Education and Credit: A Matthew Effect

Manthos D. Delis
Montpellier Business School

Yota D. Deli
University College Dublin

José-Luis Peydró
Imperial College London and *Universitat Pompeu Fabra*

Adele Whelan
Economic and Social Research Institute and *Trinity College Dublin*

For valuable comments and suggestions, we thank Hans Degryse, Paul Devereaux, Ben Elsner, Stefanie Haller, Seamus McGuinness, Oana Peia, Karl Whelan, and seminar participants at University College Dublin (UCD), the Economic and Social Research Institute (ESRI), the Irish Economics Association (IEA), Beijing University of Technology, 5th Benelux Banking Research Day (KU Leuven), RCEA Conference on Recent Developments in Economics, Econometrics and Finance (Rimini Centre for Economic Analysis), the 7th LEER Conference on Education Economics (KU Leuven), and the IMAEF conference of University of Ioannina.

Delis, M., D.: Montpellier Business School, Montpellier Research in Management, 2300 Avenue des Moulins, 34000, Montpellier, France. Email: m.delis@montpellier-bs.com; Deli, Y., D.: Economics Department, University College Dublin, School of Economics, Newman Building, Belfield, Dublin 4. Email: yota.deli@ucd.ie; Peydró, JL.: Imperial College London, Exhibition Rd., South Kensington, London SW7 2BX, United Kingdom and Catedràtic d'Economia, Universitat Pompeu Fabra, Plaça de la Mercè, 10-12, 08002 Barcelona, Spain. Email: j.peydro-alcalde@imperial.ac.uk. Whelan, A.: Economic and Social Research Institute, Whitaker Square, Sir John Rogerson's Quay, Dublin 2. Email: adele.whelan@esri.ie.

Education and Credit: A Matthew Effect

Abstract

We examine how the credit channel affects the relation between educational attainment of firms' owners and real outcomes. Using a sharp discontinuity created by the bank's credit score and the associated loan origination decision, we find that entrepreneurs who obtain university education have higher future income and wealth, while their firms enjoy higher profitability. The triggering factors of these outcomes are that these entrepreneurs have higher probability to apply for and granted a loan, and invest in more innovative projects that require better-paid employees. Overall, the initial advantage of university education is self-amplifying via the credit channel (a "Matthew Effect").

***Keywords*:** Education; Credit; Higher education; Loan application; Bank credit decisions; Firm performance; Pay Inequality

***JEL Classification*:** G21; G32; I23, I24; I26

1. Introduction

Is there a role for the credit channel in the effect of entrepreneurs' educational attainment on future individual and firm outcomes? What are the key mechanisms underlying such a potential effect? Highly educated and more skilled labor amplifies innovation and exacerbates technological advancements. Accordingly, if education plays a role in decisions to apply for credit by entrepreneurs and grant credit by banks, it can trigger a sequence of events at the managerial and firm levels, ultimately affecting firm performance and the business owners' economic outcomes. This occurs via a standard credit channel mechanism: loan origination generates liquidity and increases investment, which in turn helps firms become more innovative, more profitable, and larger. These effects are especially important for small firms that rely heavily on bank credit and do not usually have access to alternative sources of funding (Berg, 2018; Delis et al., 2021).

We use unique data on loan applications to a large (systemic) European bank with nationwide coverage. We identify entrepreneurs as majority owners of small and micro firms, following the relevant definition from the European Commission (total assets less than €10 million). We observe repeated loan applications from the same applicants and construct a balanced firm-year panel dataset over 2002-2018. Our final dataset includes 137,321 loan applications from 24,712 unique applicants (firms).

For each loan application, we have full information on the business owner's education and credit score, as well as his/her gender, income, wealth, family situation, age, etc.; we also have data on firm characteristics (including financial characteristics and region), loan characteristics (e.g., loan amount, maturity, collateral, purpose), and the bank's loan decision (granted or rejected). To safeguard our analysis from sample selection issues, we report many tests showing that our sample is fully comparable with international averages across bank, firm, and loan applicant characteristics; the bank's acceptance rate is also fully consistent with European averages.

Our empirical analysis covers three stages. Our first hypothesis is that individuals with higher levels of education (university education) are more confident, have a better understanding of the application process, and negotiate the terms of lending more efficiently. Equivalently, the bank considers education in the formation of the credit score and the probability to grant the loan. These are triggering elements of the credit channel: university education increases credit demand and access to credit.

Subsequently, our core question is how educational attainment influences the credit channel's effect on future firm outcomes (i.e., the probability of default, returns, leverage, and entrepreneurs' future income and wealth). Observing the bank's credit score is important at this stage because the distance of this score from the cutoff value (the value above which the bank originates the loan) forms a sharp discontinuity in the bank's credit decision (Lee and Lemieux, 2010; Delis et al., 2021). The key assumption for the validity of our regression discontinuity design (RDD) is that applicants cannot consistently and/or precisely manipulate their credit scores, because the bank is a value-maximizing entity. We show that this assumption holds in our setting with several relevant tests.

Third, we examine the key mechanisms driving our results. Our main hypotheses are that higher educational attainment (university degree and above) accentuates technological differences creating skill premia. Investment decisions for such entrepreneurs are oriented more toward technological innovation (R&D, intangible assets, and patents). Subsequently, within-firm pay inequality (the ratio of the owner's income to the average employee income) is lower because the firm selects high-wage workers (i.e., rising segregation). Thus, we analyze the role of within-firm pay inequality and investments in intangible assets, patents, and R&D

to show what is driving the effects of education on entrepreneurial outcomes via the credit channel.

The results from the initial (triggering) stage show that applicants who obtain a tertiary qualification during our sample period, have a higher probability to apply and obtain a loan. When we consider applicants with professional education (an MBA and/or a Ph.D.), the results become even more vigorous, potentially due to an increase in the negotiation power of these individuals and/or more sophisticated and innovative projects. At this stage, the results are from a staggered differences-in-differences (DID) approach, where identification arises from individuals who obtain higher education within our sample period ("switchers").

Most important, in the second stage, we use the RDD around the distance of the credit score from its cutoff value and find that a positive credit decision from the bank has differential effects according to levels of education, on (i) the future probability of firm default (lower for higher-education entrepreneurs), (ii) firm leverage (higher for higher-education entrepreneurs), (iii) future entrepreneurs' income and wealth (higher for higher-education entrepreneurs), and (iv) future within-firm pay inequality (lower for higher-education entrepreneurs). Importantly, these effects are more pronounced for the professional education group of entrepreneurs. These findings show that the effects identified in the first stage of our analysis (especially the differential probability of loan application and loan origination between higher education and non-higher-education entrepreneurs) trigger real differential effects among the two groups via the credit channel.

In the final stage of our analysis, we pinpoint the key mechanisms driving the real effects of the loan-origination decision. Using our RDD framework, we find differential effects of a positive bank credit decision on the ratio of intangible assets to total assets, the ratio of R&D expenses to total expenses, and the probability of a new patent. All these are considerably higher for those with higher education (and even higher for those with professional education).

Last, we show that asset intangibility and investments in high-skilled labor (low within-firm inequality) almost fully explain how a positive credit decision affects the future returns and wealth of entrepreneurs with higher and professional education. This is not the case for the future returns and wealth of non-higher-education entrepreneurs. Therefore, the combination of increased investments in innovation and lower within-firm pay inequality for entrepreneurs with higher education account for most of the positive impacts that credit origination has on future firm performance and entrepreneurs' wealth. This finding is consistent with Acemoglu (1999) and Song et al. (2019), who suggest that due to technological advancements, firms with rising returns to skill hire higher-paid employees compared to firms with lower returns to skill.

The key implication from our results is that higher education, specifically tertiary qualifications, creates a "Matthew Effect" via the credit channel.[1,2] This term refers to a cumulative advantage, where obtaining higher education increases the probability to apply and obtain a loan, creates differential technological and managerial decisions, and improves future outcomes.

Our paper proceeds as follows. Section 2 discusses the theoretical underpinnings of our study and provides testable hypotheses. Section 3 presents our dataset. Sections 4 and 5 discuss the identification, models, and results of each of the three stages of our analysis, respectively. Section 6 concludes.

2. Theoretical mechanisms and relation to the literature

2.1. The role of entrepreneurs' education in the credit channel

[1] Sociologist Robert K. Merton coined the term "Matthew Effect" to refer to his theory of cumulative advantage in science. The phenomenon was named after a verse in the Gospel of Matthew (13:12), which states that "for whoever hath, to him shall be given, and he shall have more abundance: but whoever hath not, from him shall be taken away even that he hath." Mrázová and Neary (2019) also refer to the "Matthew Effect" when examining selection effects with heterogeneous firms.

[2] From now on we refer to two groups: higher education (i.e., those with higher educational qualifications such as tertiary, MSc, MBA, and Ph.D. degrees) and non-higher-education (i.e., those without higher educational qualifications such as secondary, postsecondary, and non-tertiary education).

