Credit, Income and Inequality

Manthos Delis
Montpellier Business School

Fulvia Fringuellotti
Federal Reserve Bank of New York

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

December 31, 2020

Delis: Montpellier Business School-Montpellier Research in Management, Email: m.delis@montpellier-bs.com; Fringuellotti: Federal Reserve Bank of New York, Email: fulvia.fringuellotti@ny.frb.org; Ongena: Department of Banking and Finance, University of 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, the 9th MoFiR Workshop on Banking, the Conference on Sustainable development, firm performance and competitiveness policies in small open economies, the European Finance Association 2020 Annual Meeting, the 35th Annual Congress of the European Economic Association, and the 18th Paris Finance Meeting for many comments. We also thank seminar participants in the Essex Business School (University of Essex), Adam Smith Business School (University of Glasgow), Durham Business School, Federal Reserve Bank of New York, Montpellier Business School, and the Athens University of Economics and Business. For comments, we thank Knut Are Aastveit, Christoph Basten, Tobias Berg, Rajashri Chakrabarti, Nicola Cetorelli, Jerry Coakley, Francisco Covas, Hans Degryse, Yota Deli, Filippo Di Mauro, Robin Döttling, René Garcia, Miguel Ferreira, Pedro Gete, Mariassunta Giannetti, Linda Goldberg, Jens Hagendorff, Iftekhar Hasan, Rajkamal Iyer, Olivier De Jonghe, Sotirios Kokas, Alexandros Kontonikas, Michael Lamla, David Lucca, John Mondragon, Georgios Panos, Nicola Pavanini, Diane Pierret, Maxim Pinkovskiy, Tomasz Piskorski, Matthew Plosser, Andrea Presbitero, Simon Price, Manju Puri, Enrichetta Ravina, Jean Charles Rochet, Farzad Saidi, Asani Sarkar, László Tétényi, Wilbert van der Klaauw, and Diego A. Vera-Cossio. Ongena acknowledges financial support from ERC ADG 2016 - GA 740272 lending. This paper was previously circulated under the titles "Credit and Income Inequality" and "Credit and Income". Fulvia Fringuellotti was affiliated with the University of Zurich and the Swiss Finance Institute at the time this paper was written. The views expressed in this paper are those of the authors only and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

Credit, Income and Inequality

Abstract

Analyzing unique data on loan applications by individuals who are majority owners of small firms, we detail how a bank's credit decisions affect their future income. We use the bank's cutoff rule, which is based on the applicants' credit scores, as the discontinuous locus providing exogenous variation in the decision to grant loans. We show that application acceptance increases recipients' income five years later by more than 10% compared to denied applicants. This effect is mostly driven by the use of borrowed funds to undertake investments and is stronger when individuals are more credit constrained.

Keywords: Credit constraints; Income; Business loans; Economic mobility; 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), posing serious concerns for economic growth and social cohesion (Galor and Zeira, 1993; Alesina and Rodrick, 1994; Galor and Moav, 2004; Persson and Tabellini, 1994 Putnam, 2000; Stiglitz, 2012; Larsen, 2013; Piketty and Saez, 2013). The increase in income inequality has been associated with an increase in intergenerational social immobility in many countries, creating an upward sloping schedule commonly referred as "the Great Gatsby curve" (Corak, 2013; Kearney and Levine, 2016; Chetty et al., 2017). A lively debate ensued on the sources of this phenomenon and the proper measures to contain the problem. The role of finance is at the forefront in shaping economic opportunities of households and businesses.

This study aims to identify and quantify how banks' credit decisions (credit acceptance or rejection) affects applicants' future income. The extent to which applicants that are similar in terms of income and other traits experience significantly different incomes after the credit decision has important implications for the real effect of credit on upward mobility and inequality.

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, when a credit expansion accompanies a relaxation of credit constraints, it leads 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 limited wealth, and their exclusion from credit can hinder economic mobility and fuel persistent income 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; Demirgüç–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).

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 (income inequality) may actually drive the availability of credit (Kumhof and Rancière, 2010; Rajan, 2010). Omitted-variable bias is an additional concern due to unobserved factors that are difficult to measure (e.g., the availability of new investment ideas), which jointly affect the distribution of income and the degree of financial depth.

[Insert Figure 1 here]

Our study provides the first empirical analysis of how access to credit affects individuals' income and its distribution in a developed economy, by comparing the future incomes of accepted applicants to those of rejected applicants with approximately the same characteristics (e.g., income and credit quality). We identify this effect using a unique data set of business loan applications to a single large European bank. This is a systemic bank subject to the ECB Single Supervisory Mechanism and headquartered in a highly developed northern European country.

Our focus is on loan applications from small and micro enterprises that are majority-owned by individuals. These applicants have an exclusive relationship with the bank, meaning that they do not have a lending relationship with another regulated bank at the time of the loan application, even if their loan application is rejected. 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. The credit score is an internal rating constructed by the bank and it is not disclosed to the applicant. Then, 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 that 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 11% 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 11% increase in their incomes compared to marginally rejected applicants, thereby significantly affecting the distribution of income in the two groups. This finding is robust to several re-specifications and is not affected by sample selection or 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. Overall, our result suggests that bank credit decisions (loan origination or denial) affect individuals' income in a significant way improving upward mobility.

This finding is not trivial. In principle, a loan origination should have a positive effect on income only to the extent that credit is granted to applicants having good investment opportunities. A relevant strand of literature shows that microcredit programs conducted in several developing countries did not have a significant impact on individuals' income (Angelucci et al., 2015; Attanasio et al., 2015; Augsburg et al.; 2015; Banerjee et al., 2015; Banerjee et al., 2015b; Crépon et al., 2015; Tarozzi et al., 2015; Banerjee et al., 2018). An important contribution of our study is that it focuses on loan provision by a large commercial bank in a developed economy in Europe. In this context, even highly constrained borrowers may have valuable investment projects and banks are likely to extend credit efficiently.

In a series of extensions to our baseline model, we examine the mechanisms behind the effect of credit access on individual income. First, we show that a loan origination entails a stronger increase in the income of applicants owning young firms compared to business owners of old firms. Second, firms of accepted applicants invest more in business operations, are more likely to repay existing bank loans, experience a higher increase in profitability, and grow at a higher rate compared to firms of rejected applicants. Overall, these results reveal that access to credit is pivotal for small firms to exploit good investment opportunities, expand their business, and improve profitability. Third, we show that the effect of credit on income is more pronounced when a loan acceptance is more likely based on soft information held by the bank (for example on the quality of the investment opportunities of the firm). This confirms that the effect of a loan origination on income is far from obvious, as it depends on how efficiently the bank extends credit.

We then 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 this set 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 access to credit is improved (Greenwood and Jovanovic, 1990).

We further examine the heterogeneity of our findings in interesting subsamples reflecting additional aspects of how credit affects income and its distribution more generally. 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 on microfinance in developing countries (Angelucci et al., 2015; Attanasio et al., 2015; Augsburg et al.; 2015; Banerjee et al., 2015; Banerjee et al., 2015b; Crépon et al., 2015; Tarozzi et al., 2015; Banerjee et al., 2018). As highlighted above, these studies show that various microcredit programs did not have a significant impact on individual income in developing countries, presumably because recipients lack of good investment projects. We show that, in the context of a developed economy and a well-established financial institution, a loan origination has instead a positive and strong effect on applicants' income when credit is granted to constrained individuals with good investment opportunities.

A substantial body of related literature examines how various social and economic conditions (including race, gender, education, parents' socioeconomic class, local neighborhood, income inequality etc.) affect individual opportunities and, hence, economic mobility (Chetty et al., 2014; Chetty and Hendren, 2018a, 2018b; Bell et al., 2019, Bergman et al., 2019; Chetty et al., forthcoming). We contribute to this literature documenting that credit provision to small businesses is pivotal in fostering entrepreneurship and upward mobility.

