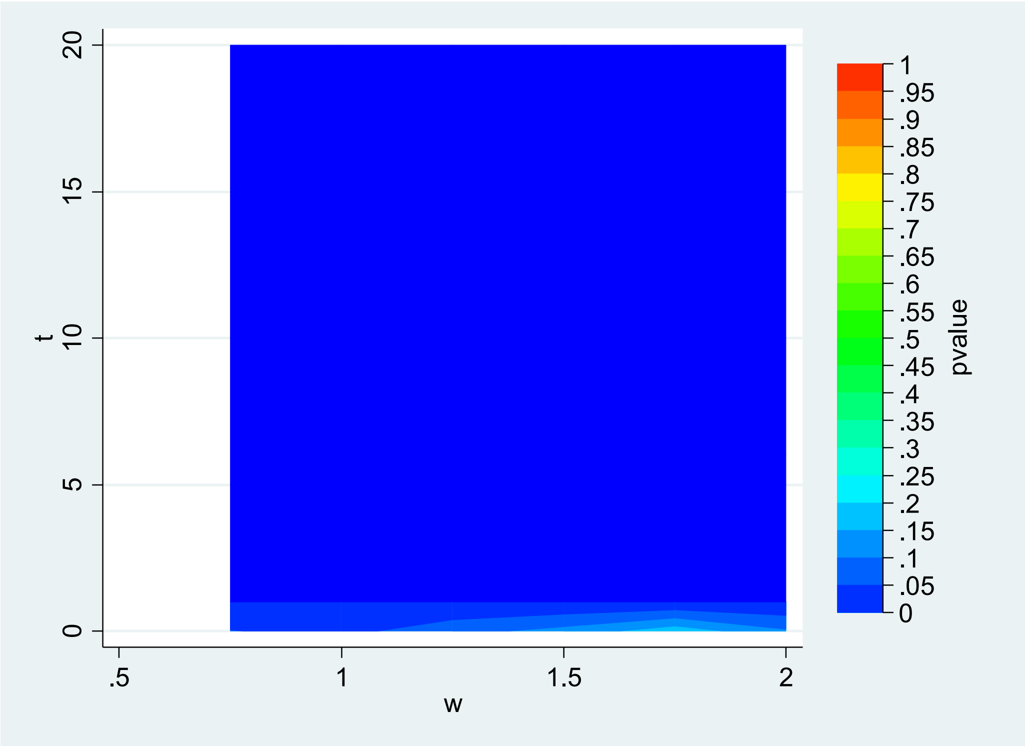


Figure 6
Sensitivity analysis for the RDD

The figure reports results from a sensitivity analysis under local randomization (see Cattaneo et al., 2016). We perform a sequence of hypotheses tests for different windows around the cutoff. Specifically, we show the test statistic of the null hypothesis of no treatment effect (x-axis) against the window length (y-axis). The p-values are calculated using randomization inference methods.



Appendix

The Appendix reports results from additional sensitivity tests. In Table A1 we include several fixed effects in the parametric model. In Table A2 we use different bandwidth-selection rules. In Table A3 we include *Initial wealth* in the parametric RDD and Figure A1 illustrates *Initial wealth* around the cutoff. Table A4 reports results from a semi-parametric model, where the parametrically estimated probability of loan application enters the non-parametrically estimated equation (2).

Table A1
Including industry, loan type, and year fixed effects in the parametric RDD