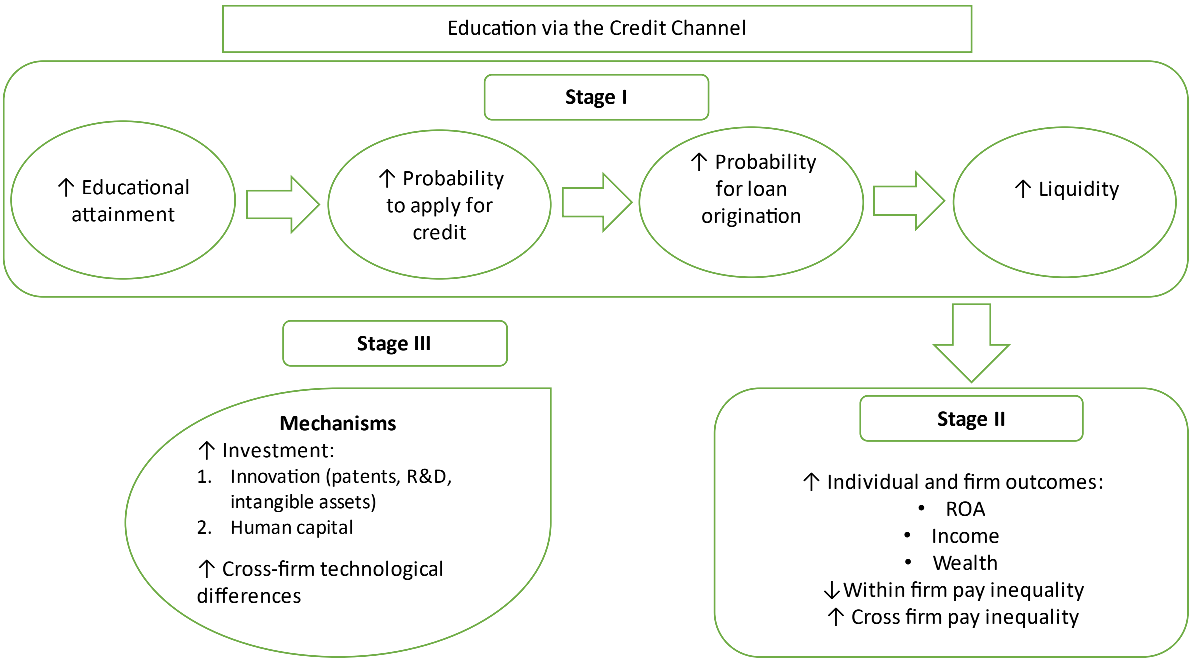
Theoretically, different levels of educational attainment can affect the investment and managerial decisions, subsequently affecting future outcomes via the credit channel. The credit channel operates as follows: a loan origination generates liquidity leading to more investment, which increases firm profitability and entrepreneurs' future income and wealth. To pinpoint the role of education in the credit channel, we identify three key stages for our analysis shown in the chart below:

Stage I. We expect that educational attainment positively affects the entrepreneurs' decision to apply for a loan and the bank's decision on the loan. Over and above their innate ability, higher educated individuals may understand the application process better, have higher levels of self-efficacy, and better negotiate their terms of lending (Zhao et al., 2005; McGee et al., 2009; Nisula and Olander, 2021; Yang and Yang, 2022). Also, we expect that the bank internalizes the applicant's educational attainment in the credit score, as it signals higher ability, affecting the bank's decision to grant credit (Becker, 1993; Spence, 1973; Goodman et al., 2017).

Stages II and III. We expect that entrepreneurs' educational attainment positively affects future firm and individual outcomes via the credit channel. The key reason for this comes from the outcomes of the first stage, whereby entrepreneurs with higher education might be able to apply for credit and granted credit more easily.

We also expect that entrepreneurs with different levels of education make different managerial and investment decisions to exploit the increased liquidity after loan origination via the credit channel. Cross-firm technological differences may affect these decisions, which in small firms are usually decided by the owners-entrepreneurs. Based on the extant literature, the key mechanisms might be different types of investment and within-firm inequality.

Concerning the former, higher education accentuates technological differences, creating a skill premium (Acemoglu, 1999; Acemoglu, xxx). Consistent with the premise that smaller firms are often the most dynamic and innovative (e.g., Klapper et al., 2006), we expect that after a loan is granted, entrepreneurs with higher education invest more in technological-oriented processes (i.e., R&D, intangible assets, and patents). Concerning within-firm pay inequality, we expect that higher-education entrepreneurs receiving credit might invest more in human capital and thus pay higher wages. This can decrease within-firm pay inequality after loan origination, increasing segregation of high-wage employees at firms investing in higher innovation (Song et al., 2019).



2.2. Relation to the extant literature

To our knowledge, our paper is the first to connect education with future firm and individual performance via the credit channel and differential managerial decisions. To this end, we build on two strands of literature. First, our paper adds to the banking literature that assesses how the credit channel affects firm performance. Delis et al. (2020) highlight how loan origination leads to better future firm performance and higher income inequality among small firms. Goodman et al. (2019) and Hartley (2019) connect individual background (education and wealth) to future financial health, showing that education plays an important role in credit score formation. Papadimitri et al. (2020) find that higher education within a firm's board of directors positively affects credit ratings. Marilanta and Nurmi (2018) and Lin et al. (2011) show how educational attainment among entrepreneurs affects their firm's performance.

Second, a substantial amount of literature documents the interplay between technology and education in firm performance. Technological changes affect inequality due to labor demand shifts toward high-skill groups, creating skill premia (see Acemoglu and Autor, 2011 for a review). Card et al. (2013) show that cross-firm wage inequality in Germany rises due to changes in workers' composition. High-wage workers are more likely to work in high-wage firms (increased "sorting") and more likely to work with one another (increased "segregation"). Song et al. (2019) observe a rise in earnings inequality in the United States and attributes one-third of that rise to within-firm pay inequality and two-thirds to cross-firm pay inequality. Acemoglu et al. (2022) focus on managers with business school degrees in the U.S. and Denmark. Their findings show that such managers reduce the wages of their employees within 5 years of their own appointment. However, these firms are not found to experience positive impacts on output, investment, employment, or growth.

3. Data

There is limited panel data on credit access and educational attainment to allow for a systematic examination of individuals over time. We empirically answer our research questions using a unique corporate loans dataset for entrepreneurs applying for loans from a major systemic European bank with nationwide coverage.

3.1. Dataset

The bank from which we obtain the data is an important systemic financial institution according to the European Banking Authority (EBA) definition. We have access to its full loan portfolio, applications, originations, and rejections from 2002 to 2018.[3] We focus on the use of data for loans to domestic small firms and micro firms (total assets of up to €10,000,000 per the EU definition) because we require that loan applicants are majority owners (own more than 50%) of the firm. This is important because otherwise the role of education in credit will be blurred by the education of other owners.

We consider all corporate loan types, including working capital loans, real estate loans, venture loans for start-ups, lines of credit, etc. For each loan application, we have detailed information on key characteristics of the applicant, firm, and loan, including the bank's loan decision (approved or rejected). Importantly, we have access to the applicant's credit score upon which the bank conditions its decision. We also know whether the applicant has an exclusive relationship with the bank. The bank records which firms apply for loans from other regulated and supervised banks (by the European Banking Authority or the country's credit register). Our bank has access to information on the timing of the loan applications and their outcomes. Applicants who have an exclusive relationship with this bank are credit constrained (even from other conventional banks) if our bank rejects their application. Using these data and repeat loan applications from the same applicants, we construct a panel data set of loan applicants over the period 2002–2018.

For most applicants, we observe more than one loan application during our sample period. To compare individuals, it is necessary to observe firm and applicant characteristics at two or more points in time. Thus, we maintain a firm-year balanced panel data set. We discard

[3] The dataset is very similar to the one used by Delis et al. (2020).

loans to applicants who never reapply for loans. Essentially, all individuals (both accepted and rejected ones) reapply for loans within a four-year period. In other words, all observed firms have a relationship with the bank from 2004 onward (the bank has information for the applicants from 2002 onward).[4]

This approach results in a total of 414,730 observations. The panel has more observations than the number of loans because firm owners do not apply for a loan every year. However, the bank continues to hold information on the applicant characteristics after the loan application because when a new application arrives in the future, the bank requests information about applicants' income and wealth retrospectively. Using this information, we generate a panel dataset of 138,633 loan applications by 24,712 unique applicants from 2002 to 2018. From these loan applications, 84.2% were originated (116,753 loans).

In relation to applicant characteristics, we observe age, gender, education, income, wealth, marital status, and the number of dependents, along with their credit score assigned by the bank. Furthermore, we have a large range of firm characteristics such as size, leverage, return on assets (ROA), liquidity, region, and industry. At the loan level, we observe the loan characteristics (i.e., spread, amount, maturity, and collateral).

We define all the variables used in our analysis in Table 1 and report summary statistics in Table 2. For illustration purposes, the mean applicant is close to having tertiary education, is approximately 45 years old, married, and has one or two dependents.

[Please insert Tables 1 & 2 about here]

3.2. Sample representativeness

[4] This comes at the expense of potentially introducing sample selection bias. We show below that running our empirical analysis on the full unbalanced sample or using estimations techniques to deal with this selection does not affect our inferences and in fact strengthens the results in the cases where these are statistically significant. Using the full unbalanced panel implies that we do not have full information (unbalance panel) on certain applicant characteristics (especially income, wealth, and changes in family status) and an observed exclusive bank-firm relationship.

In this section, we provide information on how representative our sample is, to show that the probability of having sample-selection bias is low. We address this possibility further in our empirical analysis. We consider sample representativeness (mostly comparing with European averages) across four dimensions: the bank's characteristics and loan acceptance rates, firm characteristics, loan applicants having an exclusive relationship with the bank, and the entrepreneurs' education level.

The bank operates on a global scale and provides credit to all business types. Using data from a single bank is common practice when detailed data are required (e.g., Delis et al., 2020; Berg, 2018; Iyer and Puri, 2012; Adams et al., 2009). Data on 32 other European systemic banks from Compustat suggests that the annual averages of important bank characteristics like the ratio of liquid assets to total assets, the ratio of market to book value, and return on assets are at very similar levels and significantly correlated with the respective ratios of our bank over the years in our sample (correlation coefficients equal to 0.52, 0.67, and 0.75, respectively). Moreover, data from the Survey on Access to Finance of Enterprises shows that the annual Euro Area average rejection rate is strongly correlated (0.86) with our bank's equivalent. The acceptance rate of 84.2% in our sample is slightly lower than the equivalent reported in the Survey of Access to Finance of Enterprises (SAFE). However, SAFE additionally includes a sample of relatively safer medium-size firms. In a nutshell, our bank's business model is very similar to the European average, which is also documented in Delis et al. (2020).

Second, we compare the firm characteristics. Our sample of small firms closely mimics that of other similarly-sized European firms. Appendix Figure A1 plots the annual average leverage and profitability ratios of small and micro firms in Austria, Belgium, Denmark, France, Germany, and the Netherlands against our sample averages. The data for these countries are from Orbis (information is only available since 2008). The firms in our sample have a 1.1% lower leverage ratio and a 0.76% higher ROA. The trends are very similar and these very small differences are probably due to the fact that our bank operates in a high-income European country and was not significantly affected by the economic downturn in 2010-2014.

