Enero 2023

WP4-2023-002

N° de serie

DOCUMENTO DE TRABAJO

The Effects of Formal Financial Inclusion on Informal Credit Use: Evidence from Colombia.

Autores

Héctor Posada

Carlos Gaviria

David Londoño

COLOMBIA CIENTÍFICA
Conocimiento Global para el Desarrollo

ALIANZA EFI
Economía Formal e Inclusiva

The Effects of Formal Financial Inclusion on Informal Credit Use: Evidence from Colombia.

Héctor M. Posada *   Carlos Felipe Gaviria-Garcés †

David Londoño ‡

Abstract

It is well known that formal credit can improve household's wellbeing in developing countries, however, its relationship to informal credit as a substitute or complement is uncertain. This study aims to estimate the causal effect of formal credit on the use of informal credit sources using a Regression Discontinuity Designs which exploits credit allocation rules of formal institutions. We use financial and socioeconomic information directly provided by households trough the Encuesta Longitudinal de Hogares Colombianos (ELCA) survey. With this information we obtain a credit score and estimate a decision threshold, which together replicates the credit allocation rules of formal institutions. Our results show that formal credit reduces informal credit use. Our results also show that formal credit increases the probability of living in a formal housing and owing a motorcycle.

JEL Classification: G51, O17

Keywords: Financial inclusion, Credit access, Informal credit, RDD

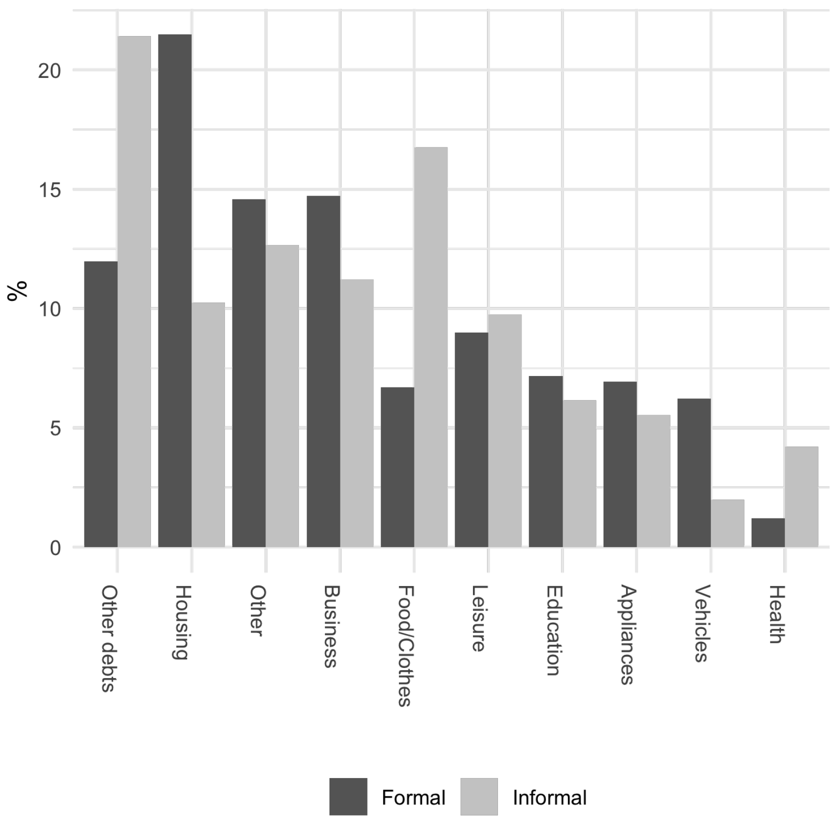
*Professor at Universidad de Antioquia. Email: hector.posada@udea.edu.co
†Professor at Universidad de Antioquia. Email: felipe.gaviria@udea.edu.co
‡Universidad de Antioquia. Email: david.londono@udea.edu.co
This paper is funded by the Colombia Científica-Alianza EFI Research Program, with code 60185 and contract number FP44842-220-2018, funded by the World Bank through the call Scientific Ecosystems, managed by the Colombian Ministry of Science, Technology, and Innovation.

1 Introduction

Several studies have proven how credit access can help disadvantaged households improve their living conditions in developing countries. Evidence has shown greater credit access leads to higher expenditure on education (Amendola et al., 2016; Doan et al., 2014), durable goods (Banerjee et al., 2015; Ruiz, 2013) and health (Doan et al., 2014). It has also been confirmed that greater credit access leads to higher savings (Aktaruzzaman and Farooq, 2017) labour supply (Niño, 2013) and children's education (Aktaruzzaman and Farooq, 2016). However, these positive effects can be limited if households access credit through informal sources. According to data from the Colombian Longitudinal Survey (Encuesta Longitudinal Colombiana) of the University of the Andes (ELCA) (Universidad de los Andes, 2016), informal credits have significantly higher interest rates and shorter payments term than formal credits. Moreover, informal credits are tied to illegal groups' activities and violent collection mechanisms under certain modalities. At the aggregate level, gains obtained by greater credit access may also be curtailed by the presence of an informal credit sector (Burgess and Pande, 2005a; Ayyagari et al., 2008; de Mel et al., 2012). Agents who grant informal credits create less added value than their formal counterparts by operating at smaller scales with unsophisticated information systems using unskilled labor, and, by definition, operating outside regulations without paying taxes. This leads to poor risk and liquidity management and low operational efficiency.

Considering the benefits of formal credit and the high costs associated with informal credit, as well as its widespread use in developing countries, national authorities have designed ambitious policies to achieve higher levels of inclusion in the formal financial system (see, for example, the proposals of national inclusion policies in the case of Colombia and Mexico: Departamento Nacional de Planeación, 2020; Gobierno de México, 2020). Nevertheless, the substitution of informal credits with formal ones –as one of the expected responses to these public policy agendas– is not necessarily guaranteed. If formal and informal credits are complementary, a greater formal credit dynamic could incite a greater demand for informally sourced credits. One of the main arguments in favor of formal and informal credit being complementary comes from the apparent differences in their uses. In Colombia, formal credits are destined to cover big purchases like housing and vehicles, while informal credits are primarily used to pay for food, clothing and other debts(see Figure 1). This could suggest both sources of financing could be considered complementary to a certain degree. Similarly, previous descriptions of the Mexican case shows households use informal products to cover credit needs that formal products do not cover (ver Vázquez, 2015). Nonetheless, 1 also show similar fraction of the credits from both sectors is intended for education, leisure, electrical appliances and promoting new ventures, suggesting a certain level of substitution between both sources of financing.