Our work also 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 a credit expansion corresponds to a relaxing of credit constraints. Denk and Cournède (2015), Jauch and Watzka (2016), and de Haan and Sturm (2017), point instead to a positive relation, suggesting that financial development improves access to credit only for the rich. Kim and Lin (2011), and Brei et al. (2018) identify a non-monotonic relation depending on the degree of financial development and the financial structure of the economy. Minetti et al. (2019) show that local banking structures affect regional income inequality. Our paper also relates to several other studies on finance and income inequality (for a thorough review, see Demirgüç-Kunt and Levine, 2009).[1] We contribute to this literature by 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 and social outcomes. Looking at the Home Owners Loan Corporation "redlining" maps drawn in the 1930s, Appel and Nickerson (2016) and Aaronson et al. (2019) show that reduced access to credit in certain city neighborhood has negative long-lasting effects on home ownership, house prices, and rents, while increasing racial segregation. Using data on loan applications (such as ours), Berg (2018) documents 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), bank conditions (Klein et al, 2002; Gan, 2007; Duchin et al., 2010; Cingano et al., 2013; Chodorow-Reich, 2014; Balduzzi et al., 2017; Bentolila et al., 2017; Choudhary and Jain, 2017; Acharya et al., forthcoming; Popov and Rocholl, forthcoming), or regulation (Duflo and Banerjee, 2014). We contribute to this literature showing that the effect of credit origination on income is stronger in low-income regions and in the crisis period, where individuals are more credit constrained.

[1] Our paper also relates to Saez et al. (2012) and Moser et al. (2018), who look at the effect of payroll taxation and credit supply, respectively, on inequality in wages.

From a methodological perspective, we use uniquely granular data from a single bank as in Iyer and Puri (2012), and Berg (2018). The detailed information on loan applications that we exploit ensures that we rigorously assess the effect of credit decisions on individual income and inequality at the micro level.

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 large European bank supervised by the ECB Single Supervisory Mechanism and headquartered in a rich northern European country. The bank provides credit to a wide array of small and large firms, as well as to consumers, households, and the public sector both domestically and abroad. Our sample is limited to loan applications from individuals, firms and administrations that are located in the country where the bank is headquartered. 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.[2] 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,

[2] 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.

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. We identify applicants having an exclusive relationship with the bank as those who do not have a lending relationship with another regulated commercial bank, even if their application(s) to our bank is (are) rejected. The exclusivity of the relationship consists in an objective fact and does not stem from any legal agreement between the firm and the bank. For two reasons, we focus on loan applications from individuals who have exclusive relationships with the bank and apply at least twice during the sample period. First, the bank has income information for these applicants for several years before and after the loan decision. Second, these applicants are generally unable to obtain credit from another bank, especially if their application is denied; moreover, they cannot access capital markets due the firm's small size.[3] These characteristics of our sample allow us to identify the effect of the bank's credit decision on applicants' income.

Each applicant is given a credit score at the time of the application, and this score is the decisive factor in loan origination. The credit score consists in a private rating constructed by the bank, which is not accessible to anyone including the applicant. For comparative purposes, we normalize the credit score to be 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

[3] We have information about this exclusivity from the bank. However, the firms can receive credit (obviously at higher rates) in the shadow-banking sector.

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.[4] 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. The bank accepts (or partially accepts) approximately 87% of loan applications and rejects 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).[5] 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 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

[4] 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.
[5] See, https://ec.europa.eu/growth/content/survey-access-finance-enterprises-safe-was-published-today_en.

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.

Using data from a single entity is not an unusual practice when the research question is detailed (Adams et al., 2009; Iyer and Puri, 2012; Berg, 2018). In our case, we take advantage of granular application-level data for one bank to document how the decision to grant or deny credit affects individuals' income. Also, the bank that we look at is a major financial institution operating on a national scale. This ensures that the bank is representative enough for the banking system, so that we can reasonably generalize the results of our study.

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 (the applicant cannot obtain credit from another regulated commercial bank if his/her application is rejected),[6] 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

[6] The bank has this information from the applicants, meaning that no other bank is able/willing to finance the same project. This feature of our sample implies that the loan applicants do not leave the sample; therefore, we do not have such attrition bias.

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.[7] This implies a sharp discontinuity in treatment as a function of credit score.[8] 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).

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

[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. Graphical Evidence

We begin our RDD analysis with a graphical inspection of the relation between access to credit and income. Figure 4 shows applicants' income five years after the loan decision against the credit score. There is 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]

Also, the plot shows a linear relation between applicants' income and the credit score on both sides of the cutoff. The relation looks slightly increasing below the cutoff and almost flat above. This evidence suggests that the econometric analysis should focus on a linear model allowing for a different slope on each side of the cutoff. More importantly, the upward discontinuity in applicants' income at the cutoff, as well as the flat relationship between income and credit score above the cutoff, reveal that access to credit plays a preeminent role in shaping the future income path of small business owners.

B. 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.

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.[9] 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).

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]

[9] 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.

C. 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 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.[10] 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.

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

Overall, the estimates of the treatment effect are comparable to those in the corresponding regressions of Table 4. Given the small discrepancy in the results between the parametric and nonparametric RDD and the advantages of the nonparametric 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]

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.[11]

[Insert Figure 6 about here]

In specifications 4-6 of Table 5 we estimate the effect of credit on income controlling for a wide set of loan, firm, and applicant characteristics, including the requested loan amount and maturity. The lending rate applied on a new loan determines the future stream of payments and, hence, may affect the recipient's future income. Specifically, we would expect that the higher is the credit score of a borrower, the lower is the interest rate applied. Figure 7 shows that the income of accepted applicants considered in the non-parametric RDD one year after the loan decision is a flat function of the lending rate. This means that the interest rate charged on newly granted loans does not influence the effect of loan acceptance on individual income.

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

[Insert Figure 7 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. This link is not obvious. As documented in various studies on microfinance in developing countries, access to credit may have no impact on individual income (Angelucci et al., 2015; Attanasio et al., 2015; Augsburg et al.; 2015; Banerjee et al., 2015; Banerjee et al., 2015b; Crépon et al., 2015; Tarozzi et al., 2015; Banerjee et al., 2018). Intuitively, a loan origination improves individual income only if credit is granted to applicants having good investment opportunities. This is likely to be the case for our bank, which is a major financial institution in Europe. Therefore, our findings reveal that access to credit has a positive effect on individual income when lending decisions are taken efficiently. Also, the magnitude of this effect is substantial, suggesting that credit provision to small businesses impacts significantly the firm owner's economic opportunities and upward mobility.

The large increase in income experienced by accepted applicants vis-à-vis rejected applicants with similar attributes might show that the bank overlooks good investment opportunities. As mentioned before, the percentage of denied applications of this bank is in line with the European averages reported in the Survey on access to finance for enterprises (SAFE) published by the European Commission and the ECB. This suggests that the bank may limit its lending capacity as a result of an optimization process. However, further looking into that optimization process is beyond the scope of this paper and we leave it for further research.

D. 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 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 considers a balanced panel of individuals with an exclusive relationship with the bank. These individuals are firm owners who do not have a lending relationship with another regulated bank at the time of the loan application, and who apply at least twice during the sample period so that we have information on their income for several years before and after the loan decision. While working on such balanced panel limits concerns of attrition bias and allows us to estimate the treatment effect focusing on individuals for which we have comprehensive information, there is a downside related to the potential introduction of a selection bias. This is because we overlook one-time applicants who may drop out of the sample because they turn to another lender or decide to stop operating their business (for example after a denied application). We also discard firm owners who have credit relationships with multiple banks. If these applicants differ in a substantial way from individuals having a long-lasting exclusive relationship with the bank, we may either underestimate or overestimate the effect of credit of income.