Third, for small firms, having an exclusive relationship with a bank is common. This is the case for 65% of the firms in our full sample. This figure is fully consistent with previous studies on multiple or exclusive lending relationships. Berger et al. (2011) document a 71% exclusive relationship between banks and SMEs in three European countries (Germany, Italy, and the UK), but this is less often the case in the United States (Berger et al., 2014, document a 57% rate). Farinha and Santos (2002) report similar statistics for Portugal (70% of firms with fewer than 10 employees have one bank relationship). More recently, Bonfim at al. (2018) report a mean value of two banks for small Portuguese firms, but the Portuguese banking sector is much less concentrated compared to our bank's country. Essentially, the available evidence suggests that the percentage of exclusive relationships in our sample is comparable to previous papers on relationship banking.

A final important issue is the representativeness of business owners with respect to their education levels. In our sample, highly educated entrepreneurs are 50.3% of all loan applicants. An exploration of the EU Labor Force Survey (EU-LFS) Q4 2020, for North-European countries shows that 47.1% of self-employed individuals have higher levels of educational attainment (i.e., tertiary, bachelors, masters, and PhD). These countries, namely Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Netherlands, and the UK, range from 35% to 56% of highly educated self-employed individuals. Our country falls within this range, close to the upper end of the scale, and in line with the country being a relatively rich North European one.

3.3. Key variables

We group entrepreneurs into six levels of education: (i) no secondary; (ii) secondary; (iii) postsecondary/non-tertiary; (iv) tertiary (university); (v) Master of Science degree (MSc); (vi) Master of Business Administration (MBA) or Doctor of Philosophy (Ph.D.). In the last group, the vast majority concerns MBA holders. A key aspect is that 2,711 individuals (*Switchers*) change from non-tertiary (university) education to tertiary education, creating a time-series element that is important for empirical identification.[5]

Table 3 reports summary statistics separately for *Education* and provides a first indication of a "Matthew effect" (i.e., a significant increase in the probability of applying for a loan, increased credit scores, and better firm outcomes as *Education* increases). Also, we observe that entrepreneurs with higher education (university degree and above) get better loan terms (i.e., amount, spread, maturity, and provisions) and their firms are less likely to default.

[Please insert Table 3 about here]

Figure 1 shows the coefficient estimates and confidence intervals in the probability of loan application by education level. The point estimates are for those with: (i) secondary or below; (ii) postsecondary/non-tertiary; (iii) tertiary and MSc; and (vi) MBA or Ph.D. We observe a positive relationship between the probability of receiving a loan application and educational attainment. We observe the most significant increase when comparing higher education to non-higher education applicants.

[Please insert Figure 1 about here]

Credit score is a statistical tool financial institutions construct to determine the credit health of an individual or a firm. In our panel, *Credit score* ranks the entrepreneurs' credit risk; banks use it to decide whether to extend or deny credit, as well as the lending terms. If a credit

[5] When we do not know the precise year of the change (i.e., there is no loan application in two consecutive years), we assume that this change happens in the middle of the time interval between the two loan applications. We make the same assumption for marital status. We also complete the observations with the last credit score calculated by the bank. Thus, if there is a loan application in year t but not one in year t+1, we impute in year t+1 the credit score in year t. Different timing assumptions do not affect our main results.

score is above a specific cutoff point, the bank originates the loan; if a credit score is below this cutoff, the bank denies the loan (or suggests reexamination later). We are not permitted to disclose the precise cutoff, therefore we normalize it to zero.

The *Credit score* is a key control because it encompasses information on all firm-applicant time-varying characteristics (a mix of hard and soft information by the bank). Hard information refers to all information systematically recorded on paper (on the application files). Soft information refers to the residual; what explains the credit score that is not included explicitly on paper. For example, soft information contains the bank's perception of the applicant/firm, quality of the investment idea, the strength of the bank-firm relationship, etc. Any control variable explicitly used in a regression essentially extracts information from the credit score and should thus not affect the adjusted R-squared. This also holds for the education variables. Moreover, controlling for the credit score should substantially increase the adjusted R-squared and strongly limit the possibility of omitted-variable bias affecting the estimates on the education variables in both the staggered DID and the RDD models that we use in our empirical analysis.

3.4. Other control variables

As noted in the previous section, controlling for the credit score essentially renders all other control variables redundant. We show this by including variables containing hard information. We use *Gender*, *Age*, *Income*, *Wealth*, *Marital status*, or *Dependents.* For example, previous research finds that males are more likely to apply for credit than females. Also, entrepreneurs who are younger (on average), married, and with fewer dependents are also more likely to apply for and obtain loans. Further, higher *Wealth* and *Income* are positively correlated with access to education and credit (Morgan and David, 1963; Delis et al., 2020). Finally, we include firm characteristics such as *Size, Leverage, Return on assets* (*ROA*), *Liquidity*, and firm region and industry (Jimenez et al., 2014).

In the third stage of our empirical analysis, we include additional variables to pinpoint the key mechanisms of our main findings. We estimate future within-firm *Pay inequality* as the annual salary of the owner divided by the mean salary of employees (excluding the owner). *Intangible assets* is the ratio of intangible assets to total assets. *R&D expenses* is the ratio of R&D expenses to total expenses. We also use a dummy variable to indicate the probability of a new patent (*Patents*).

4. Stage I: Loan application and origination

4.1. Empirical model and identification

We first examine the relation between education and the probability of loan application, loan origination, and the lending terms (i.e., amount, collateral, and spread). In a preliminary analysis and consistent with Figure 1, we find that higher education plays the most significant role compared to other educational attainment levels. We thus estimate the following model:

$$Apply_{it}(Granted_{it}) = a_0 + a_1 HigherEducation_{it} + a_2 CreditScore_{it} + a_3 x_{i(f)t} + u_{it} \quad (1)$$

Apply is a binary variable taking the value 1 if individual *i* in our sample applies for a loan in year *t* (and 0 otherwise). *Granted* is a binary variable equal to 1 if the bank originates the loan (i.e., the credit score is positive) and 0 if the bank rejects the loan application (i.e., the credit score is negative). *Higher education* is a dummy variable that takes the value 1 if the individual (*i*) has completed higher (tertiary) education and 0 otherwise.

In alternative specifications, we use *Professional education*, which takes the value 1 if the individual (*i*) has completed professional education (MBA/Ph.D.) and the value 0 if that individual has not completed any higher education. *Credit score* is a continuous variable normalized around 0, taking a value above (below) when the bank grants (rejects) the loan application. The vector *x* represents control variables reflecting individual (*i*) or firm (*f*) characteristics. All specifications include individual and year fixed effects.

We estimate equation 1 using a linear probability model, which fares better compared to non-linear models in the presence of several fixed effects. When we estimate equation 1 with *Apply* as our dependent variable, we use the full sample of 414,730 individual-year observations. When *Granted* is our dependent variable, we use the sample of 137,321 granted loan applications (for the rest observations this is obviously missing as individuals have not applied).

Our identification strategy is a staggered DID, considering switchers of education, i.e., individuals who obtain higher education during our sample period (2,711 cases) and thus see a change in *Higher education* from 0 to 1. We do this by including individual fixed effects, which capture any time-invariant applicant characteristics (e.g., innate ability) that might not be fully observed to the bank despite the bank-firm relationship, interviews, etc. Then, our estimates on *Higher education* essentially compare the outcome variables for the same individuals/firms before and after obtaining a university degree. Importantly, as extensively discussed above, controlling for the credit score further limits the possibility of unobservables driving our inferences.

Switchers and non-switchers are very similar in their observable characteristics at the time of the switch; thus, introducing sample selection bias along this dimension is unlikely. For example, the mean values across the two groups on *Apply* are 0.338 vs. 0.335, *Income* 10.99 vs. 10.94, *Wealth* 12.11 vs. 12.07, *Gender* 0.804 vs. 0.803, *Age* 44.98 vs. 44.94, *Marital status* 0.589 vs. 0.589, *Dependents* 1.899 vs. 1.898, *Firm size* 12.896 vs. 12.893, *Leverage* 0.207 vs.

0.206, *ROA* 0.080 vs. 0.079, *Credit score* 0.659 vs. 0.655, *Applications* 6.835 v. 6.844. Obviously, these differences are economically trivial.

A recent literature criticizes the OLS results from staggered DID models because of the potential presence of treatment effect heterogeneity. Callaway and Sant'Anna (2021) consider robust DID estimation in the presence of variation in treatment timing and when the "parallel trends assumption" holds potentially only after conditioning on observed covariates (as in our case with the credit score). We use this approach in a robustness test. Last, on top of the discussion of sample representativeness in section 3.2, we use the full sample of firms in the bank's country, including one-time applicants and other similarly-sized firms in the Orbis database. This is an unbalanced panel. We then estimate a two-stage Heckman model, where in the first stage we estimate the probability that any observed firm in the extended sample applies for credit in our bank. The second stage is a *Granted* specification. This approach further limits the possibility that our baseline sample includes firms that yield selection bias in our estimates.

4.2. Estimation results

Table 4 reports the estimation results from equation 1 (staggered DID model). In all specifications, we control for the credit score, individual and firm characteristics, and the fixed effects noted in the lower part of the table. We cluster the standard errors by individual applicants. As expected, the estimation results show that the *Credit score* plays the most significant role in application decisions and loan origination and increases the adjusted R-squared to the very high levels shown in the lower part of the table. Excluding the credit score yields an adjusted R-squared of 0.42, and further excluding the fixed effects (using only the controls) limits the adjusted R-squared to 0.19. Moreover, the adjusted R-squared remains as reported in Table 4 when we exclude all other control variables. These tests show that the credit score alone adequately encompasses the hard information of the rest of the controls.

In the first column (Panel A), the results show that obtaining *Higher Education* (when previously an individual did not, given the individual fixed effects) has a statistically and economically significant effect on the probability of applying for a loan (1.9 percentage points). This becomes 2.3 percentage points for applicants with a *Professional Education* (MBA/Ph.D.), as reported in column 2.[6] On the same line, the third and fourth columns show that individuals obtaining *Higher education* and *Professional education* have equally a 0.8 percentage-point higher probability of getting a loan.