Figure 1: Credit use



This article studies the causal effect of formal credit access on the use of informal credit in Colombia. To estimate the causal effect, we propose a Regression Discontinuity Design (RDD) that exploits the decision rules financial institutions use for credit approval. To implement our identification strategy, information on households' credit scores and the threshold where they are considered creditworthy is needed. However, the scores used by financial institutions are not readily accessible, and even less so in conjunction with other socioeconomic data. To overcome these difficulties, we use the socioeconomic and financial information provided by ELCA to applydimensionality reduction methods, such as Principal Component Analysis (PCA) and Factor Analysis (FA), in order to recreate a credit score similar in nature to those employed by formal financial institutions. Additionally, since the decision threshold is unknown in our case, we will follow the threshold identification procedure proposed by Khan (2020) and inspired on Card et al. (2008). This procedure combines a structural change test with a search for outliers through a statistical learning method. On top of informal credit, we review evidence about the effects of financial inclusion on other financial behaviours, housing conditions and ownership of durable goods, which can significantly impact quality of life. Our findings reveal that approval of credit in the formal sector leads to a statistically significant reduction in the use of informal credits. We also demonstrate households included in the formal financial sector are less likely to live in precarious housing and more likely to own a motorcycle.This article is divided as follows. Section 2 is the literature review. Section 3 presents the context, followed by Section 4 describing the empirical strategy and data sources, and Section 5 with the results. The main conclusions are discussed in the last Section.

2 Literature Review

Our study contributes to the literature examining the effects of financial inclusion on households in developing countries. Regarding access to savings products, evidence from different studies using randomised experiments or quasi-experiments show that providing household members with savings accounts or simple electronic systems to deposit money leads to higher levels of savings (Aportela, 1999; Ashraf et al., 2006; Masino and Niño-Zarazúa, 2020; Lyons et al., 2020), female empowerment and consumption (Ashraf et al., 2010), expenditure on preventive health, productive investment, productivity and income (Dupas and Robinson, 2013b,a; Ashraf et al., 2010) and to reductions in poverty (Lyons et al., 2020). As for credit products, evidence shows increased access results in lower poverty levels (Burgess et al., 2005), greater income and a lower probability of being unemployed (Bruhn and Love, 2014). Greater access to microcredit has been proven to lead to investment on education (Amendola et al., 2016; Doan et al., 2014), durable goods and business creation (Banerjee et al., 2015); health (Doan et al., 2014), and borrowers' children's education (Aktaruzzaman and Farooq, 2016). It has also been shown to favourably impact savings levels (Aktaruzzaman and Farooq, 2017) and the likelihood of owning electronic devices, large appliances and furniture (Ruiz, 2013). Access to microcredits can also stimulate labour supply (Niño, 2013) and generate positive effects on consumption, economic independence, mental health and well-being (Kaboski and Townsend, 2011, 2012; Karlan and Zinman, 2010; Khandker, 2005; Pitt and Khandker, 1998). At the aggregate level, evidence indicates that credit restrictions are associated with slower growth (Ayyagari et al., 2008) and that, conversely, greater access would increase returns on the capital of micro-enterprises (de Mel et al., 2012) and reduce poverty levels (Burgess and Pande, 2005b). Additionally, there is evidence that limiting credit access restricts entrepreneurship and business growth (Banerjee and Duflo, 2007; Beck et al., 2005), which would generate significant losses in social welfare.

Meanwhile, the effects of greater formal credit access on the use of informal credit products have been less studied. Descriptive evidence reveals Mexican households use informal products to cover credit needs that are not adequately met by formal credits (ver Vázquez, 2015), suggesting households don't replace informal credit with formal given the chance. In contradiction, the work of Iregui et al. (2018) argues that when a household's income and education are increased, the probability of having formal credit rises while the chance of having an informal credit decreases. In aggregate form, the World Bank (2014) and World Bank (2022) reports reveal the use of informal sources of credit decreases as countries move towards a higher income level. To the best of our knowledge, our work is the first to present causal evidence of the effects of credit approval in the formal sector on the use of informal credit.

Our work also contributes to the growing empirical literature identifying a causal effect through RDD with an unknown cut-off threshold. One of the first antecedents can be found in Card et al. (2008), which investigate the permanence of white people in a neighbourhood as a function of the fraction of other ethnic populations. The study uses a structural change search technique based on Hansen (2000) to identify tipping points triggering mass migration. The first methodological proposal designed specifically for RDD is that of Porter and Yu (2015). The authors proved cut-off point estimation does not affect the efficiency of the treatment effect estimator and, therefore, RDD can be carried out as if the cut-off point were known from the outset. Unlike the structural change approach, Herlands et al. (2018) propose a method from statistical learning to detect discontinuities. This article uses the method proposed by Khan (2020), which makes use of both approaches –structural change and statistical learning –to detect discontinuities. In Khan (2020), Andrews (1993) test is initially applied to detect structural change and then bands (or bandwidth) defining a neighbourhood for the assignment variable around the detected cut-off point are calculated. Finally, with only the sample within the neighbourhood, a search for discontinuity is conducted again using the indicator saturation method of Pretis et al. (2018). Other proposals for detecting cut-off points include Hansen (2017) for when discontinuity is on the slope of the treatment function (the Regression kink case) and Yang (2019) for cut-off points that vary depending on another variable. Empirical studies have used these methodologies to detect cut-off points and estimate causal effects with RDD. Some examples of these applications include studies on the effect of mergers on university efficiency (Agasisti et al., 2021); the impact of secondary education on human capital (Ozier, 2018); the assessment of service provision programs on the employment and housing of households in extreme poverty (Carneiro et al., 2019); and the effect of carbon taxes on emissions (Pretis, 2022).

3 Context

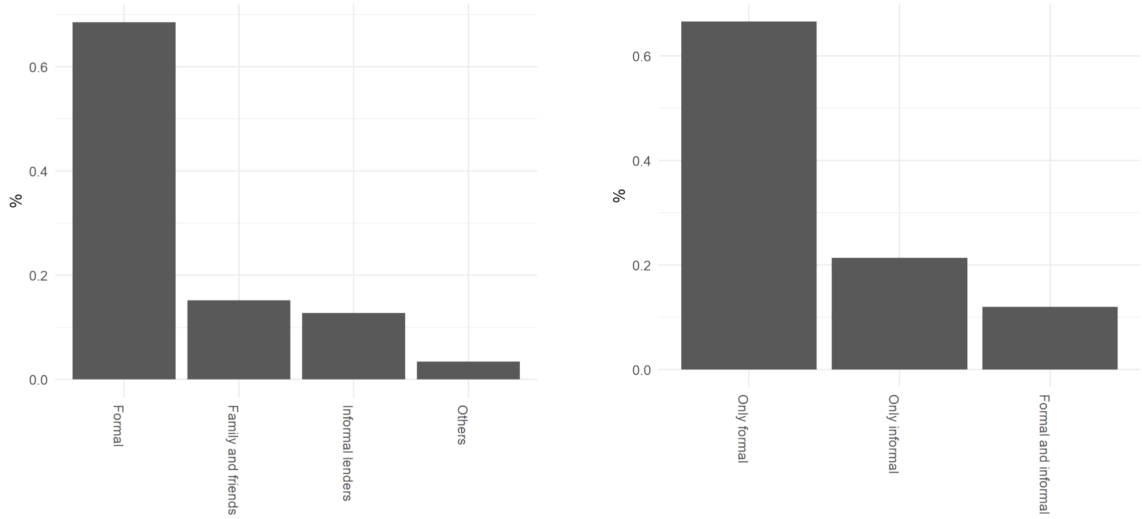
Informal credit is closely related to development levels. Table 1, based on Global Findex 2021 data (World Bank, 2022), shows households in high-income countries typically acquire credit through formal financial institutions, whereas a smaller proportion of loans are obtained from friends and family. In contrast, in developing areas such as Latin America, sub-Saharan Africa and southeast Asia, access to credits from formal institutions is limited, while credit from family and friends prevails. In all regions the use of savings clubs to obtain credit is found to be low, the only exception being sub-Saharan Africa where 11% of credit is obtained from such clubs.