To address this issue, we use a parametric two-stage selection model as in, e.g., Heckman (1976), Dass and Massa (2011), and Jiménez et al. (2014). In the first stage, we estimate the probability that a bank customer having a long-lasting exclusive relationship with the bank applies for a loan in a specific year (probit model). We run this regression on the full sample, which consists in an unbalanced panel including all applicants, irrespective of whether they have an exclusive relationship with the bank or not. The right-hand side variables in the first stage encompass the applicant's attributes of columns 4-6 of Tables 4 and 5, excluding the credit score (which is unknown to the applicant) and including *Gender*.[12] 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 (Mills ratio) from the first stage as an additional control variable.[13]

Table 7 reports the estimation results. The first-stage results show that income, wealth and education positively and strongly affect the probability of a loan application by an individual with a long-lasting exclusive relationship with the bank. The same holds for owners of more leveraged

[12] We find that *Gender* is significantly correlated with the probability of a loan application by an individual with a long-lasting relationship with the bank but does not explain income in the baseline specifications.
[13] Given that the sample of our baseline RDD is a balanced panel of bank customers with an exclusive credit relationship and these customers appear in the panel irrespective of whether they apply for a loan in a given year, we can also model the probability of receiving a loan application in the baseline setup. The results of this exercise are similar to those of Tables 7-8 and are available upon request.

firms. Interestingly, we also find that male applicants are 0.8% more likely to apply for credit than female applicants. The second-stage results are fully in line with Table 4, with the Mills ratio having a positive but insignificant coefficient. This suggests that the selection effect is very low and the estimation of the treatment effect using a balanced panel of individuals having a long-lasting exclusive relationship with the bank delivers reliable results.

To account for selection of loan applicants, we prefer to use the conventional parametric model because it is standard in the applied economics/finance literature, whereas the nonparametric models are quite rare in this respect.[14] However, we do an 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]

E. Economic Channels

In our baseline models, we estimate the effect of loan origination on individual income controlling for the size of the loan. Intuitively, the treatment intensity should be stronger the higher is the loan amount. In the first two specifications of Panel A of Table 9, we replicate the regression of column 6 of Table 5, splitting our full sample into small loans and large loans based on the median loan amount. As expected, we find that the effect of credit access on individual income is stronger for larger loan amounts. In particular, the income of approved applicants rises by 11.8% five years ahead of the loan origination for large loans (column 2), versus 10.5% for small loans (column 1).

[14] In a two-stage linear Heckman model we also can correctly adjust the standard errors.

Next, we more broadly examine the mechanism behind the observed effect of a positive credit decision on the income of small business owners. In principle, an accepted applicant may use the borrowed funds to invest and expand the business or to smooth consumption over time. To test these economic channels, we rely on a wide set of econometric models where we consider different subsamples and various firm outcomes as dependent variable.

We start by replicating our baseline regression of column 6 of Table 5, separating our sample into new firms and old firms, which are identified as the 25th and the 75th percentile of the distribution of firm age respectively. The last two specifications of Panel A of Table 9 show that, five years after a bank's credit decision, accepted applicants owning a new firm experience an increase in income of 16.7% (column 3), which is more than double the increase in income observed for those who own old firms (column 4). The difference of 10.5 percentage points in the effect of loan origination on individual income for business owners of young versus old firms is economically very meaningful. This suggests that access to credit is crucial at the early stage of a business to allow firm investments that foster growth and expansion.

This conjecture is confirmed from the results presented in Panel B of Table 9, where we explicitly look at the evolution of various firm outcomes in response to a credit origination. From a methodological perspective, we use a similar econometric model to that of column 6 of Table 5, the difference being the dependent variable, which consists in the following firm outcomes: the amount of credit borrowed for working capital, which is used to finance everyday business (column 1); a dummy equal to one if the firm is repaying previous loan obligations with the bank and zero otherwise (column 2); firm profitability as captured by the return on assets (column 3); and the growth rate of firm assets (column 4). We find that, five years after the credit decision, firms of accepted applicants invest more in short-term business operations, are more likely to repay existing bank loans, experience a higher increase in profitability, and grow at higher rate compared to firms of rejected applicants. Overall, these results suggest that access to credit is crucial for small firms to undertake investments, expand their business, and be more profitable. This, in turn, has positive repercussions on future income of the business majority owner. More generally, our findings reveal that credit provision to small businesses (having good investment opportunities) is pivotal to foster entrepreneurship and economic mobility.

[Insert Table 9 about here]

F. Hard Information and Soft Information

As a further extension, we explore the role played by hard and soft information held by the bank in driving the real effect of credit decisions on individuals' income. Hard information consists in the observable characteristics listed in Table 1. Soft information includes any other relevant feature of the applicant and the firm that is unobservable, such as the quality of the investment opportunities of the firm, the bank's perception of the loan applicant, etc. While both hard information and soft information contribute to the bank's credit decision, what leads the effect of credit on income is far from clear.

To decompose the credit score into hard information and soft information, we regress the credit score on the set of observables capturing hard information (income, wealth, education, firm size, firm leverage, loan amount, maturity, availability of collateral, and use of loan covenants). We then interpret the residuals as the component of the credit score ascribable to soft information. We find that 77% of the credit scored is explained by hard information. As a second step, we replicate the nonparametric regressions in columns 4-6 of Table 5, splitting the data in two subsamples on the basis of the sign of the residuals (positive residuals in the first subsample and negative or equal to zero in the second).

Table 10 reports the results. Despite soft information explaining only 23% of the credit score, we find that the effect of credit origination on individuals' income is stronger when soft information makes a loan acceptance more likely. In particular, five years after a bank's credit decision, accepted applicants experience an increase in income of 13.5% when soft information enters positively into the credit score (column 3), compared to 7% when soft information contributes negatively (column 6). This finding suggests that access to credit improves individual income especially when a loan acceptance is favored by a positive assessment of the bank on unobservable characteristics of the applicant. This further confirms that the effect of loan origination on income is far from trivial, as it depends on the level of efficiency of the bank in granting credit.

[Insert Table 10 about here]

G. Reflection on Income Inequality

A natural implication of our key finding is that the 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.[15]

Panel A of Table 11 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.

[Insert Table 12 about here]

In Panel B of Table 12, 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 the whole income distribution.[16] The first concerns the role of applicant location based on regional income, distinguishing between low-income regions and high-income regions.[17] In Table 12, we replicate

[15] 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.

[16] A natural way to analyze the effect of credit access on the whole income distribution would be to replicate the non-parametric RDD by splitting the sample into low income and high-income individuals based on, e.g., the 25th quantile and the 75th quantile of the income distribution at t-1. We are unable to perform this exercise because no applicant with a past income above the 75th percentile of the income distribution is rejected.

[17] This analysis is in the same spirit of Agarwal et al. (2018), who document an income-based geographical heterogeneity in the effect of a micro-credit program on financial access in Rwanda.

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.[18]

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 12 about here]

In the second 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. Thus, we expect that loan origination has a stronger effect on applicant income 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.[19]

---

[18] In our sample, the mean value of *Granted* in high-income regions is 0.880; it is 0.853 in the low-income regions.
[19] See https://data.worldbank.org/indicator/FS.AST.PRVT.GD.ZS?locations=XC.

Table 13 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 by 10.4% five years ahead of the loan origination during 2000-2008 (column 6), versus 11.2% 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 13 about here]

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 large European bank, we study and quantify how a bank's credit decision (acceptance or rejection) affects individuals' future incomes. Subsequently, we use our results to quantify the role of the bank's credit decisions in the distribution of income and inequality.

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 with that bank (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 6% higher than the income of denied applicants one to three years ahead of the loan decision; this jumps to 11% five years ahead. This finding is robust to several re-specifications and robustness tests.

We next investigate the economic channels behind the observed positive impact of credit access on individual income. We show that firms of accepted applicants use the borrowed funds to make investments and expand their business, ultimately experiencing higher profitability and growth rates compared to firms of rejected applicants. We also show that the effect of credit origination on income and its distribution is more pronounced when a loan acceptance is favored by soft information held by the bank (for example on the quality of the investment opportunities of the firm).