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is OLS on the RDD model of equation (2). Specifications (1) to (3) do not include any covariate besides the treatment and assignment variables. More covariates are included in specifications (4) to (6). All specifications include industry, loan type, and year fixed effects. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Granted	0.0534***	0.0751***	0.0713***	0.0536***	0.0754***	0.0718***
	(0.0063)	(0.0066)	(0.0072)	(0.0063)	(0.0066)	(0.0072)
Credit score	-0.0051	0.0029	0.0089**	-0.0056	0.0027	0.0084*
	(0.0038)	(0.0040)	(0.0044)	(0.0039)	(0.0041)	(0.0044)
Granted x Credit score	0.0021	-0.0089	-0.0172***	0.0025	-0.0087	-0.0168***
	(0.0052)	(0.0055)	(0.0059)	(0.0053)	(0.0056)	(0.0060)
Income t-1				0.0975***	0.0657***	0.0447***
				(0.0053)	(0.0056)	(0.0058)
Education				0.0023	-0.0017	0.0004
				(0.0016)	(0.0017)	(0.0019)
Firm size				-0.0004	0.0030	-0.0015
				(0.0021)	(0.0022)	(0.0024)
Firm leverage				0.1872***	0.2877***	0.2435***
				(0.0672)	(0.0745)	(0.0778)
Loan amount				-0.0008	-0.0023	-0.0014
				(0.0020)	(0.0021)	(0.0023)
Maturity				0.0004**	0.0001	0.0002
				(0.0002)	(0.0002)	(0.0002)
Constant	0.0429***	0.0297***	0.0209***	-0.0020	-0.0004	0.0005
	(0.0029)	(0.0030)	(0.0032)	(0.0038)	(0.0039)	(0.0041)
Observations	53,585	45,333	37,210	53,585	45,333	37,210
Clustering	Individual	Individual	Individual	Individual	Individual	Individual

Table A2
Alternative bandwidth selection methods

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the conventional RD estimates with conventional variance estimator. The specifications do not include any covariate besides the assignment variable (credit score). Specifications (1), (3), and (5) use the two mean squared error (MSE)-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect. Specifications (2), (4), and (6) use one common coverage error (CER)-optimal bandwidth selector for the RD treatment effect. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively. Obs. is the original number of observations. Effective obs. are the effective number of observations (determined by the bandwidth) left and right of the cutoff. BW estimate is the estimate of the bandwidth and BW bias is the associated bias.

	(1)	(2)	(3)	(4)	(5)	(6)
	Income t+1	Income t+1	Income t+3	Income t+3	Income t+5	Income t+5
	0.0611***	0.0716***	0.0610***	0.0645***	0.103***	0.0956***
	(0.0127)	(0.0167)	(0.0131)	(0.0178)	(0.0159)	(0.0215)
Obs.	57,766	57,766	49,514	49,514	41,391	41,391
Eff. obs. left of cutoff	7,743	5,053	8,260	4,373	5,180	2,599
Eff. obs. right of cutoff	10,530	5,284	7,802	4,536	4,831	2,738

Table A3
Controlling for "initial" wealth: OLS results

The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is OLS on the RDD model of equation (2). The table essentially replicates columns (3) to (6) of Table 4, the difference being the inclusion of Wealth t-5 as a control variable. The *, **, and *** marks denote statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
	Income t+1	Income t+3	Income t+5
Granted	0.0514***	0.0726***	0.0814***
	(0.0072)	(0.0080)	(0.0094)
Credit score	-0.0071	-0.0023	0.0003
	(0.0044)	(0.0050)	(0.0059)
Granted x Credit score	0.0028	-0.0020	-0.0083
	(0.0060)	(0.0068)	(0.0079)
Income t-1	0.0816***	0.0600***	0.0450***
	(0.0051)	(0.0056)	(0.0064)
Education	0.0032*	-0.0027	0.0013
	(0.0018)	(0.0021)	(0.0024)
Firm size	-0.0001	0.0024	-0.0007
	(0.0024)	(0.0027)	(0.0031)
Firm leverage	0.1898**	0.1764**	0.2908***
	(0.0765)	(0.0850)	(0.1051)
Loan amount	0.0001	0.0014	0.0006
	(0.0023)	(0.0026)	(0.0030)
Maturity	0.0004*	-0.0000	0.0001
	(0.0002)	(0.0002)	(0.0003)
Wealth t-5	0.0215***	0.0148***	0.0046
	(0.0032)	(0.0035)	(0.0040)
Constant	9.9057***	10.2427***	10.5395***
	(0.0736)	(0.0803)	(0.0929)
Observations	36,856	28,604	20,481
Clustering	Individual	Individual	Individual

Figure A1
Initial wealth around the cutoff

The figure reports Wealth t-5 (first instance of wealth before the loan application) against the Credit score. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of each covariate below and above the cutoff.