Table 1: Credit sources

Region	Formal institution	Family and friends	Savings club
East Asia and Pacific	0.23	0.29	0.02
Europe and Central Asia	0.22	0.28	0.01
High income	0.47	0.14	0.01
Latin America and Caribbean	0.19	0.25	0.02
Middle East and North Africa	0.08	0.38	0.02
South Asia	0.11	0.35	0.03
Sub-Saharan Africa	0.1	0.38	0.11

According to the Global Findex 2017 report of the World Bank(Demirgüç-Kunt Asli and Hess, 2018), Colombia made important advances in the banking of its population between 2011 and 2017. In 2011, 30.4% of the population over 15 years of age had an account

Figure 2: Credits



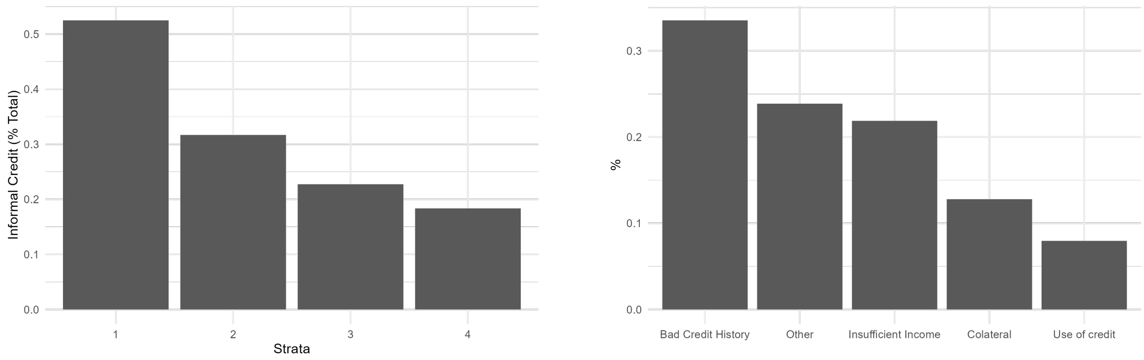
(a) % Credits by origin (b) % Households by credit source

Source: Author's calculations. ELCA.

in the financial sector, and this percentage increased to 44.9% in 2017. Nevertheless, the country significantly lags behind the rest of Latin America and other medium-high income countries. By 2017, the 'bankarisation' rate was 53.5 % in the rest of Latin America and 72.8% in middle-income countries. This lag is reflected in the lack of access to other financial products, crucially credit access. According to ELCA data, only 54% of Colombian households held at least one loan in 2016. Colombians access credit from a variety of sources. An important portion of the credits is obtained through family and friends (around 15%) and lenders (around 15%) who either operate informally (such as pawnbrokers and catalogue sellers) or illegally (such as loan sharks) (see Figure 2a). In turn, 21% of the households with some type of credit only had informal credits and another 12% had both formal and informal credits (see 2b). The high proportion of households holding informal loans exposes a significant level of financial exclusion, particularly in the case of poor households. Figure 3a shows informal credits account for more than 50% of the loans obtained by households with the worst socioeconomic conditions (that is, classified within the Stratum 1 band according to the socioeconomic stratification system used in Colombia), a proportion that drops to less than 20% for the households in the sample with the best socioeconomic conditions.

Compared to formal credit, informal credit is a disadvantageous option. First, informal lenders charge significantly higher interest rates than those from the formal sector. A study by La República (Lopez, 2015) found the average interest rate for consumer loans in the formal sector is 19.21% effective per year, while this same interest rate in the informal market would be close to an effective annual rate of 81%. Data from ELCA reports a more extreme situation: the median of the monthly interest rate in the formal sector is 2.4% compared to 16% in the informal sector. In addition, a sizeable portion of these credits has daily, weekly or fortnightly payment periods which, compared to monthly instalments, can make payment difficult given the irregular income of the poorest households. Other characteristics of informal credit include a higher proportion of microcredits and shorter

Figure 3: Credits characteristics



(a) Informal credit and Strata

(b) Credit Denial Reasons

Source: Author's calculations. ELCA.

payment terms. According to ELCA data, 69% of informal credits have initial values of less than COP $1,000,000 as compared to 8.2% of formal credits. Similarly, Asobancaria (2014) reveals formal credits provide larger amounts of money than informal loans and their payment terms are also longer and more convenient. In short, lack of access to formal credit forces a significant proportion of low-income households to use more expensive credits, with more frequent payments, smaller amounts and shorter terms, limiting their ability to make investments that improve their quality of life or sustain it in the face of unforeseen events.

In addition to lack of income, a number of other factors hamper households from obtaining formal credit. The following reasons for rejecting new credit applications, listed in order of importance, can be observed in Figure 3b: delay in payment of either financial or other types of obligations, insufficient income, lack of collateral and use of the credit. Overall, financial entities face asymmetric information issues in identifying whom to grant credit, the risk factor and a suitable interest rate to cover the risk. These information asymmetries partly arise because low-income households do not usually save through formal savings instruments and have not built a credit history enabling the bank to estimate their probability of default. Additionally, restrictions imposed by law on interest rates prevent banks from covering this greater risk with higher interest rates, limiting the access of risky households to credit.

4 Empirical strategy

4.1 Data and Variables

The present analysis uses the Longitudinal Survey of Colombian Households (*Encuesta Longitudinal de Hogares Colombianos*) of the University of the Andes (ELCA). In its urban chapter, the survey collects information on households and their members, and is representative for socioeconomic strata 1 to 4[1] and for five geographic regions (the Atlantic, Pacific, Central and Eastern regions, as well as Bogotá). The survey collected information across four different periods: 2010, 2013, 2016 and 2019 (although the information for 2019 is not yet public) on a wide range of socioeconomic indicators relating to households and their members, including financial behaviours, housing conditions, durable goods, among others.

In Colombia, as in many other countries, financial information is confidential and may only be obtained directly from banks and other formal financial institutions. The ELCA overcomes these barriers to data availability by collecting financial information directly from households on areas such as savings and loans. In particular, ELCA contains information on credit sources, uses and terms, allowing us to identify if the household got a formal credit approved and it uses informal credit sources. This survey collected information from 5,446 urban households in 2010, from 4,681 households in 2013 and from 5,275 households in 2016, a low sample loss for longitudinal surveys. To construct the indicator for informal credit holding, household credits are classified according to the legal status of their lenders. Formal lenders include banks, chain stores, employee funds, among other organisations. In turn, informal lenders include private lenders, relatives, friends and pawnshops. Among the housing quality variables, housing informality is defined following ONU-Habitat (2003), where housing is considered informal if quality standards are not met in terms of durable material components, legal tenure, overcrowding, access to improved water sources and access to sanitation facilities. Likewise, the poverty variable is constructed through the multidimensional poverty index defined for Colombia by de Planeación (2014). Lastly, the durable goods variables are dummies that indicate the ownership or lack thereof of a particular good, regardless of the amount owned.