Overall, these results suggest that (an efficient) credit provision to small businesses has a positive impact on individual upward mobility. We next investigate the implications of our such effect in terms of income inequality. We show that, while the Gini and Theil indices increase for indices increase for individuals around the cutoff, the distribution of income is tighter among accepted applicants and wider among rejected applicants. We also document that loan acceptance has a 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), hereby affecting the distribution of income in the economy. These results are in line with the theory pointing to a negative relation between credit availability and income inequality.

Our findings have two key and interrelated economic implications. First, efficient 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 among individuals who prior to the credit decision have similar credit scores.

In general, the evidence that efficient credit decisions affect positively economic mobility and reduce inequality provides support to policy interventions aimed at increasing credit access to loan applicants rejected by the banking system due to lack of credit history or collateral. 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, and of the Small Business Administration, which guarantees loans to small firms lacking access to credit but having good business financials. We leave the thorough examination of the effects of these policies to future research.

References

Aaronson, D., Hartley, D., Mazumder, B., 2019. The effects of the 1930s HOLC "Redlining" Maps. Federal Reserve Bank of Chicago Working Paper No. 2017-12.

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.

Adams, W., Einav, L., Levin, J., 2009. Liquidity constraints and imperfect information in subprime lending. *American Economic Review* 99 (1), 49-84.

Agarwal, S., Kigabo, T., Minoiu, C., Presbitero, A., Silva, A., 2018. Financial inclusion under the microscope. IMF Working Paper No. 18/208.

Alesina, A., Rodrik, D. 1994. Distributive politics and economic growth. *Quarterly Journal of Economics* 109 (2), 465-490.

Angelucci, M., Karlan, D., Zinman, J., 2015. Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco. *American Economic Journal: Applied Economics* 7 (1), 151-182.

Appel, I., Nickerson, J., 2016. Pockets of poverty: The long-term effects of redlining. Working paper.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., Harmgart, H., 2015. The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *American Economic Journal: Applied Economics* 7 (1), 90-122.

Augsburg, B., De Haas, R., Harmgart, H., Meghir, C., 2015. The impacts of microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics* 7 (1), 183-203.

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.

Banerjee, A., Duflo, E., Glennerster, R., Kinnan, C., 2015. The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics* 7 (1), 22-53.

Banerjee, A., Karlan, D., Zinman, J., 2015. Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics* 7 (1), 1-21.

Banerjee, A., Duflo, E., Hornbeck, R., 2018. How much do existing borrowers value microfinance? Evidence from an experiment on bundling microcredit and insurance. *Economica* 85, 671-700.

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.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., Van Reenen, J., 2019. Who becomes an inventor in America? The importance of exposure to innovation. *Quarterly Journal of Economics*, 134 (2), 647-713.

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.

Bergman, P., Chetty, R., DeLuca, S., Hendren, S., Katz, L. F., Palmer, C., 2019. Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. NBER Working Paper No. 26164.

Brei, M, Ferri, G., Gambacorta, L., 2018. Financial structure and income inequality. BIS Working Papers No. 756.

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.

Chetty, R., Grusky, D., Hell, M., Hendren, N., Manduca, R., Narang, J., 2017. The fading American dream: Trends in absolute mobility since 1940. *Science* 356 (6336), 398-406.

Chetty, R., Hendren, N., Kline, P., Saez, E., 2014. Where is the land of opportunities? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics* 129 (3), 1553-1623.

Chetty, R., Hendren, N., 2018a. The effects of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics* 133 (3), 1107-1162.

Chetty, R., Hendren, N., 2018b. The effects of neighborhoods on intergenerational mobility II: County level estimates. *Quarterly Journal of Economics* 133 (3), 1663-1228.

Chetty, R., Hendren, N., Jones, M. R., Porter, S. R., 2019. Race and Economic Opportunity in the United States: An Intergenerational Perspective. *Quarterly Journal of Economics* forthcoming.

Chetty, R., Hendren, N., Katz, L., 2016. The effects of exposure to better neighborhoods on children: New evidence from the moving opportunity experiment. *American Economic Review* 106 (4), 855-902.

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.

Choudhary, M. A., Jain, A., 2017. Finance and Inequality: The Distributional Impacts of Bank Credit Rationing. International Finance Discussion Papers No. 1211.

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.

Corak, M., 2013. Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspective* 27 (3), 79-102.

Crépon, B, Devoto, F., Duflo, E., Parienté, W., 2015. Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *American Economic Journal: Applied Economics* 7 (1), 123-150.

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.

Demigrüç-Kunt, A., Levine, R., 2009. Finance and inequality: Theory and evidence. *Annual Review of Financial Economics* 1, 287-318.

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.

Duflo, E., Banerjee, A., 2014. Do firms want to borrow more? Testing credit constraints using a direct lending program. *Review of Economic Studies* 81, 572-607.

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., Moav, O. 2004. From physical to human capital accumulation: inequality and the process of development. *Review of Economic Studies* 71 (4), 1001-1026.

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.

Iyer, R., Puri, M., 2012. Understanding bank runs: The importance of depositor-bank relationships and networks. *American Economic Review* 102 (4), 1414-1445.

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.

Kearney, M. S., Levine, P. B., 2016. Income Inequality, social mobility, and the decision to drop out of high school. *Brookings Papers on Economic Activity*, Spring 2016, 333-380.

Kim, D. H., Lin, S. C., 2011. Nonlinearity in the financial development-income inequality nexus. *Journal of Comparative Economics* 39 (3), 310-325.

Klein, M. W., Peek, J., Rosengren, E. S., 2002. Troubled banks, impaired foreign direct investment: The role of relative access to credit. *American Economic Review* 93 (3), 664-682.

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.

Larsen, A. C., 2013. The rise and fall of social cohesion: The construction and de-construction of social trust in the US, UK, Sweden and Denmark. Oxford University Press.

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.

Minetti, R., Murro, P., Peruzzi, V., 2019. One size does not fit all. Cooperative banking and income inequality. CASMEF Working Paper No. 2, September 2019.

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

Mookherjee, D., Ray, D., 2003. Persistent inequality. *Review of Economic Studies* 70 (2), 369-393.

Moser, C., Saidi, F., Wirth, B, 2018. The effects of credit supply on wage inequality between and within firms. 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.

Persson, T., Tabellini, G., 1994. Ins inequality harmful for growth? *American Economic Review* 84 (3), 600-621.

Piketty, T., 1997. The dynamics of wealth distribution and the interest rate with credit rationing. *Review of Economic Studies* 64 (2), 173-189.

Piketty, T., Saez, E., 2013. Top incomes and the great recession: Recent evolutions and policy implications. *IMF Economic Review* 61 (3), 456-478.

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

Putnam, R., 2000. Bowling alone: The collapse and revival of American community. Simon & Schuster Paperbacks.

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.

Saez, E., Matsaganis, M., Tsakloglou, P., 2012. Earnings determination and taxes: Evidence from a cohort-based payroll tax reform in Greece. *Quarterly Journal of Economics* 127, 493-533.

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.

Sitglitz, J., 2012. The price of inequality: How today's divided society endangers our future. W. W. Norton & Company.

Tarozzi, A., Desai, J., Johnson, K., 2015. The impacts of microcredit: Evidence from Ethiopia. *American Economic Journal: Applied Economics* 7 (1), 54-89.