[Please insert Table 4 about here]

We next use a restricted sample to analyze more closely individuals around the cutoff of the credit score, specifically five percentage-points above and below the zero-cutoff point. These individuals are very similar with respect to their characteristic (as reflected in their credit score), making inferences even tighter. The results in column 5 (Panel B) show that obtaining *Higher education* has a more potent effect on the probability of applying for a loan (2.7 percentage points) compared to our baseline specifications. This becomes 3.5 percentage points for applicants with a *Professional education* (MBA/Ph.D.), as reported in column 6. Furthermore, the results show that individuals obtaining *Higher education* and *Professional education* have 1.7 and 4.1 percentage-point higher probability of being granted a loan, respectively (columns 7 and 8).

Given the potential problem of treatment heterogeneity under the staggered DID models (e.g., Baker, Larcker, and Wang, 2022), we next use the model of Callaway and Sant'Anna (2021). This study considers robust DID estimation in the presence of variation in treatment

[6] For all our results, we run an alternative specification to examine whether the effect is more potent when we combine education with gender. We persistently find no significant effect from the interaction of education with gender. The results are available on request.

timing and when the "parallel trends assumption" holds potentially only after conditioning on observed covariates. We report the results in Panel C of Table 4, noting that Stata does not report the results on control variables from this estimation. The results are economically slightly stronger than our baseline, reflecting that treatment heterogeneity is not an important problem in our panel and, if anything, our baseline results are conservative.

We consider several additional robustness checks. First, to ensure that focusing on switchers appropriately captures the characteristics of our whole sample, we exclude the individual fixed effects (results in appendix Table A1). Second, we use the full unbalanced sample (551,354 observations) of loan applications, thus also including the one-time applicants (results in the first four columns of appendix Table A2). Third, we estimate a two-stage Heckman model, where in the first stage we regress the probability of observing a firm in our loan applications sample from the universe of similarly sized firms in the bank's country that are available in the Orbis database plus the firms in our sample. The first-stage covariates include *Firm size*, *Firm ROA*, *Firm leverage*, and *Firm cash*, as well as the ratio of interest income to total income of our bank (if the firm applies to our bank) vs. the mean of the same ratio of the other major banks in the country. The results are in the last two columns of appendix Table A2 and are equivalent to our baseline. Interestingly, Heckman's lambda is statistically insignificant, implying that our data are consistent with no selection.

As a residual exercise, we consider the effects of *Education* on *Loan amount*, *Loan spread*, and *Collateral*. Panel A of Table 5 shows that higher education significantly lowers the loan spread (by approximately 6 basis points) but does not affect the loan amount or the probability that the loan has collateral. Interestingly, considering individuals obtaining *Professional Education* in panel B, we find that apart from a statistically significant effect on the loan spread, those individuals get loans that are 2% larger (statistically significant at the

5% level). An increase in the negotiation power of these individuals and/or the nature of their projects, which might be more expensive and technologically sophisticated, may potentially explain this result.

[Please insert Table 5 about here]

5. Stages II and III: Future firm and individual outcomes

5.1. Empirical models and identification

The core of our empirical analysis is to consider the effect of education on future outcomes, such as the probability of firm default (*Default*), firm profitability (*ROA*) and leverage, within-firm pay inequality, and individual outcomes such as income and wealth. Our identification strategy comes from the dichotomy between the bank granting or not granting the loan (*Granted* =1 versus *Granted* =0). This dichotomy creates a sharp RDD (e.g., Berg, 2018; Delis et al. 2021). The credit score is the strict tool the bank uses to reach its credit decision; for credit scores above (below) a cutoff point (here normalized to 0), the bank always grants (rejects) the loan. The theoretical channel behind this design is that loan origination generates liquidity and increases firm investment, which in turn increases profitability and decreases the probability of default. The key assumption for the validity of this RDD is that applicants cannot consistently and precisely manipulate their credit scores, because the bank is a value-maximizing entity aiming to minimize non-performing loans. To this end, we estimate the following model:

$$Forward\ outcome_{i,t+3} = a_0 + a_1 Granted_{it} + a_2 x'_{i(f)t} + u_{it}. \quad (2)$$

Forward outcome is either *Default*, *Forward ROA*, *Forward leverage*, *Future pay inequality*, and individual *Future income* and *Wealth*, observed three years after the bank's credit decision (i.e., at *t*+3). The credit score is the assignment (also referred to as "the running" or "the forcing") variable (Imbens and Lemieux, 2008; Lee and Lemieux, 2010).

Equation 2 examines the heterogeneous effect of granting a loan to higher education and non-higher education applicants. Using an RDD with interaction terms to infer heterogeneous effects is not common practice in the related literature; thus, we identify the effect of *Education* by estimating equation 2 twice for each of the two groups (Cattaneo et al., 2021).[7] We use a nonparametric local linear regression, which has the advantage of assigning higher weights to observations closer to the cutoff value of 0. We determine the optimal bandwidth using the approach in Calonico et al. (2014), and for efficient estimation we base our inference on the local-quadratic bias-correction in Calonico et al. (2018) and Cattaneo et al. (2018).

5.2. RDD validation and estimation results

In Figure 2, we provide a graphical representation of the relation between *Credit score* and *Forward ROA* for the full sample of loan applicants (i.e., *Apply* = 1), as well as for the separate samples of applicants with and without higher education. The points represent local sample means of the applicant's ROA for a set of disjointed 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.[8] The continuous line represents a fourth-order polynomial fit used to approximate the conditional mean of applicants' incomes below and above the cutoff. All the figures show clear upward shifts in *Forward ROA*. This suggests that the treatment (*Granted* = 1) entails a sharp discontinuity in both the outcome variables for the

[7] In general, the advantage of using two separate regressions is that the slopes of all the right-hand-side variables are allowed to differ, and this is preferable when these variables have largely different correlations by education. In our context, the two separate regressions have another important advantage. The "rdrobust" Stata tools by Calonico et al. (2014), Cattaneo et al. (2016), Calonico et al. (2018), Cattaneo et al. (2018), and related papers allow identifying the validity of the RDD and produce robust estimates. These imply improved inference and associated transparency. However, these tools come at the expense of flexibility, especially as we cannot introduce interaction terms. In the technically most relevant recent study, Berg (2018) uses a local linear regression and more standard software allowing the regression function to differ on both sides of the cutoff point (see also Lee and Lemieux, 2010, p. 318). Using such an approach does not affect our main inferences (results in Appendix Table A3).

[8] Essentially, these represent the "interesting" bins as selected by the software and not the full set of observations.

full sample and for the separate samples. In that sense, the local linear regression helps with identification, as the family of nonparametric models is better suited to account for nonlinearity.

[Please insert Figure 2 about here]

In Figure 3, we run a manipulation test proposed by Cattaneo et al. (2018). The test uses the local quadratic estimator with cubic bias-correction and a triangular kernel. Consistent with the validity of a sharp RDD, the formal test shows no statistical evidence of manipulation of the assignment variable. This is theoretically plausible because it is highly unlikely that loan applicants systematically manipulate their credit scores.[9] Moreover, all our control variables do not jump at the cutoff (a full set of figures is available on request).

[Please insert Figure 3 about here]

Following the validity tests, we report our baseline RDD results in Table 6. We report the bias-corrected RDD estimates with a conventional variance estimator. The equivalent results with a robust variance estimator are almost the same. For the estimation, the RDD method uses a specific number of observations right and left of the cutoff (reported as effective observations in Table 6); this also implies that the approach is less sensitive to differences in the sample size between those with and without higher education. Columns 1 to 3 report the effects, three years after the bank's decision to grant the loans, on *Default, Future ROA*, and *Future leverage* for individuals with a higher education. Columns 4 to 6 present the equivalent for individuals without higher education; columns 7 to 9 report the results for individuals with professional education.

[Please insert table 6 about here]

[9] Moreover, in the bank's country there is no evidence of fraud in loan applications, not even in the years prior to the global financial crisis.

The estimate in column 1 suggests that a positive credit decision lowers the probability of default for applicants with higher education by a substantial 16.4 percentage points. The equivalent estimate for applicants without higher education (column 4) is an even higher 24.5 percentage points. This eight-point difference is highly statistically significant (at the 1% level) and suggests that applicants without higher education rely much more on loan origination to avert default. Considering applicants with professional education, this difference is even higher at 9.5 percentage points. These findings are fully consistent with our stage I analysis, whereby entrepreneurs with higher and professional education are more likely to apply for a loan (or reapply after being rejected) and get it.

The corresponding effects on *Forward ROA* and *Forward leverage* are even more indicative. We find that a positive credit decision increases *Forward ROA* for applicants with higher (professional) education by 0.06 (0.16) points more than for applicants without higher (professional) education. This is a large difference given the mean average ROA is 0.068 in our sample. Interestingly, the effect of a positive credit decision on *Forward leverage* highlights a different pattern between the groups. Entrepreneurs with higher education are more willing to increase *Future leverage*, with the effect being statistically and economically significant; leverage increases by 1.3 percentage points and is statistically significant (at the 5% level). In contrast, the effect is statistically insignificant for those without higher education. This picture is even more pronounced comparing entrepreneurs with professional education with entrepreneurs without higher education.

Apart from the effects of education on standard firm outcomes, we observe that higher levels of education, through the credit channel, also affect the relative wages of the firm owners compared to the rest of the employees (within-firm pay inequality). We have two ways to capture this wage inequality. First, we observe whether education affects individual *Future income* and *Wealth* through the credit channel. Second, we examine how different levels of education affect the future within-firm *Pay inequality*. Results are in Table 7. Once again, we estimate equation 2 for applicants with and without higher education, as well as for applicants with professional education. The dependent variable for columns 1 and 3 is *Future income*; for columns 2 and 4 is *Future wealth*; and for columns 3 and 6 is *Future pay inequality*.

[Please insert Table 7 about here]

We find that a positive credit decision from the bank leads to a 3.8 (5)-percentage-point increase in income for entrepreneurs with higher (professional) education, whereas the equivalent effect for the applicants without higher education is 2.1 percentage points. Similar differences are observed for *Future wealth,* which is 1.4 (3.5) percentage points higher for those with higher (professional) education. These results are consistent with our premise that less education, via the credit channel, exacerbates income and wealth inequality, contributing to a Matthew effect. Interestingly, from columns 3 and 6 we observe that entrepreneurs with higher education are more likely to reward their employees with salaries closer to their own. We find that loan origination has no significant effect on within-firm pay inequality for entrepreneurs with higher education, whereas the effects are statistically and economically significant for entrepreneurs without higher education. For the latter, we find that future pay inequality increases after the loan origination by 4 percentage points.