4.2 Identification Strategy

This study aims to estimate the causal effect of financial inclusion, i.e., the approval of a credit application in the formal sector, on the use of informal credit. To achieve this, the difference between the average outcome for households included in the formal financial sector and the average outcome that would have been observed for these same households had they not been included (the counterfactual outcome) must be calculated. Since the counterfactual outcome cannot be observed, the above comparison must be made with a

[1]Stratification is a household socioeconomic classification system used to differentiate the rates applied to public services (electricity, water consumption and refuse collection) as well as the taxes imposed on land and housing property. A household can be assigned to a specific stratum band (1 to 6) depending on the physical conditions and the location of its place of residence. According to this classification, a higher socioeconomic stratum implies better socioeconomic conditions.

control group of households. A first approximation to identify the causal effect of financial inclusion consists of estimating the parameter β in the following equation:

$$y_i = \beta D_i + \delta^{'} \boldsymbol{X}_i + \varepsilon_i \tag{1}$$

where i is the index for households, y_i is the outcome variable of interest, D_i is a binary variable that takes the value of one when the household is included in the formal financial sector and zero when it is excluded; $\boldsymbol{X}_i$ is a vector of observable characteristics that explain the outcome of interest (including the intercept); and ε_i is the error term of the model. Estimating Equation 1 through ordinary least squares typically results in a biased and inconsistent estimator $\hat{\beta}$ since the treatment indicator D_i and the error term ε_i are correlated. This occurs because unobservable variables that explain household participation in the formal financial sector and the outcome variable are irremediably omitted from Equation 1. For instance, increased household head confidence in future events may lead to more active credit management in both the formal and informal sectors. The correlation between the treatment variable and the error term may also result from a reverse causality situation, for example, when the status of inclusion in the formal financial sector is affected by the informal sector's credit management.

To overcome the challenges associated with identifying the causal effect of financial inclusion, this research will exploit the decision rules used by formal financial institutions to approve household applications for credit products. Financial institutions use credit scores to determine if a person or household is likely to default on any bank credit product (such as a credit card). Based on these scores, financial institutions define a cut-off threshold c (cut-off) to determine which credit applications will be approved, which creates a discontinuity in access to formal credit products. In this study we exploit this discontinuity to estimate the causal effects of formal financial inclusion on different household socio-economic outcomes using a Regression Discontinuity Design (RDD). Because for some households with a good credit score ($s > c$) credit applications are not approved while for some households with a bad credit score ($s < c$) credit applications are approved (owing to banks considering multiple criteria to credit approval), the treatment probability jumps by a value of less than one at the cut-off score c. This type of design is known as a fuzzy RDD and, in this case, the causal effect is given by (see Hahn et al. (2001)):

$$\tau_f = \frac{\lim_{\epsilon \downarrow 0} E[Y_i | s = c + \epsilon] - \lim_{\epsilon \uparrow 0} E[Y_i | s = c + \epsilon]}{\lim_{\epsilon \downarrow 0} E[D_i | s = c + \epsilon] - \lim_{\epsilon \uparrow 0} E[D_i | s = c + \epsilon]}. \tag{2}$$

This is the Wald estimator with a binary instrument (the instrument T_i is defined by the cut-off score $s = c$ such that $T_i = 1$ if $s > c$and $T_i = 0$ if $s < c$) and a binary treatment variable D_i. For the above ratio to be interpreted as causal, monotonicity and exclusion assumptions must be met, and the functions $E[Y_i(1)|s]$ and $E[Y_i(0)|s]$ must be continuous in s (see Hahn et al. (2001)). In this case, τ_f is interpreted as a local treatment effect for the sub-population of households that receive treatment when they are above the cut-off score and which would otherwise not have received it. Due to its local nature, the causal effect is estimated using a local regression method that, in their simplest version, estimate the function of interest with a linear regression on both sides of the cut-off score c. In our study we use the procedure proposed by Imbens and Kalyanaraman (2012) y Calonico et al. (2014) to define an optimal neighbourhood around the cut-off score c, which minimises the mean squared error (MSE) of the estimator. This neighbourhood enables obtaining an estimator with desirable properties, such as consistency and a fast decrease in the MSE. Its implicit specification bias, however, renders traditional inference unsuitable. Therefore, we adopt the proposal of Calonico et al. (2022) to perform inference with optimal confidence intervals that minimise the coverage error rate (CER).

4.3 Building a credit score

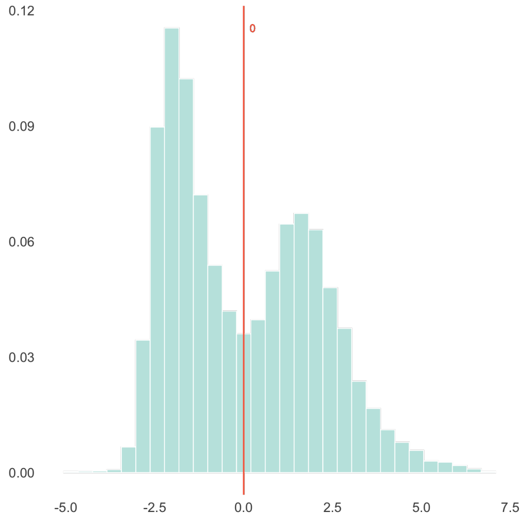
Credit scores used by banks are built by private entities from household data across multiple dimensions such as income, expenses, quality of work, credit history and asset holdings, among others. A widely used credit score is the FICO (an acronym for Fair, Isaac and Company, the names of its creators), a 3-digit indicator that predicts default risk based on information from loan and debt records, including payment history, debt size, diversification of banking products, loan period, loan history and new loan applications. In Colombia, banks and other commercial and financial institutions also use scores that measure the risk of households defaulting on payment. These scores are built by Datacrédito from information collected on income, type of employment, expenses, assets, education, credit history, etc.

The credit scores s of each household in the sample are needed to perform this estimation adequately. However, accessing this information poses important practical and methodological problems. As previously mentioned, credit scores are confidential information and access to them is therefore cumbersome, even more so in conditions that would allow households to be identified –a necessary condition for matching them with other types of socioeconomic information successfully. More importantly, in cases where credit scores are supplied to the researcher with identifiers, it is unlikely other researchers will be able to reproduce the results in their entirety due to confidentiality agreements. To overcome these drawbacks, this study exploits the socioeconomic and financial information provided by ELCA in its three rounds (2010, 2013, and 2016) to build credit scores similar to those used by financial institutions. Credit scores are obtained from implementing widely used dimensionality reduction methods such as Principal Component Analysis (PCA) and Factor Analysis (FA). The methods used in this study to build credit scores are summarised in Section A of the Appendix and the full list of variables employed is provided in Section B of the Appendix.