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 variables	
Income	The euro amount of individuals' total annual income (in log).
Working capital loan	Log of the amount of a working capital facility.
Debt repay	A dummy variable equal to 1 if the borrower is repaying previous loan obligations and 0 otherwise.
ROA	The ratio of firm's net income to total assets.
Firm growth	The annual growth rate of firm assets.
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.
Firm age	The firm's age in years.
Gender	A dummy variable equal to 1 if the applicant is a male and 0 otherwise.
Loan amount	Log of the requested loan amount in thousands of euros.
Maturity	Requested loan duration in months.
Collateral	A dummy variable equal to 1 if the requested loan is secured by collateral and 0 otherwise.
Covenant	A dummy variable equal to 1 if there is one or more covenants associated with the requested loan and 0 otherwise.
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	16.12
Firm leverage	61,863	0.207	0.0249	0.143	0.917
Firm age	61,863	14.20	14.87	0	182
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
Working capital loan	61,863	1.925	0.714	0.679	5.825
ROA	61,863	0.094	0.160	-0.711	0.836
Firm growth	61,863	0.193	0.386	-1.938	6.484

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)
Dependent variable	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)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	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)
Observations	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)	(2)	(3)
Dependent variable	Income t+1	Income t+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)
Observations	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 sample selection in the parametric RDD

The table reports coefficients and standard errors (in parentheses) from a two-stage Heckman model. The first stage models the probability that individuals having a long-lasting exclusive relationship with the bank apply for a loan in a given year (probit model). These individuals correspond to firm owners who do not have a lending relationship with another regulated bank at the time of the loan application and for which we observe at least two loan applications in the sample period, so that we have information on their income at time *t* and in the following years (as reflected in the dependent variable of the second stage). The first stage is estimated on the full sample which consists in an unbalanced panel including all applicants, irrespective of whether they have an exclusive relationship with the bank or not. The second stage is equivalent to the estimation of equation (2) as in columns 4-6 of Table 3, but including the fitted value of the *Mills ratio* (i.e., the instantaneous probability of loan application) 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.

	Second-stage results		
	(1)	(2)	(3)
Dependent variable	Income t+1	Income t+3	Income t+5
Granted	0.0533***	0.0761***	0.0795***
	(0.0179)	(0.0185)	(0.0188)
Credit score	-0.0021	-0.0011	-0.0051
	(0.0311)	(0.0350)	(0.0205)
Granted x Credit score	0.0184	0.0038	0.0087
	(0.0367)	(0.0401)	(0.0233)
Mills ratio	0.9150	0.9683	0.6129
	(1.3962)	(1.3121)	(0.8163)
Observations	53,585	45,333	37,210
Controls as in Table 4	Yes	Yes	Yes
Clustering	Individual	Individual	Individual
	First-stage results		
Dependent variable	Pr. application t	Pr. application t	Pr. application t
Income	0.0739***	0.0767***	0.0781***
	(0.0083)	(0.0083)	(0.0108)
Wealth	0.0580**	0.0625**	0.0642**
	(0.0270)	(0.0305)	(0.0316)
Education	0.0245***	0.0220***	0.0237**
	(0.0072)	(0.0079)	(0.0094)
Firm size	0.0014	0.0026*	0.0034**
	(0.0024)	(0.0015)	(0.0014)
Firm leverage	0.2870***	0.3022**	0.3147**
	(0.0331)	(0.0610)	(0.1103)
Gender	0.0081***	0.0081***	0.0074***
	(0.0023)	(0.0028)	(0.0031)
Observations	228,507	228,507	228,507
Clustering	Individual	Individual	Individual

Table 8

Controlling for sample selection in the non-parametric RDD

The table reports coefficients and standard errors (in parentheses) from a quasi-two-stage Heckman model. The table essentially replicates the analysis of columns 4-6 of Table 5, the difference being the inclusion of the *Mills Ratio* obtained in the first stage regressions of Table 7 as a control variable in the non-parametric RDD estimation. The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the bias-corrected RD estimates with robust variance estimator. The *, **, and *** marks denote statistical significance at the 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.

	Second-stage results		
	(1)	(2)	(3)
Dependent variable	Income t+1	Income t+3	Income t+5
Robust	0.0601***	0.0613***	0.106***
	(0.014)	(0.0163)	(0.0182)
Observations	53,585	45,333	37,210
Eff. obs. left of cutoff	8,203	6,049	4,080
Eff. obs. right of cutoff	8,480	6,261	4,197
BW estimate	62.4	56.13	45.09
BW bias	96.25	87.24	79.11

Table 9
Economic Channels

The table report coefficients and standard errors (in parentheses). The dependent variable is given in the first row of each panel and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the bias-corrected RD estimates with robust variance estimator. All specifications include the control variables as in specifications 4-6 in Table 4. The first two specifications of panel B distinguish between small and large loans, which are identified as the 25th and the 75th percentile of the distribution of *loan amount*, respectively. The last two specifications of panel B distinguish between new and old firms, which are identified as the 25th and the 75th percentile of the distribution of *firm age*, respectively. The dependent variables in panel B consist in various firm outcomes, including the amount of a working capital loan taken to finance short-term operations, a dummy equal to one if the firm is repaying previous loan obligations, the return on asset and the growth rate of the firm. 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).

Panel A. Small vs large loans, new vs old firms				
	Small loans	Large loans	New firms	Old firms
	(1)	(2)	(3)	(4)
Dependent variable	Income t+5	Income t+5	Income t+5	Income t+5
Robust	0.105***	0.118***	0.167***	0.0623***
	(0.0171)	(0.0216)	(0.0386)	(0.0162)
Observations	8,226	3,507	2,727	13,245
Eff. obs. left of cutoff	1,499	403	662	2,015
Eff. obs. right of cutoff	2,022	416	679	2,026
BW estimate	14.69	8.67	10.07	14.55
BW bias	16.52	10.11	12.81	17.39

Panel B. Firm outcomes				
	(1)	(2)	(3)	(4)
Dependent variable	Corporate purpose t+5	Debt repay t+5	ROA t+5	Firm growth t+5
Robust	0.131***	0.048**	0.048**	0.035***
	(0.019)	(0.022)	(0.0207)	(0.0118)
Observations	27,628	7,311	41,391	41,391
Eff. obs. left of cutoff	5,211	1,361	4,815	4,927
Eff. obs. right of cutoff	5,440	1,407	5,003	5,093
BW estimate	20.6	13.24	61.27	67.91
BW bias	22.46	15.72	95.16	107.18

Table 10
The role of soft information

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 bias-corrected RD estimates with robust variance estimator. All specifications include the control variables as in specifications 4-6 in Table 4. The table replicates the analysis of columns 4-5 of Table 5 on different subsamples depending on the residuals of a linear regression of the credit score on a set of observables (income, wealth, education, firm size, leverage, loan amount, maturity, and two dummies reflecting the use of collateral and covenants) capturing hard information. The residuals of this regressions are interpreted as soft information held by the bank. Specifications 1 to 3 are estimated on the subsample where the residuals are positive and specifications 4-6 where the residual are negative or zero. 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.

	Residuals>0			Residuals≤0		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
Robust	0.0764***	0.0595**	0.135***	0.0856***	0.0391	0.0695*
	(0.0244)	(0.0234)	(0.0293)	(0.0319)	(0.0318)	(0.0378)
Observations	30,998	27,016	23,136	26,768	22,498	18,255
Eff. obs. left of cutoff	4,649	3,927	2,549	3,748	3,375	2,373
Eff. obs. right of cutoff	4,937	4,118	2,720	4,549	3,373	2,556
BW estimate	56.13	54.27	47.11	54.20	52.29	41.28
BW bias	94.29	93.18	79.26	92.16	90.25	76.64

Table 11
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 12
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 bias-corrected RD estimates with robust variance estimator. All specifications include the control variables as in specifications 4-6 in Table 4. 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)
Dependent variable	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)
Observations	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 13
Pre-post crisis