Considering that the effect on future income is higher for the high-educated entrepreneurs and the effect on pay inequality is lower for them, it must be that high-educated owners pay significant higher average salaries to their employees. Reestimating equation 2 using the average employees' salary (total personnel expenses to the number of employees) as the outcome variable (this variable is the denominator of the pay inequality ratio), we find that this is indeed the case. Specifically, the coefficient for the higher-education group equals 0.046 (significant at the 1% level), indicating that obtaining higher education leads to a 4.6% increase in the firm's mean salary after loan origination. For the `non high-education group, the equivalent estimate is -0.017 and statistically insignificant.

Last, we run a falsification test based on the RDD of the lagged (at t-1) outcome variables in equation 2. We expect that *Granted* enters with a statistically insignificant coefficient in these regressions, as the treatment effect has not materialized. We report the results on firm outcomes (*Default*, *ROA*, and *Leverage*) in appendix Table A3 and the results on individual outcomes (*Income*, *Wealth*, and *Pay inequality*) in appendix Table A4. As expected, all estimates are statistically insignificant at conventional levels, confirming that any effects are triggered by the treatment at $t = 0$.

5.3. Mechanisms

In this final stage, we examine the mechanisms driving our results. According to our theoretical hypotheses in section 2, we expect that entrepreneurs with higher education undertake different managerial and investment decisions. First, they may invest in innovation capabilities, such as R&D, patents, and intangible assets. In these technological frontier firms, such investments may result in higher future firm performance and individual outcomes after loan origination. Second, consistent with the results in the previous section, entrepreneurs with higher education may hire employees with similar education, creating skill premia in their employees' wages, reducing within-firm pay inequality. These effects might be even more potent considering entrepreneurs with professional education.

To pinpoint these mechanisms, first we re-estimate equation 2 with *Asset intangibility*, *R&D expenses*, and *Patents* as dependent variables. Again, we use our RDD framework. Second, using a similar setup, we estimate *Future ROA* and *Future wealth* equations while controlling for asset intangibility and within-firm-pay inequality to infer their impact on the estimate for *Granted*.

Table 8 reports that firms owned by entrepreneurs without higher education indeed have higher within-firm pay inequality after loan origination, whereas the effect is insignificant for firms owned by entrepreneurs with higher education. This is a first indication consistent with our hypothesis that entrepreneurs with higher education hire employees at wages similar to their own. To further explain this finding, we examine whether entrepreneurs with higher education use credit to invest more in R&D, patents, and intangible assets, which in turn increases their firms' profitability and their own future income and wealth.

[Please insert Table 8 about here]

In column 1 of panel A, we first show that entrepreneurs with higher education who got loans invest, on average, 11 percentage points more in intangible assets than applicants with higher education who did not get a loan. In column 7, the equivalent effect for entrepreneurs with professional education is 13 percentage points. In contrast, the effect for the less educated entrepreneurs (column 4) is statistically insignificant. Also, when we take the difference of the coefficients between columns 1 and 4, we find that entrepreneurs with higher education invest, on average, 11 percentage points more in intangible assets (the coefficient for non-higher education entrepreneurs in column 4 is statistically insignificant).

Similarly, the results in columns 3 and 6 of panel A show that applicants with higher education who have their loans originated are 8 percentage points more likely to use patents than applicants with higher education who were not granted a loan. There is no significant effect on asset intangibility or patent use for applicants without higher education, indicating that they do not direct more credit toward innovation after a loan origination. The effect of loan origination on *R&D expenses* is positive for entrepreneurs with and without higher education, but again the effect is stronger for the higher-education group (10 percentage points versus 6 percentage points, respectively). Importantly, the equivalent differences between the professional education and no tertiary education groups are even more pronounced, which pinpoints that moving to higher and more sophisticated forms of education explains firm performance-related outcomes via the credit channel.

Next, we examine how *Granted* affects firm and individual outcomes (Table 8, panel B) by directly controlling within the RDD for *Asset intangibilit*y and *Within-firm pay inequality* (separately and combined) to examine their impact on the coefficient on *Granted*. In specifications 1 to 6, we first replicate the results in Tables 6 and 7 for illustrative purposes. Next, in specifications 7 to 18, we find that sequentially adding these controls significantly lowers the impact of *Granted* on *Future ROA* and *Future wealth* for the higher-education entrepreneurs and the professional-education entrepreneurs. Adding both controls (specifications 19 to 24) accounts for almost all the statistically significant impact of *Granted* in the higher-education and professional-education groups. For higher education, the relevant coefficient falls from 0.067 (0.031) in the *Future ROA* (*Future wealth*) specification without these controls to 0.035 (0.021) in the specification with both controls. The estimates in specifications 19 and 20 are barely statistically significant at the 10% level or insignificant, and the original estimates without the controls in specifications 1 and 2 are statistically significant at the 1% level. The results draw a very similar picture for the entrepreneurs with professional education (specifications 23 and 24).

Evidently, this is not the case for those without higher education (as shown on the right-hand-side specifications of panel B). In these specifications, adding *Asset intangibilit*y and *Within-firm pay inequality* in the baseline specifications does not lower the coefficient on *Granted* as much. Comparing the results in columns 15 and 16 to those in columns 3 and 4, we find only small reductions in the economic and statistical significance of the coefficients on *Granted*. In a nutshell, a key driver of the significantly higher firm *Future ROA* and individual *Future wealth* for entrepreneurs with higher education are investments in intangible assets and lower within-firm pay inequality financed through loan origination. These findings highlight how differences in entrepreneurs' educational attainment generates higher income and wealth differences via the credit channel, whereby investment in intangible assets and high-quality employees play a key role.

6. Conclusions

This paper examines how educational attainment affects real firm and individual outcomes, via the credit channel. Our analysis uses a unique sample of corporate bank loans to majority owners of small firms and microenterprises from a major European bank. Our empirical identification exploits the sharp discontinuity generated by the bank's credit score, in accepting or rejecting the loans applications.

We find changes in the credit channel initiated by entrepreneurs obtaining higher education during our sample period: higher probability of loan application, higher probability to grant the loan by the bank, and better lending terms. These translate to significantly enhanced future firm outcomes (firm profitability, probability of default, leverage), and higher future individual income and wealth. Our results highlight a Matthew effect, where the initial advantage of higher education magnifies over time to produce greater firm and individual outcomes, via the credit channel.

We identify that the key mechanisms driving our findings are the differential managerial and investment decisions by highly educated entrepreneurs, which accentuate cross-firm technological differences and within-firm pay inequalities. Investment decisions for highly educated entrepreneurs are increasingly oriented towards technological innovation (R&D, intangible assets, and patents). Equivalently, their managerial decisions focus on investments in human capital and selecting higher-wage workers (i.e., rising segregation).

References

Acemoglu, D., 1999. "Changes in unemployment and wage inequality: An alternative theory and some evidence." *American Economic Review* 89(5), 1259–1278.

Acemoglu, Daron and David H. Autor., 2011. "Skills, tasks and technologies: Implications for employment and earnings." Handbook of Labor Economics 4B, 1043–1171.

Acemoglu, D., He, A. and le Maire, D., 2022. "Eclipse of rent-sharing: The effects of managers' business education on wages and the labor share in the US and Denmark." *National Bureau of Economic Research*, No. w29874.

Adams, R.B. and Ferreira, D., 2009. "Women in the boardroom and their impact on governance and performance." *Journal of Financial Economics* 94(2), 291–309.

Becker, G.S., 1993. "Human capital: A theoretical and empirical analysis." *Reading in Labor Economics and Labor Relations*. 3rd ed. USA: Prentice-Hall.

Berg, T., 2018. "Got rejected? Real effects of not getting a loan." *Review of Financial Studies* 31, 4912–4957.

Berger, A.N., Goulding, W. and Rice, T., 2014. "Do small businesses still prefer community banks?" *Journal of Banking & Finance* 44, 264–278.

Berger A.N., Schaeck, K., 2011. "Small and medium-sized enterprises, bank relationship strength, and the use of venture capital." *Journal of Money, Credit and Banking* 43, 461–490.

Bonfim, D., Dai, Q. and Franco, F., 2018. "The number of bank relationships and borrowing costs: The role of information asymmetries." *Journal of Empirical Finance* 46, 191–209.

Callaway, B. and Sant'Anna, P.H.C., 2021. Difference-in-differences with multiple time periods. *Journal of Econometrics* 225, 200-230.

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* 113, 767–779.

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

Card, D., J. Heining, and P. Kline, 2013. "Workplace heterogeneity and the rise of West German wage inequality." *The Quarterly Journal of Economics* 128(3), 967–1015.

Cattaneo, M.D., Jansson, M., Ma, X., 2018. "Manipulation testing based on density discontinuity." *Stata Journal* 18, 234–261.

Cattaneo, M.D., Keele, L., Titiunik, R., 2021. Covariate Adjustment in regression discontinuity designs. Available at: https://arxiv.org/abs/2110.08410.

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

Delis M. Hasan I., Iosifidi M., Ongena S., 2020. "Gender, credit, and firm outcomes." *Journal of Financial and Quantitative Analysis* 1–31.

Delis M. D., Fringuelotti F., Ongena S., 2021. "Credit, income, and inequality." *FRB of New York Staff report* No 929.

Duranton, G. & Turner, M. A., 2012. "Urban growth and transportation." *Review of Economic Studies* 79(4), 1407–1440.

Farinha, L.A. and Santos, J.A., 2002. "Switching from single to multiple bank lending relationships: Determinants and implications." *Journal of Financial Intermediation* 11(2), 124–151.

Goodman, S., Henriques, A. & Mezza, A., 2017. "Where credit is due: The relationship between family background and credit health." *Finance and Economics Discussion Series* 32.

Hartley, D., Mazumder, B. and Rajan, A., 2019. "How similar are credit scores across generations? ." *Chicago Fed Letter*.