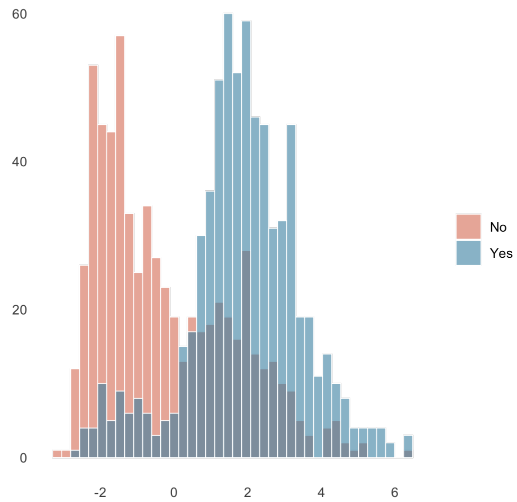
The credit score obtained from the PCA methodology has a bimodal distribution for the set of all households (see Figure 4a), suggesting the presence of two clearly differentiated groups. In fact, when grouped histograms are plotted from the credit acceptance categories (see Figure 4b), it is observed that the distribution of the score for households whose credit application is accepted is to the right of the distribution for households whose credit application is not accepted[2]. The same pattern is observed with the score obtained from the factor analysis method (see 4c and 4d). Finally, the score distributions are compared considering also the categories of all households and households not applying for credit, again observing a higher average score and a higher median for households whose credit applications were approved (see Figure 5).

[2]It must be noted that credit approval status is not used in the score-building process

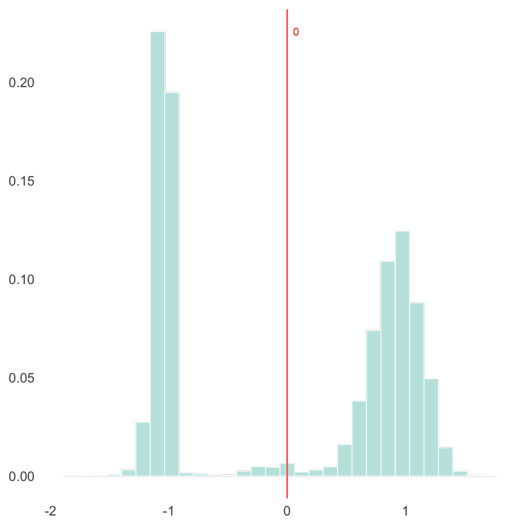
Figure 4: **Score distribution among methods applied to reduce dimensionality**



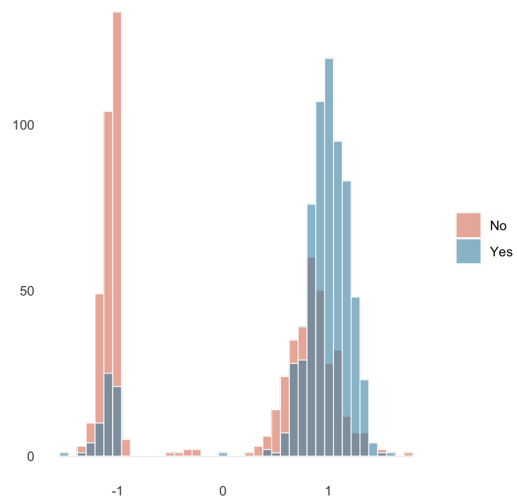
(a) Pca



(b) Comparative pca



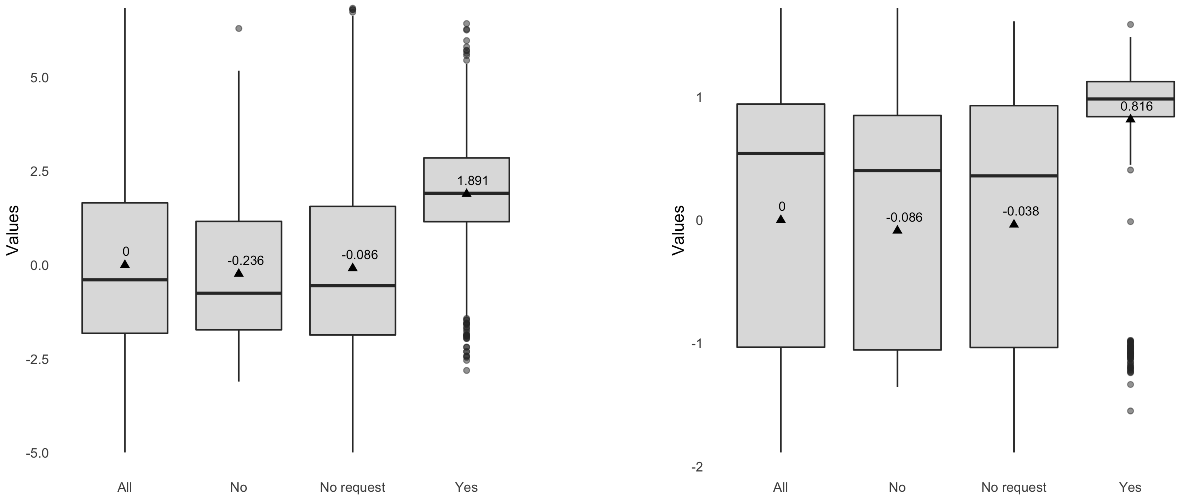
(c) Factor



(d) Comparative factor

Source: Author's calculations. ELCA.

Figure 5: **Comparison of scores result of methods to reduce dimensionality**



(a) Pca (b) Factor

Source: Author's calculations. ELCA.

4.4 Estimation of classification threshold

In a traditional regression discontinuity design, the cut-off point c is known to the researcher. In our study the cut-off point is unknown as the assignment variable is an independently constructed credit score. To identify the cut-off point, we employ the three-step procedure proposed by Khan (2020), combining a structural change test with a search for outliers through a statistical learning method (a review of this recent literature is presented in section C of the appendix). In the first step of the procedure, the Andrews (1993) test is implemented to detect unknown breaks in the probability of financial inclusion. Bands are then estimated which define a neighbourhood of the assignment variable around the detected cut-off point. Finally, using only the sample within the neighbourhood, discontinuity is again searched for through the indicator saturation method of Pretis et al. (2018). It should be clarified that both the Andrews test and the Pretis procedure are methods originally developed for time series. To use them, a pseudo-index is created by sorting the assignment variable in ascending order.