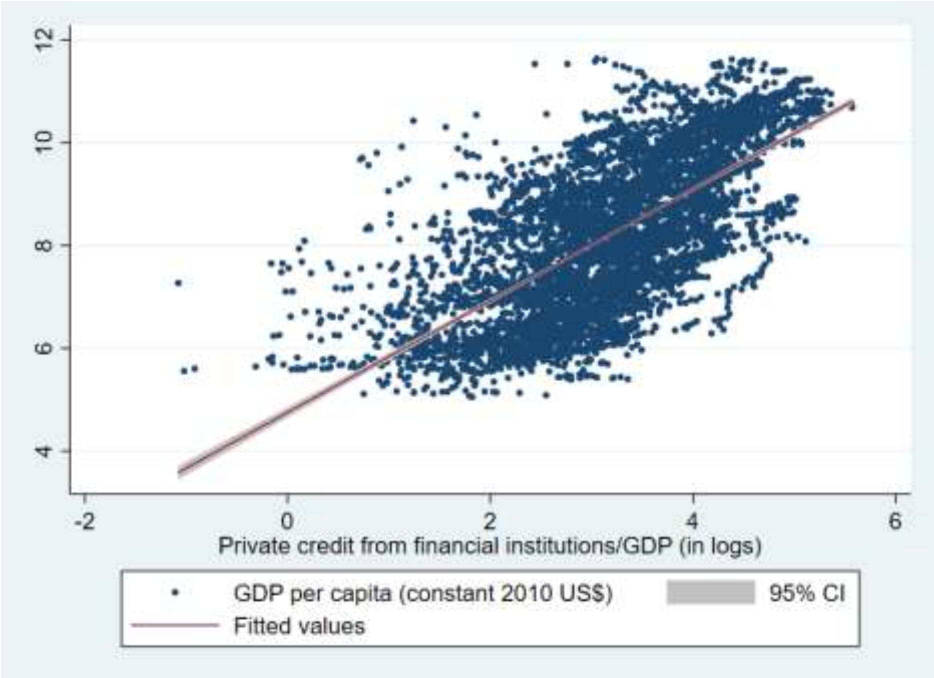
The table reports coefficients and standard errors (in parentheses). The dependent variable is given in the first row of the table and all variables are defined in Table 1. Estimation method is the local linear regression with triangular kernel. For each specification, we report the bias-corrected RD 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)
Dependent variable	Income t+1	Income t+3	Income t+5	Income t+1	Income t+3	Income t+5
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)
Observations	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.



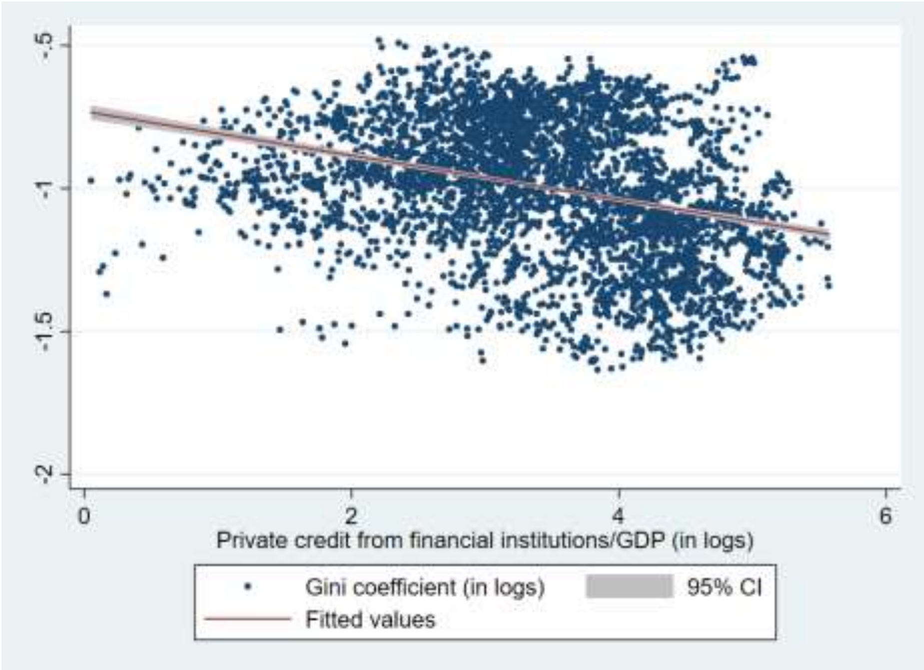
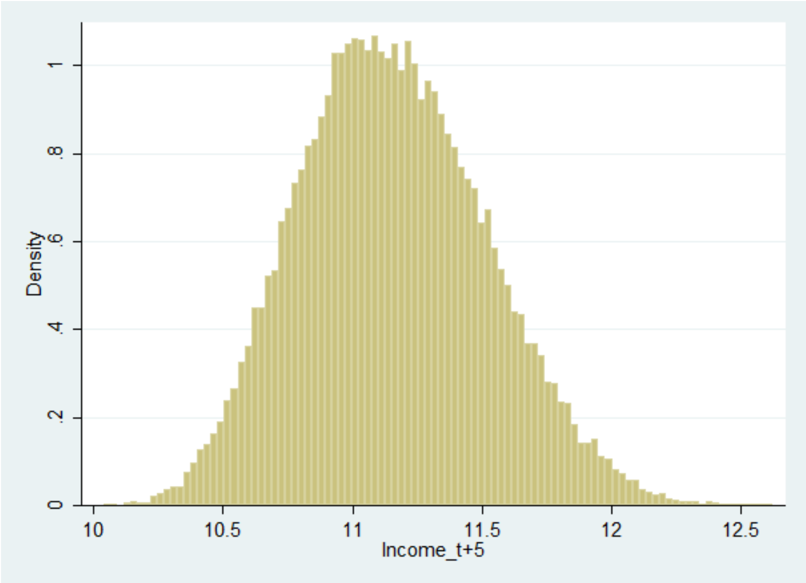


Figure 2
Densities of outcome and assignment variables

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



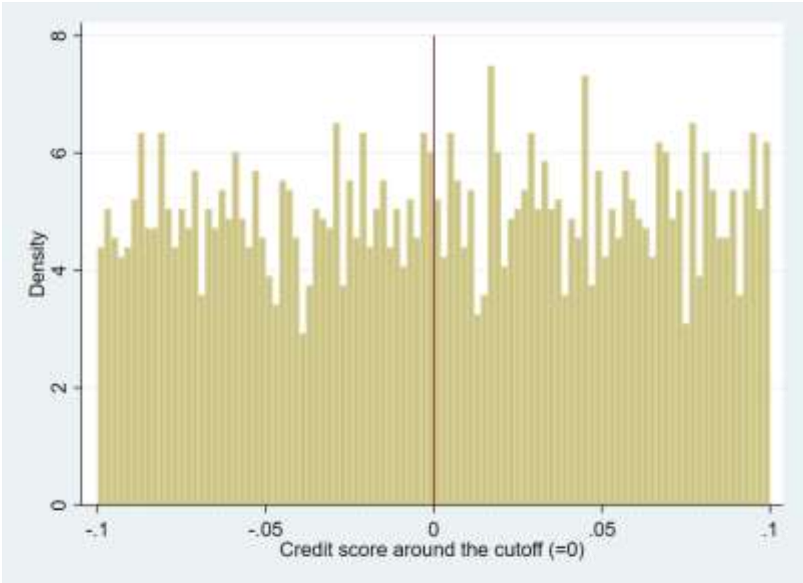


Figure 3
Manipulation test

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.