Huang J., Kisgen D. J., 2013. "Gender and corporate finance: Are male executives overconfident relative to female executives?" *Journal of Financial Economics* 108, 822–839.

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

Iyer, R. and Puri, M., 2012. "Understanding bank runs: The importance of depositor-bank relationships and networks." *American Economic Review* 102(4), 1414–45.

Jiménez, G., Ongena, S., Peydró, J.L. and 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.

Klapper, L., Laeven, L. & Rajan, R., 2006. "Entry regulation as a barrier to entrepreneurship." *Journal of Financial Economics* 82(3), 591–629.

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

Lin C., Lin P, Song F.M., Li C., 2011. "Managerial incentives, CEO characteristics and corporate innovation in China's private sector." *Journal of Comparative Economics* 39, 176–190.

Marilanta M., and Nurmi S., 2018. "Business owners, employees, and firm performance." *Small Business Economics* 52, 111–129.

McGee, J. E., Peterson, M., Mueller, S. L. & Sequeira, J. M., 2009. "Entrepreneurial self–efficacy: Refining the measure." *Entrepreneurship Theory and Practice*, 33(4), 965–988.

Morgan J., David M., 1963. "Education and income." *The Quarterly Journal of Economics*, 77, 3, 423–437.

Mrázová, M. & Neary, J. P., 2019. "Selection effects with heterogeneous firms.” *Journal of the European Economic Association* 17(4), 1294–1334.

Nisula, A.M. and Olander, H., 2021. "The role of creativity in knowledge workers’ entrepreneurial intentions: The moderating effect of general self-efficacy.” *Journal of Small Business Management*, 1–27.

Papadimitri, P., Pasiouras, F., Tasiou, M. and Ventouri, A., 2020. "The effects of board of directors’ education on firms’ credit ratings.” *Journal of Business Research* 116, 294–313.

Song, J., Price, D.J., Guvenen, F., Bloom, N. and Von Wachter, T., 2019. "Firming up inequality.” *The Quarterly Journal of Economics* 134(1), 1–50.

Spence M., 1973. "Job market signaling.” *The Quarterly Journal of Economics* 87, 3, 355–374.

Yang, F. & Yang, M. M., 2022. "Does cross-cultural experience matter for new venture performance? The moderating role of socio-cognitive traits.” *Journal of Business Research* 138, 38–51.

Zhao, H., Seibert, S. E. & Hills, G. E., 2005. "The mediating role of self-efficacy in the development of entrepreneurial intentions.” *Journal of Applied Psychology* 90(6), 1265.

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 2002-2018 and the loan is either originated (fully or at least 75% of the requested loan amounted) or rejected (bank advises against proceeding with the application, fully rejects, or only originates up to 25% of the requested loan amount). Due to the exclusive relationship, the bank holds information on the applicants even outside the year of loan application.
Year	Our sample covers the period 2002-2019. Applications end in 2018 and we use one more year of firm financial ratios (2019) to examine future firm outcomes.
B. Variables	
Apply	A dummy variable equal to 1 if the individual applied for a loan in a given year and 0 otherwise.
Education	An ordinal variable ranging between 0 and 5 if the individual completed the following education. 0: No secondary; 1: Secondary; 2: Postsecondary, non-tertiary; 3: Tertiary; 4: MSc; 5: MBA or Ph.D.
Higher education	A dummy variable equal to 1 if the individual completed tertiary education or higher (i.e., Education > 2) and 0 otherwise (i.e., Education < 3).
Professional education	A dummy variable equal to 1 if the individual completed MSc/MBA/Ph.D. education (i.e., Education > 3) and 0 if the individual did not complete tertiary education (i.e., Education < 3).
Income	The euro amount of individuals' total annual income (in log) in the year of the loan application and the two years before the application. For the missing years, we input the predicted value of the regression of the last available observation of income on the mean income by region, year, and industry.
Wealth	The euro amount of individuals' total wealth other than the assets of the firm and minus total debt (in log). The bank observes this in the year of the loan application and the two years before the application. For the missing years, we input the predicted value of the regression of the last available observation of wealth on the mean wealth by region, year, and industry.
Gender	A dummy variable equal to 1 if the applicant is a male and 0 otherwise.
Age	The applicant's age.
Marital status	A dummy variable equal to 1 if the applicant is married and 0 otherwise.
Dependents	The number of dependents.
Firm size	Total firm's assets (in log).
Firm leverage	The ratio of firm's total debt to total assets.
Firm ROA	The ratio of firm's after tax profits to total assets.
Firm cash	The ratio of cash holdings to total assets.
Forward ROA	The mean *Firm ROA* in the three years after the year of the loan application.
Forward growth	The mean increase in *Firm size* in the three years after the year of the loan application.
Forward leverage	The mean *Firm leverage* in the three years after the year of the loan application.
Credit score	The credit score of the applicant, as calculated by the bank. There is a 0 cutoff: positive values indicate that the loan is granted, and negative values indicate that the loan is denied.
Applications	The number of applications to the same bank before the current loan application.
Granted	A dummy variable equal to 1 if the loan is originated (Credit score>0) and 0 otherwise (Credit score<0).

Default	A dummy variable equal to 1 if the firm defaults up to three years after the loan origination, and 0 otherwise.
Loan amount	Log of the loan facility amount in thousands of euros.
Loan spread	The difference between the loan rate and the LIBOR (in basis points).
Maturity	Loan maturity in months.
Loan provisions	A dummy variable equal to 1 if the loan has performance-pricing provisions, and 0 otherwise.
Collateral	A dummy variable equal to 1 if the loan has collateral guarantees and 0 otherwise.
Regional education	The share of entrepreneurs with university (or professional) education to total entrepreneurs by region, industry, and year, 15 years before the loan application.

Table 2. Summary statistics

The table reports the number of observations, mean, standard deviation, minimum, and maximum for the variables use in the empirical analysis. The variables are defined in Table 1, except from *Application probability*, which is obtained from the estimation of equation (1).

	Obs.	Mean	St. dev.	Min.	Max.
Panel A: Full sample					
Apply	414,730	0.331	0.471	0	1
Education	414,730	2.997	1.015	0	5
Higher education	414,730	0.503	0.473	0	1
Professional education	414,730	0.109	0.314	0	1
Income	414,730	10.94	0.428	9.734	12.78
Wealth	414,730	12.07	0.615	7.212	14.29
Gender	414,730	0.802	0.399	0	1
Age	414,730	44.94	15.87	20	78
Marital status	414,730	0.589	0.463	0	1
Dependents	414,730	1.898	1.491	0	7
Firm size	414,730	12.89	0.440	9.960	16.12
Leverage	414,730	0.206	0.124	0.123	0.831
ROA	414,730	0.079	0.100	-0.409	0.583
Cash	414,730	0.080	0.033	0.066	0.255
Credit score	414,730	0.652	0.604	-0.773	3.500
Applications	414,730	6.833	1.464	1	9
Granted	137,321	0.845	0.370	0	1
Default	414,730	0.017	0.098	0	1
Loan amount	137,321	3.509	1.988	0.686	11.41
Loan spread	114,641	340.7	246.1	33.45	985.7
Maturity	137,321	47.9	37.29	4	278
Loan provisions	114,641	0.407	0.451	0	1
Collateral	114,641	0.695	0.499	0	1
Regional education (university)	414,730	0.496	0.285	0.388	0.594
Regional education (professional)	414,730	0.193	0.087	0.125	0.256
Application probability	414,730	0.259	0.027	0.140	0.611

Table 3. Means of key variables by level of educational attainment

The table reports the means for key variables of the model per incremental level of educational attainment. The last lines report individuals at each level as a proportion of educational attainment for the total sample and for the sample of the individuals who were granted loans. The variables are defined in Table 1.

	Below secondary	**Secondary**	**Postsecondary/ Non-tertiary**	**Tertiary**	**MSc**	**Ph.D./MBA**
Apply	0.291	0.326	0.328	0.335	0.345	0.348
Income	10.525	10.864	11.946	10.978	10.990	11.000
Wealth	11.722	12.001	12.076	12.102	12.112	12.123
Gender	0.788	0.799	0.802	0.804	0.802	0.803
Age	44.413	44.913	44.937	44.957	44.963	44.928
Marital status	0.592	0.589	0.588	0.589	0.590	0.585
Dependents	1.887	1.893	1.904	1.896	1.847	1.820
Firm size	12.871	12.888	12.896	12.895	12.897	12.905
Leverage	0.201	0.205	0.206	0.207	0.207	0.207
ROA	0.075	0.078	0.079	0.080	0.079	0.080
Cash	0.077	0.079	0.080	0.080	0.080	0.080
Credit score	0.397	0.591	0.655	0.687	0.708	0.729
Applications	6.706	6.813	6.830	6.853	6.843	6.877
Granted	0.820	0.829	0.836	0.861	0.868	0.875
Default	0.018	0.019	0.017	0.017	0.017	0.016
Loan amount	0.763	3.345	3.528	3.601	3.618	3.646
Loan spread	355.32	350.14	352.19	340.20	330.88	331.72
Maturity	43.560	47.454	47.020	47.775	48.042	49.227
Loan provisions	0.465	0.415	0.413	0.407	0.383	0.339
Collateral	0.642	0.695	0.710	0.709	0.608	0.613
Share in the sample (all applications)	0.003	0.209	0.285	0.301	0.093	0.109
Share in the sample (granted)	0.003	0.197	0.248	0.338	0.108	0.106

Table 4. Probability of loan application and positive credit decision

The regressions examine how *Higher education* or *Professional education* affects the probability of applying for a loan (dependent variable is *Apply*) and the probability that the bank grants the loan (dependent variable is *Granted*). The table reports coefficient estimates from OLS estimation and standard errors (in parentheses) clustered by individual. All variables are defined in Table 1. Panel A uses the full sample, whereas Panel B restricts the analysis to +-0.05 points around the zero-cutoff value on the credit score. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

Panel A. Results from the full sample

	1	**2**	**3**	**4**
Dependent variable:	Apply	Apply	Granted	Granted
Higher education	0.019***		0.008***	
	(0.002)		(0.002)	
Professional education		0.023***		0.008***
		(0.002)		(0.003)
Credit score	0.316***	0.320***	0.585***	0.585***
	(0.032)	(0.035)	(0.033)	(0.033)
Observations	414,730	414,730	137,321	137,321
R-squared	0.91	0.91	0.97	0.97
Control variables	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes