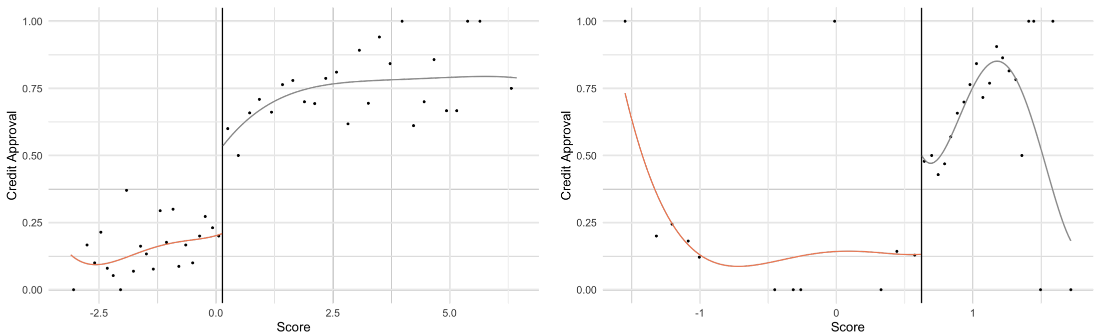
This procedure is advantageous in that it does not require testing for the presence of a selection effect or treatment before estimating the cut-off point since one or both can be tested simultaneously (Khan, 2020). Other advantages of using this method include the ability to differentiate between outliers and discontinuities, to differentiate whether these discontinuities are found in the function or in its derivative, and the ready availability of statistical software to implement the three steps. For these reasons, this methodology will be used to identify discontinuity in the treatment probability in our exercise. The technical details of the methodology are presented in section D of the Appendix.

5 Results

5.1 Evaluating identification assumptions

Our estimation of the causal effect through an RDD is possible as long as a discontinuity (jump) in the probability of credit application approval in the formal sector is observed at the classification threshold. To verify the presence of such a discontinuity, we use RDD graphs for the probability of treatment following the procedure proposed by Calonico et al. (2014). Figure 6 confirms that there is indeed a jump in the probability of treatment around the cut-off c for both the scores constructed from PCA and FA. That is, evidence is shown in favour of $\lim_{\epsilon \downarrow 0} Pr[D = 1|s = c + \epsilon] < \lim_{\epsilon \uparrow 0} Pr[D = 1|s = c + \epsilon]$. P-values associated with the null hypothesis that there are no differences in the likelihood of treatment on either side of c are 0.005 and 0.03 for PCA and FA.

Figure 6: Treatment discontinuity



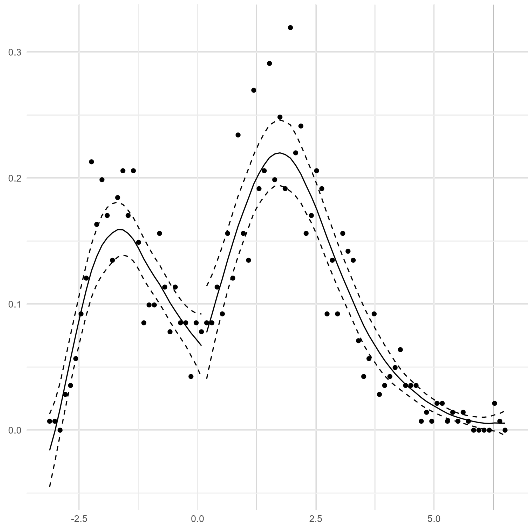
(a) Pca (b) Factor

Source: Author's calculations. ELCA.

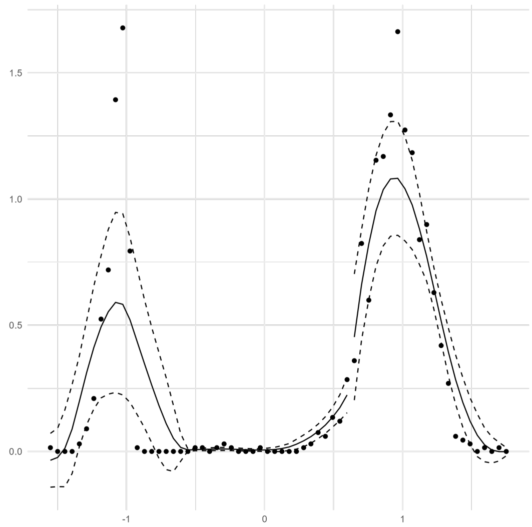
If households self-select around the cut-off score, then treated and untreated households will be systematically different. In this case the RDD strategy would not allow us to obtain an unbiased and consistent estimator of the causal effect of interest. Next, in order to confirm whether households self-select, the MacCrary test (see McCrary (2008)) is used to verify if there is any discontinuity in the distribution of the assignment variable around the cut-off point. Figure 7 shows that for both PCA and FA the null hypothesis that there is no discontinuity in the estimated density for the assignment variable around the cut-off point is not rejected (p-values associated with the null hypothesis are 0.55 and 0.10 for PCA and FA respectively).

To support the evidence that there are no systematic differences between treated and untreated households, balance of means testing is also implemented around the cut-off score for a set of predetermined variables. Table 2 reveals that the null hypothesis of equality of means is not rejected for most of the covariates used in the exercise. The above evidence suggest the fulfillment of the assumption of continuity and local randomization, which is needed for correctly identifying the causal effect of financial inclusion on the outcome variables.

Figure 7: MacCrary test



(a) Pca



(b) Factor

Source: Author's calculations. ELCA.

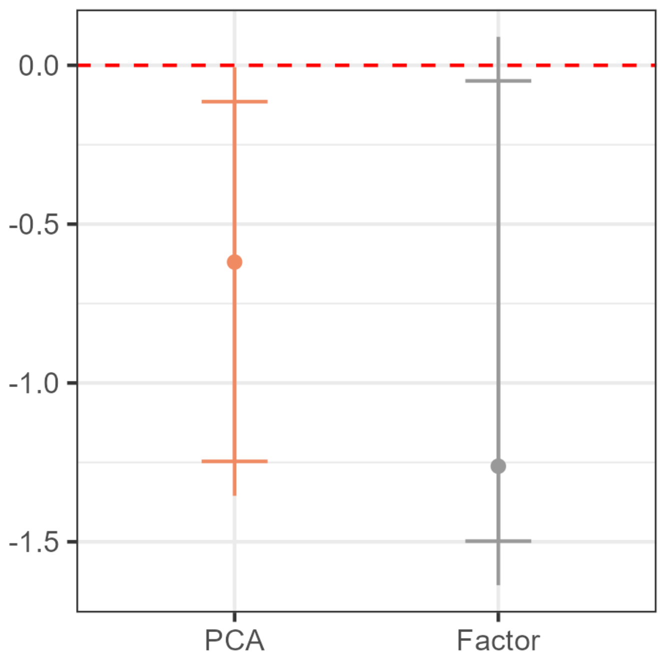
Table 2: Covariates Balance

	PCA		Factor	
Variable	Coef	P-value	Coef	P-value
Assets	0.05	0.71	0.17	0.15
Adults	-0.41	0.16	-0.02	0.50
Age	-0.32	0.30	-0.34	0.12
Education	-0.28	0.47	0.65	0.01
Unemployed	0.02	0.49	-0.04	0.74
Formal worker	0.02	0.91	0.16	0.10
Inactive	-0.01	0.97	0.08	0.16
Informal worker	-0.06	0.68	-0.17	0.16
Strata 1	0.24	0.15	-0.23	0.16
Strata 2	-0.12	0.41	0.09	0.83
Strata 3	-0.19	0.14	0.10	0.30
Strata 4	0.05	0.47	0.04	0.16
Income Percap	-0.08	0.78	0.65	0.02
Income	-0.36	0.16	0.32	0.12
Credits	0.29	0.05	0.20	0.11
Default (yes)	0.05	0.75	-0.29	0.07
Balance Owed	0.02	0.91	-0.01	0.87
Gender (Female)	0.18	0.45	-0.11	0.39
Owner	-0.26	0.14	0.14	0.19