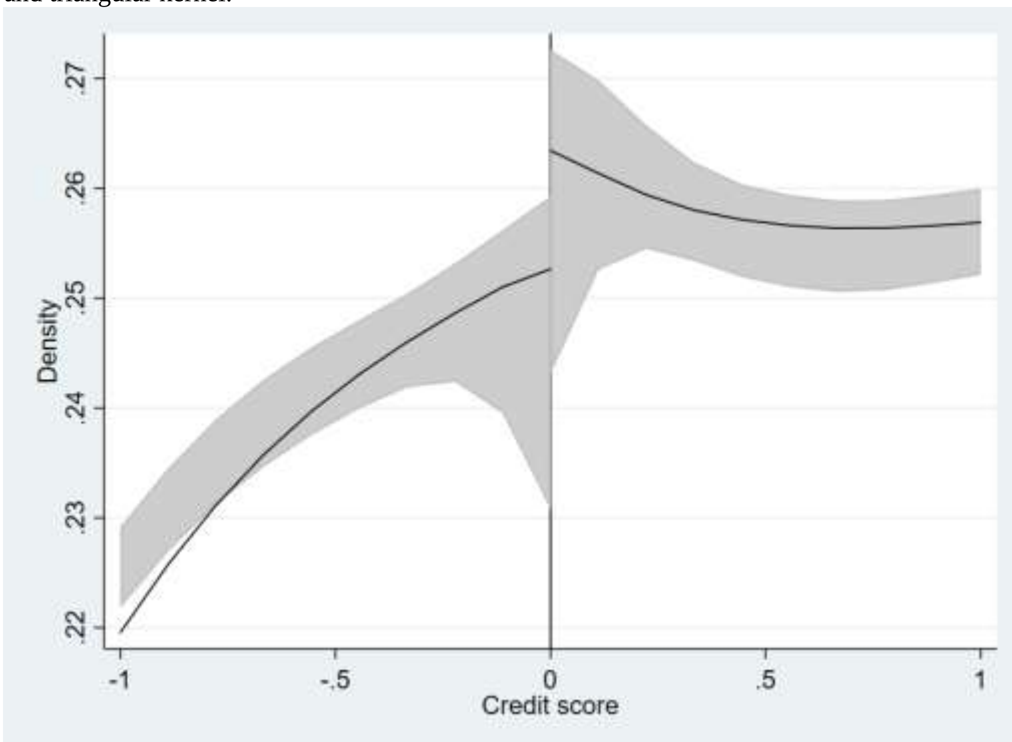


Figure 4
Applicants' income around the cutoff

The figure depicts applicants' Income five years after the loan decision (y-axis) against the Credit score (x-axis). The points represent local sample means of the applicant's income for a set of disjoint bins of control and treatment units spanning the full sample. We select evenly spaced bins that mimic the underlying variability of the data using spacings estimators. The continuous line represents a fourth order polynomial fit used to approximate the conditional mean of applicants' income below and above the cutoff.

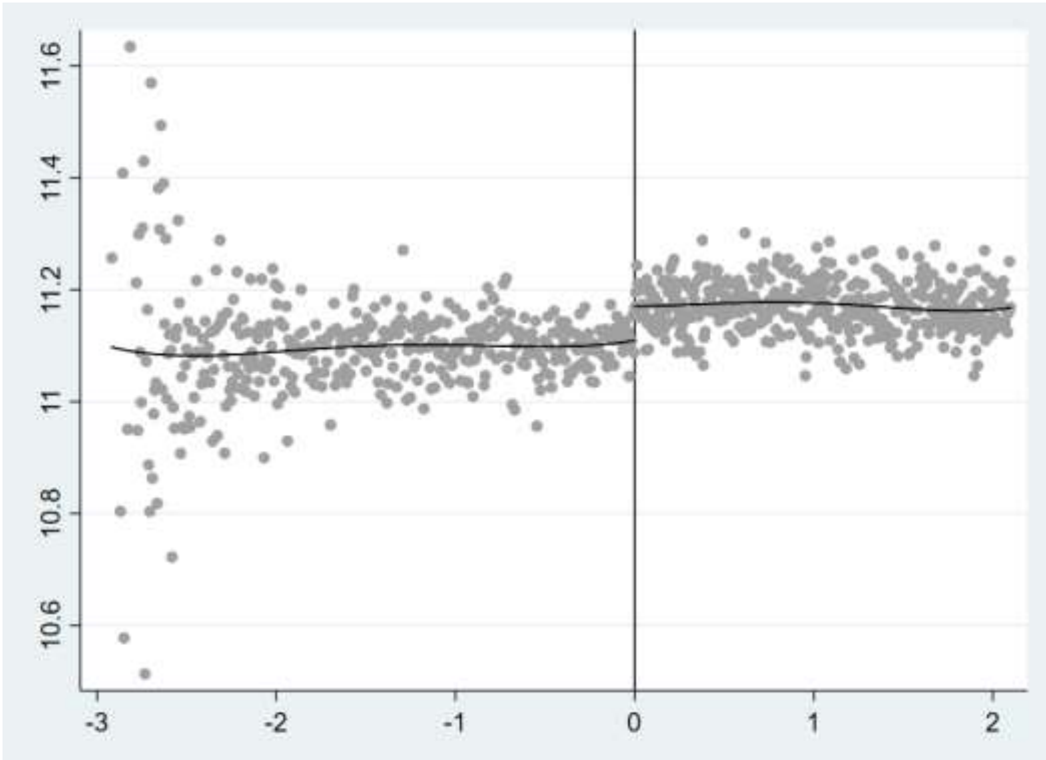
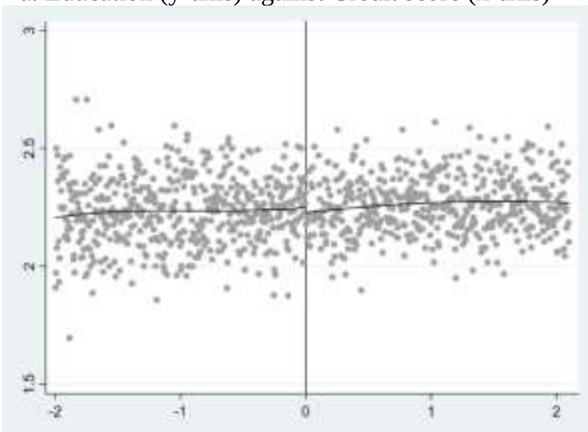


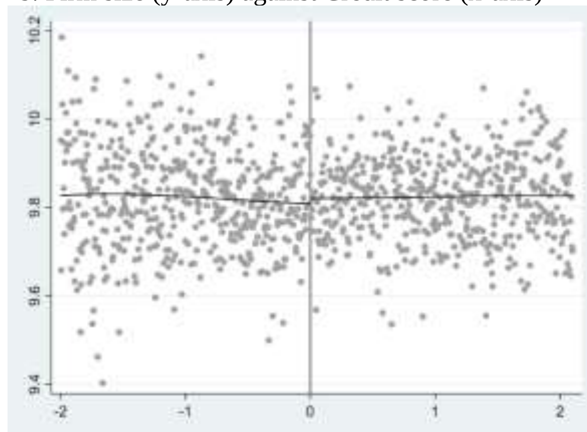
Figure 5
Covariates around the cutoff

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.

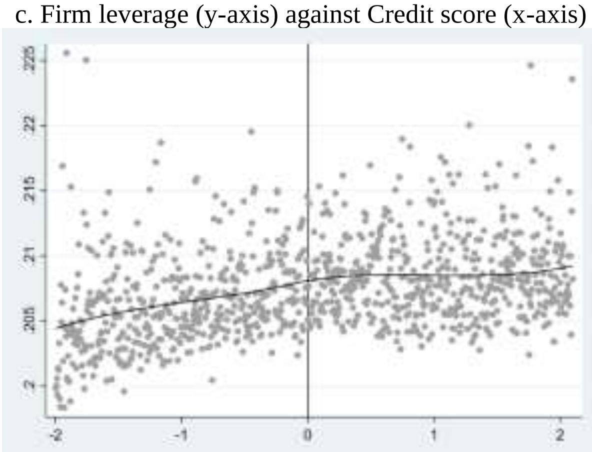
a. Education (y-axis) against Credit score (x-axis)



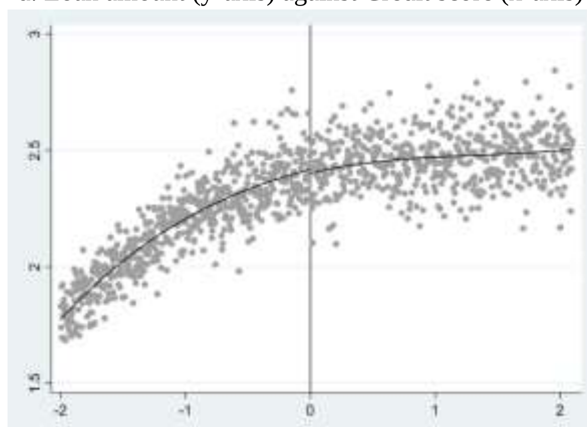
b. Firm size (y-axis) against Credit score (x-axis)