Panel B. Results from the sample around the cutoff

	5	**6**	**7**	**8**
Dependent variable:	Apply	Apply	Granted	Granted
Higher education	0.027***		0.017***	
	(0.006)		(0.004)	
Professional education		0.035***		0.041***
		(0.007)		(0.011)
Credit score	0.929***	0.930***	1.442***	1.450***
	(0.108)	(0.108)	(0.156)	(0.153)
Observations	19,063	19,063	6,353	6,353
R-squared	0.95	0.95	0.98	0.98
Control variables	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes

Panel C. Results from the model of Callaway and Sant'Anna (2021)

	9	10	11	12
Dependent variable:	Apply	Apply	Granted	Granted
Higher education	0.021***		0.011***	
	(0.004)		(0.003)	
Professional education		0.029***		0.012***
		(0.005)		(0.004)
Observations	414,730	414,730	137,321	137,321
Control variables + credit score	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes

Table 5. Loan amount, spread, and collateral

The table reports coefficient estimates and standard errors clustered by individual (in parentheses) from the estimation of equations for loan amount, loan spread, and collateral; the dependent variable is noted on the first line of table. In panel A, the main dependent variable is *Higher education* and in panel B *Professional education*. All variables are defined in Table 1. Results are from the sample of originated loans. All specifications are estimated using OLS. The lower part of the table denotes the rest of the control variables (same as in Table 3), fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

Panel A: Higher education

	1	**2**	**3**
Dependent variable:	Loan amount	Loan spread	Collateral
Higher education	0.006	-5.503**	0.001
	(0.0015)	(2.561)	(0.002)
R-squared	0.92	0.94	0.92
Observations	114,641	114,641	114,641

Panel B: Professional education

	4	**5**	**6**
	Loan amount	Loan spread	Collateral
Professional education	0.020**	-7.316**	0.002
	(0.010)	(3.650)	(0.002)
R-squared	0.92	0.94	0.92
Observations	63,053	63,053	63,053
Other controls + Credit score	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes

Table 6. Credit decision, education, and future firm outcomes

The table reports coefficients and standard errors clustered by individual (in parentheses). The dependent variable is on top of each regression, 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 1% and 5% levels. The table includes all the control variables in Table 4.

	1	**2**	**3**	**4**	**5**	**6**
	Applicants with higher education			Applicants without higher education		
Dependent variable:	Default	Future ROA	Future leverage	Default	Future ROA	Future leverage
Granted	-0.164***	0.067***	0.013**	-0.245***	0.061***	0.008
	(0.029)	(0.015)	(0.006)	(0.031)	(0.016)	(0.006)
Observations	75,801	75,801	75,801	61,520	61,520	61,520

	7	**8**	**9**
	Applicants with professional education		
Dependent variable:	Default	Future ROA	Future leverage
Granted	-0.150***	0.077***	0.020***
	(0.038)	(0.023)	(0.006)
Observations	14,556	14,556	14,556

Table 7. Credit decision, education, and future income and wealth

The table reports coefficients and standard errors clustered by individual (in parentheses). The dependent variable is on top of each regression, 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 1% and 5% levels. The table includes all the control variables in Table 4.

	1	2	3	4	5	6
	Applicants with higher education			Applicants without higher education		
Dependent variable:	Future income	Future wealth	Future pay inequality	Future income	Future wealth	Future pay inequality
Granted	0.038***	0.031***	0.016	0.021***	0.017**	0.040***
	(0.011)	(0.013)	(0.012)	(0.008)	(0.007)	(0.013)
Observations	75,801	75,801	75,801	61,520	61,520	61,520

	7	8	9
	Applicants with professional education		
Dependent variable:	Future income	Future wealth	Future pay inequality
Granted	0.050***	0.035***	0.021*
	(0.013)	(0.017)	(0.011)
Observations	14,556	14,556	14,556

Table 8. Higher education, credit decision, and the role of asset intangibility

The table reports coefficients and standard errors clustered by individual (in parentheses). The dependent variable is on top of each regression, and all variables are defined in Table 1. The 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 1% and 5% levels. The table includes all the control variables in Table 4. The number of observations is as in the respective parts of Table 6 for applicants with higher education, applicants without higher education, and applicants with professional education.

Panel A: Effect of the credit decision on asset intangibility, R&D expenses, and patents

	1	2	3	4	5	6
	Applicants with higher education			Applicants without higher education		
Dependent variable:	Asset intangibility	R&D expenses	Patent dummy	Asset intangibility	R&D expenses	Patent dummy
Granted	0.112***	0.098***	0.083***	0.054	0.061**	0.007
	(0.023)	(0.015)	(0.028)	(0.031)	(0.029)	(0.023)

	7	8	9
	Applicants with professional education		
Dependent variable:	Asset intangibility	R&D expenses	Patent dummy
Granted	0.130***	0.152***	0.119***
	(0.028)	(0.029)	(0.040)

Panel B: Heterogeneous effect of the credit decision on firm and individual outcomes due to asset intangibility

	Applicants with higher education		Applicants without higher education		Applicants with professional education	
	Future ROA	Future wealth	Future ROA	Future wealth	Future ROA	Future wealth
	1	2	3	4	5	6
Granted[10]	0.067***	0.031***	0.061***	0.017**	0.077***	0.035***
	(0.015)	(0.013)	(0.016)	(0.007)	(0.023)	(0.017)
	7	8	9	10	11	12
Granted (with Asset intangibility control)	0.048***	0.026**	0.059***	0.016**	0.044**	0.027**
	(0.016)	(0.013)	(0.018)	(0.007)	(0.021)	(0.012)
	13	14	15	16	17	18
Granted (with Pay inequality control)	0.054***	0.024***	0.055***	0.014**	0.059***	0.025**
	(0.016)	(0.013)	(0.019)	(0.007)	(0.020)	(0.011)
	19	20	21	22	23	24
Granted (with Asset intangib. and Pay inequality controls)	0.035*	0.021	0.054***	0.014*	0.029*	0.019
	(0.018)	(0.014)	(0.020)	(0.008)	(0.015)	(0.012)

[10] As seen previously in Tables 8 and 9.

Figure 1. Point increments in education and probability of loan application

The figure reports coefficient estimates and confidence intervals from the estimation of the probability of loan application (as in Table 5) but including four dummy variables for *Education* (*Education* equals 1+2, to *Education* equals 5).

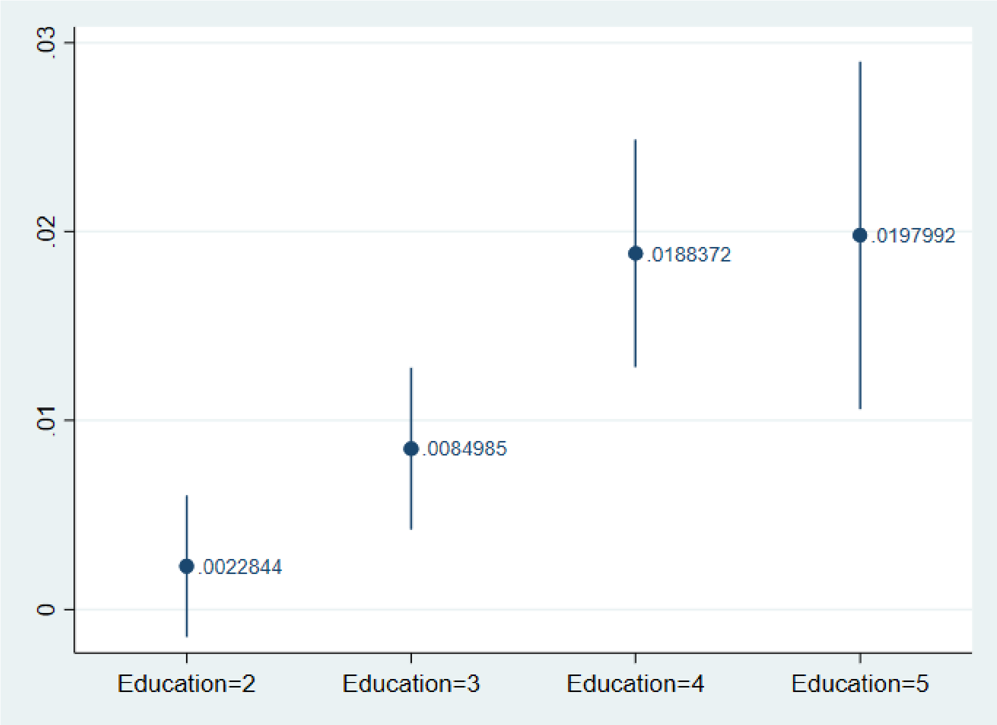
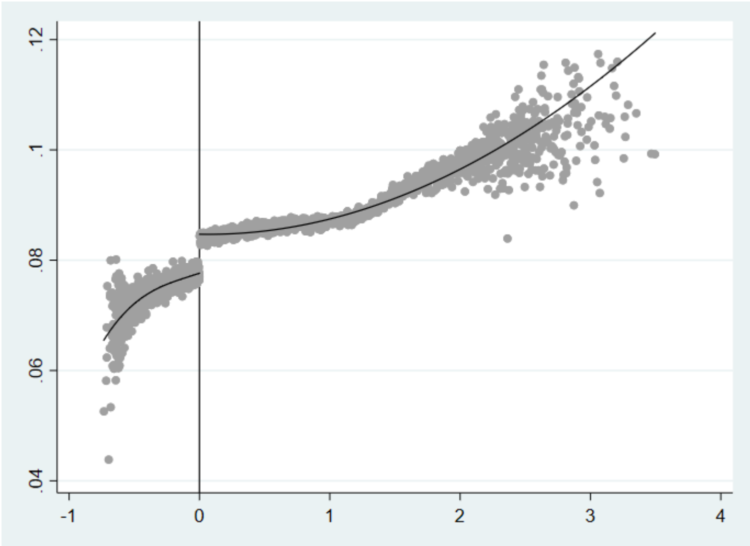
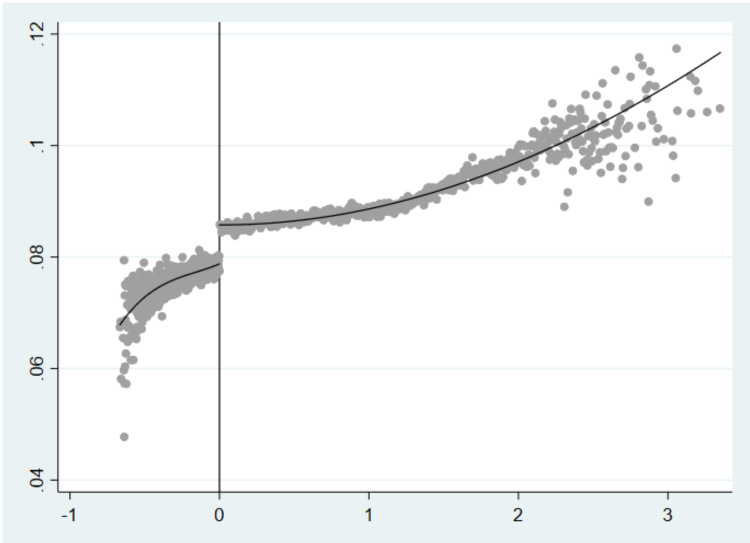