5.2 Effects of credit access on households' financial behaviour

Our results reveal the approval of a credit application in the formal sector reduces the probability of use by households of informal credits (see Figure 8). This evidence argues against what is described in Vázquez (2015), which suggests that informal credit is primarily a complement to formal credit. As previously discussed, informal credit is significantly more expensive than formal credit and may be provided by individuals and organisations that are closely linked to illegal armed groups, meaning collection mechanisms for late payments may be violent. Consequently, once households have the possibility of accessing credit in the formal sector, they cease to request and make use of informal credits. This reinforces the idea informal credit constitutes a lesser (inferior) financing option for households, a choice made primarily by persons without access to formal sources. These findings support the descriptive evidence presented by Iregui et al. (2018), Demirgüç-Kunt Asli and Hess (2018) and the World Bank (2022). Considering that most households do not have multiple credits but a single one, and that our outcome variable also indicates whether or not households have at least one informal credit, the above finding reveals informal credits are strongly excluded or substituted by formal credits, making this substitution between credit types more plausible than a mixture of both.

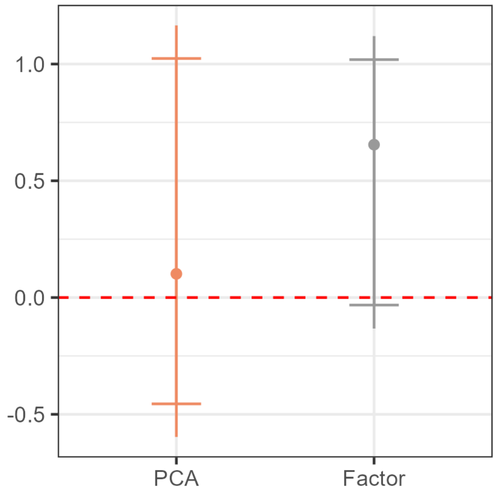
Figure 8: Treatment Effects on Informal Credit



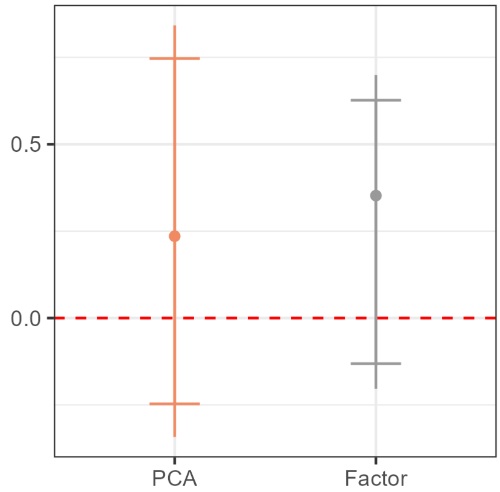
Source: Author's calculations. ELCA.

The effects of greater credit access on savings are also identified in this analysis. Even though debt service may represent a significant burden for low-income households, access to credit may discourage savings by enabling, unlike the latter, prompt financing of investment projects, consumption, or expenses resulting from emergencies. On the other hand, gaining access to the financial system through credit may encourage households to

Figure 9: Treatment Effects on Savings



(a) Savings



(b) Savings Accounts

Source: Author's calculations. ELCA.

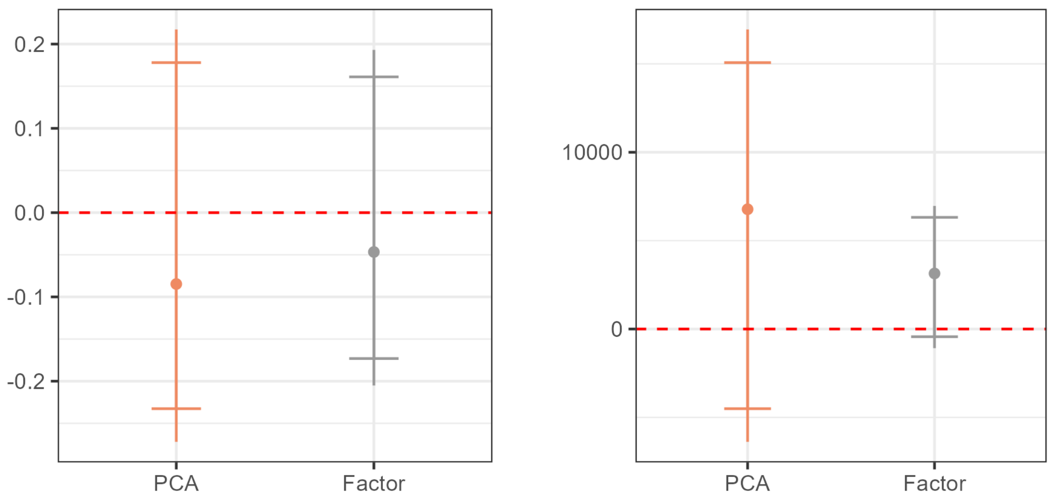
save by educating them about the various savings alternatives and tools that generate long-term profitability. Credit access could also indirectly increase savings levels if credits are used to financing investment projects that will generate returns in the future. Our results show receiving a formal loan does not affect families' savings decisions (see **??**), which indicates that the opposite effects described previously may be counteracting each other. Similarly, our results reveal obtaining formal credit has little impact on the choice to save money in the financial system, which is closely related to the previous finding, given that if no overall changes are observed in savings decisions, it is unlikely any significant changes will occur in those savings moving to bank accounts.

5.3 Effects of credit access on households' socioeconomic outcomes

This section will assess the impact of formal credit access on various households' socioeconomic outcomes. Figures 10 and 11 show the results of the estimation on the multidimensional poverty index, education expenditure and housing informality. These results show formal credit access does not produce statistically significant changes on either the multidimensional poverty index or education expenditure, but it does create a statistically significant reduction (at a significance level of 10% for the score built from factor analysis) in the housing informality (or precariousness) indicator. As observed above in Section 3, unlike informal credits, formal credits offer longer payment terms, more regular payment schedules and higher financed amounts, which makes them particularly suitable for financing purchases or real estate improvements. In fact, as shown in Figure 1, formal credits are mainly destined for housing expenses. It is therefore not surprising greater formal credit access causally increases the probability of owning a home in better physical conditions.

Now, given that our causal estimators are local and defined for households marginally included in the formal credit market, access to formal credit providing larger amounts of money (and greater impact) may continue to be restricted for these households. Consequently, the estimated effect of credit on housing precariousness may not have the scope previously mentioned. Even so, low-income households' housing allows for a progressive adaptation to the minimum standards of quality by means of using limited financial resources. Consider, for example, the case of a home improvement through changing the exterior wall material or connecting to a recently constructed water system. Improvement of housing conditions can occur even if the new access to formal credit is aimed at spending other than on such improvements. Consider, for example, the situation in which households have more money available to spend on home improvements as a result of lower borrowing costs due to changes in financing sources.