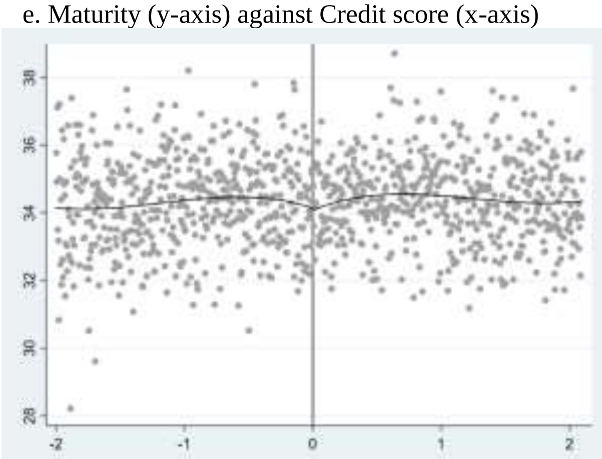
c. Firm leverage (y-axis) against Credit score (x-axis)



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)

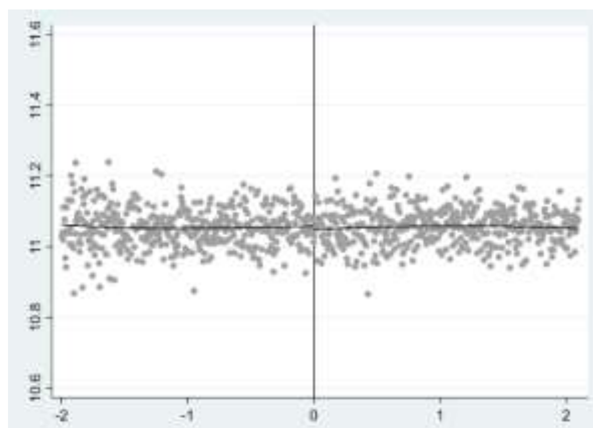


Figure 6
Sensitivity analysis for the RDD

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.

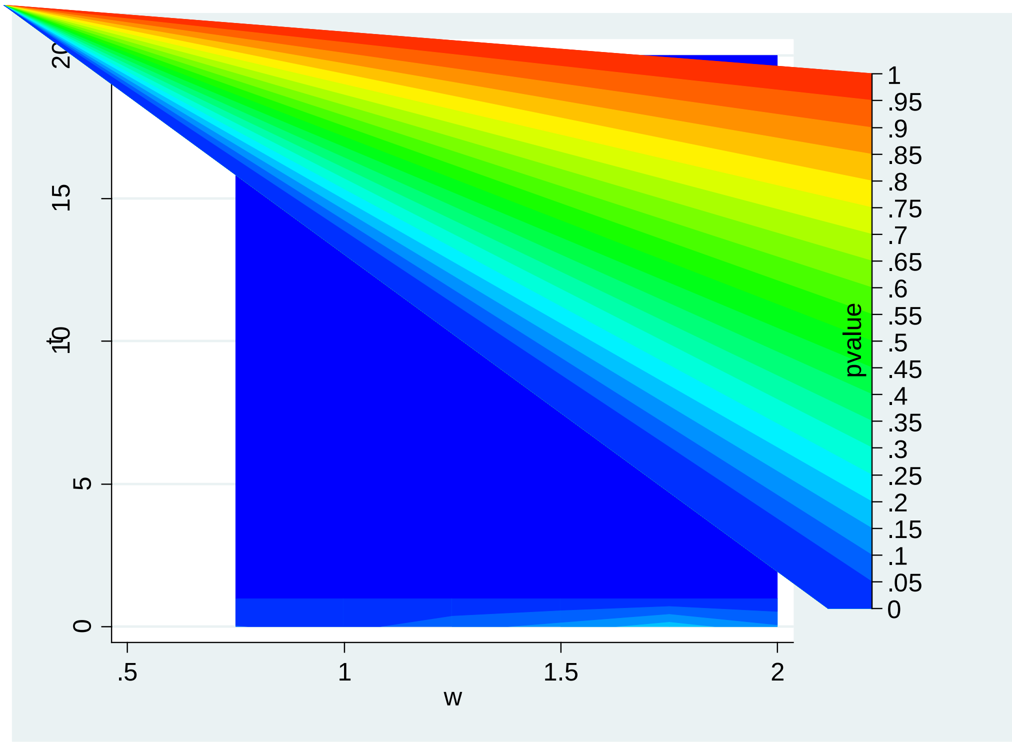
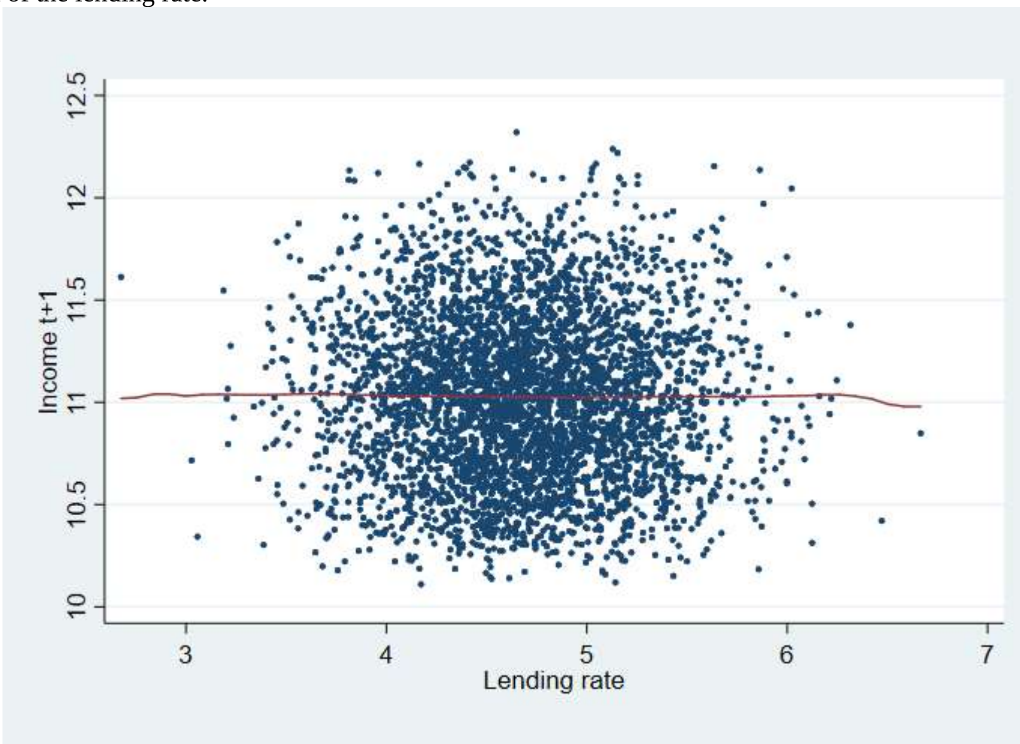


Figure 7
Applicants' income and lending rate around the cutoff

The figure depicts applicants' Income one year after the loan decision (y-axis) against the Lending rate (x-axis). The points represent local sample means of the applicant's income for a set of disjoint bins of control and treatment units spanning the restricted sample where we estimate the non-parametric RDD of Table 7. We select evenly spaced bins that mimic the underlying variability of the data using spacings estimators. The continuous line represents a local polynomial smoother of order zero (i.e. local mean smoother) used to approximate the mean of applicants' income as a function of the lending rate.



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 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)
Dependent variable	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)
Dependent variable	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)
Observations	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)
Dependent variable	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.