Figure 2. Response of forward ROA at the credit score's cutoff

The figures show the responses of forward ROA (y-axis) at the credit score's cutoff value (=0 on the x-axis). The figure follows Table 11. In particular, the first figure uses the full sample of loan applicants, the second is for applicants with higher education, and the third for applicants without higher education. 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.





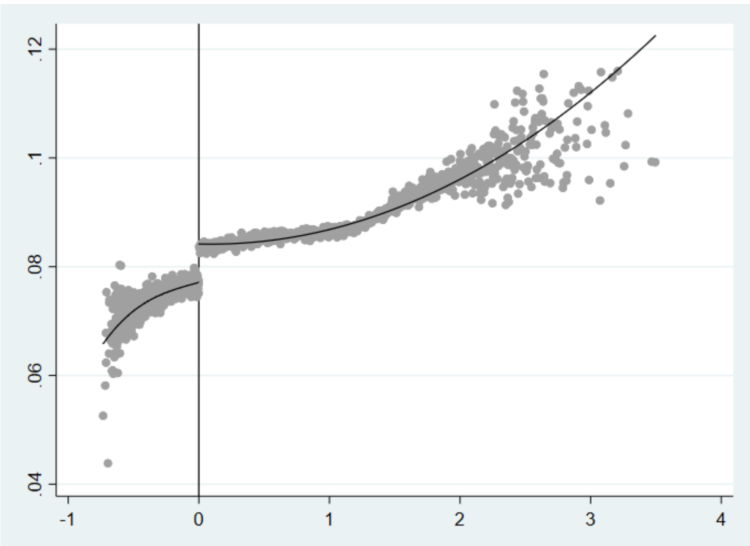
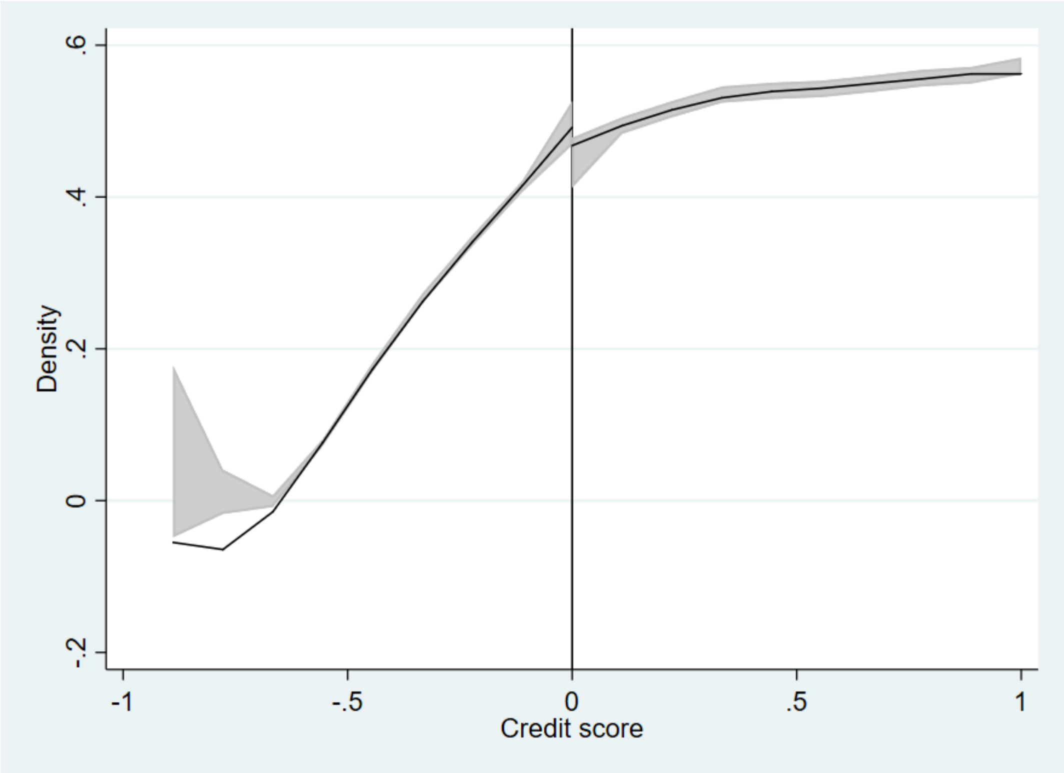


Figure 3. Manipulation test

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.



Appendix
Education and Credit: The Matthew Effect

This appendix, intended for online use only, provides additional robustness tests. Specifically, we replicate the results of Tables 4 to 7 with standard error clustering by region (Table A1) and without individual fixed effects (Table A2). Table A3 reports equivalent results from the full unbalanced sample of loan applications (including one-time applicants) as well as results from Heckman regressions. Figure A1 graphs over time the average leverage and ROA for firms in our sample vs. other European firms.

Table A1. Results without individual fixed effects

This table replicates the regressions of the first panel of Table 4 without including individual fixed effects. The dependent variables are given for every regression, and all variables are defined in Table 1. The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

	1	**2**	**3**	**4**
Dependent variable:	Apply	Apply	Granted	Granted
Higher education	0.020***		0.023***	
	(0.002)		(0.002)	
Professional education		0.009***		0.008***
		(0.002)		(0.003)
Observations	414,730	414,730	137,321	137,321
R-squared	0.91	0.91	0.97	0.97
Control variables	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	No	No

Table A2. Results from all available observations and Heckman regressions

The table reports coefficient estimates and standard errors (in parentheses) clustered by individual. Dependent variable is the binary variable *Apply*, and all variables are defined in Table 1. All specifications include the control variables of Table 4. Specifications 1 to 4 are estimated with OLS, and specifications 5 and 6 with Heckman's model. The dependent variable in the first stage of specifications 5 and 6 is the probability that an observed firm (all firms in the bank's country that are available in Orbis plus the firms observed in our usual panel) appears in our bank's sample. The first stage in these specifications includes *Firm size*, *Firm ROA*, *Firm leverage*, and *Firm cash*, as well as the ratio of interest income to total income of our bank (if the firm applies to our bank) vs. the mean of the same ratio of the other major banks of that country. The lower part of the table denotes the fixed effects, number of observations, and adjusted R-squared (if applicable). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels.

	1	**2**	**3**	**4**	**5**	**6**
Dependent variable:	Apply	Apply	Granted	Granted	Granted	Granted
Higher education	0.020***		0.008***		0.010***	
	(0.002)		(0.002)		(0.003)	
Professional education		0.024***		0.009***		0.011***
		(0.003)		(0.002)		(0.003)
Lambda					-0.169	-0.174
					(0.290)	(0.283)
Observations	551,354	551,354	216,420	216,420	551,354	551,354
R-squared	0.91	0.91	0.97	0.97		
Other controls + Credit score	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	No	Yes	Yes	Yes

Table A3. Credit decision, education, and future firm outcomes: Lagged outcomes

The table reports coefficients and standard errors clustered by individual (in parentheses). The dependent variable is on top of each regression and all of them enter lagged one year. 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 1% and 5% levels. The table includes all the control variables in Table 4.

	1	**2**	**3**	**4**	**5**	**6**
	Applicants with higher education			Applicants without higher education		
Dependent variable:	Default	ROA	Leverage	Default	ROA	Leverage
Granted	-0.007	0.005	0.002	-0.026	0.009	0.001
	(0.024)	(0.016)	(0.006)	(0.034)	(0.015)	(0.006)
Observations	75,801	75,801	75,801	61,520	61,520	61,520

	7	**8**	**9**
	Applicants with professional education		
Dependent variable:	Default	ROA	Leverage
Granted	-0.022	0.005	0.003
	(0.040)	(0.023)	(0.006)
Observations	14,556	14,556	14,556

Table A4. Credit decision, education, and income and wealth: Lagged outcomes

The table reports coefficients and standard errors clustered by individual (in parentheses). The dependent variable is on top of each regression and all of them enter lagged one year. 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 1% and 5% levels. The table includes all the control variables in Table 4.

	1	2	3	4	5	6
	Applicants with higher education			Applicants without higher education		
Dependent variable:	Income	Wealth	Pay inequality	Income	Wealth	Pay inequality
Granted	0.001	-0.000	0.000	0.001	0.002	-0.001
	(0.011)	(0.015)	(0.011)	(0.008)	(0.008)	(0.016)
Observations	75,801	75,801	75,801	61,520	61,520	61,520

	7	8	9
	Applicants with professional education		
Dependent variable:	Income	Wealth	Pay inequality
Granted	0.002	0.004	-0.003
	(0.011)	(0.019)	(0.015)
Observations	14,556	14,556	14,556

Figure A1. Leverage and ROA in North European small firms vs. our sample

The figure plots the annual mean of leverage and ROA of small and micro firms in Austria, Belgium, Denmark, France, Germany, and the Netherlands (solid lines) and the equivalent for firms in our sample (dashed lines).