Figure 10: Treatment Effects on Poverty and Spending on Education

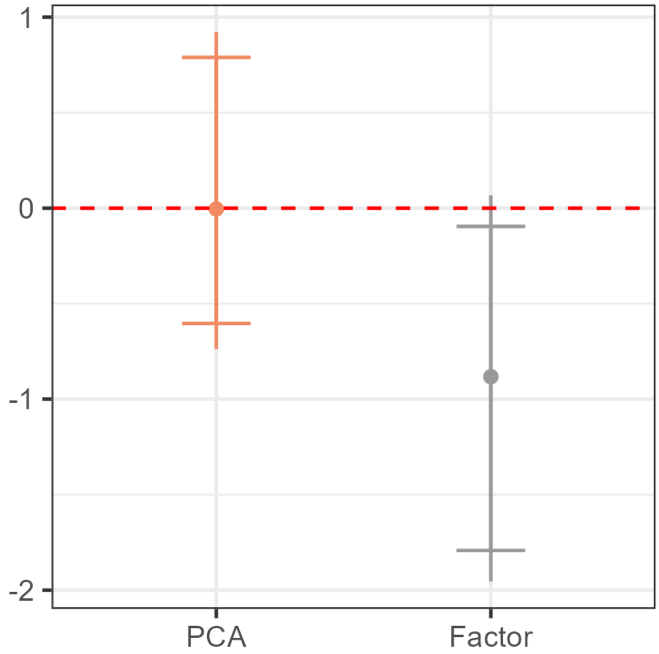


(a) Poverty

(b) Spending on Education

Source: Author's calculations. ELCA.

Figure 11: Treatment Effects on Informal Housing



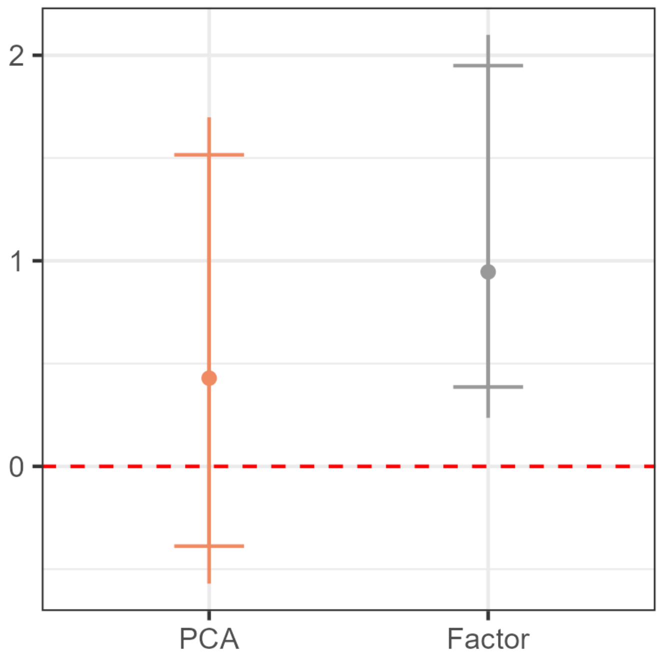
Source: Author's calculations. ELCA.

5.4 Effects of credit access on the ownership of durable goods

Finally, we calculated the effect of having access to a new credit on the ownership of a set of durable goods. Figures 12 and 13 show the effects of credit access on motorcycle, computer and TV ownership indicators. Overall, it is evinced that credit access does not generate statistically significant changes in the ownership of durable goods, with the notable exception of motorcycles. Specifically, credit access produces a statistically significant increase in the probability of owning a motorcycle (however, this is true only in the case of the score built from the factor analysis). These results suggest households normally decide to invest in a motorcycle when obtaining credit, which implies a greater investment as compared to common household appliances.

Motorcycles have great social and economic relevance in the Colombian and Latin American context. Hence, encouraging greater access to them (for example through financial inclusion policies) can create significant improvements in social welfare levels. In first place, motorcycles have several advantages over other alternative modes of transportation: compared to public transportation, they help avoid traffic congestion and thus reduce travel times; offer greater flexibility of use; and also benefit from lower acquisition, operation and maintenance costs in contrast to other private transport means, such as the automobile (Rodríguez et al., 2015). Second, owning a motorcycle increases the chances of finding a job by bringing households closer to more employment options. Finally, motorcycles are also commonly used as an asset for income generation. According to ANDI (2017) in 2016, 75.6% of motorcycles in Colombia were purchased only as a means of transportation while 22.4% were also bought as a means of increasing income. Motorcy-

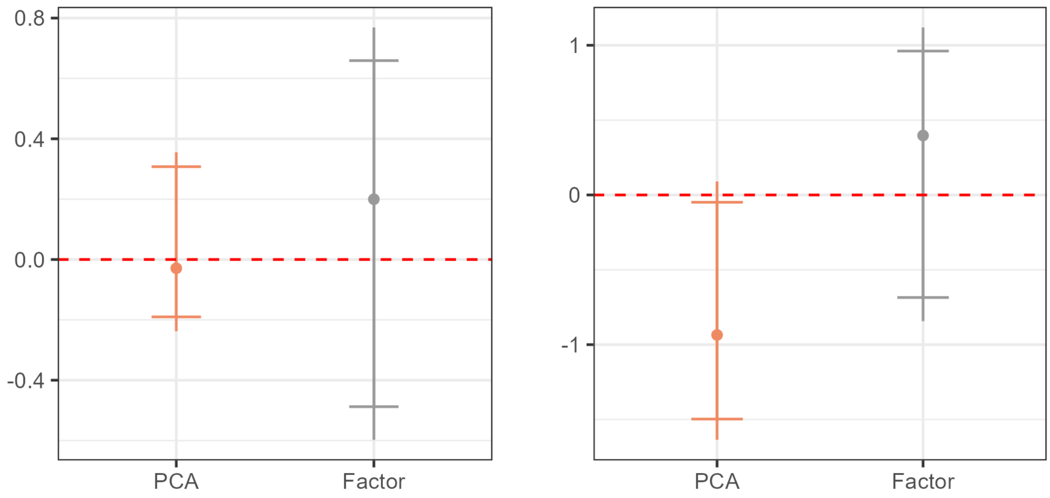
Figure 12: Treatment Effects on Motorcycles



Source: Author's calculations. ELCA.

cles are instrumental in offering courier, transport and delivery services. For example, in recent years, home deliveries have become an important source of income in response to the rapid adoption of mobile applications and the increased demand stemming from the coronavirus pandemic. Considering all this, and that 70% of motorcycle buyers' incomes in 2016 amounted to fewer than two minimum wages and 45.2% only had completed basic primary and secondary studies ANDI (2017), it can be concluded that investing in a motorcycle is a significant way to help many low-income families improve their quality of life and fight poverty.

Figure 13: Treatment Effects on durables

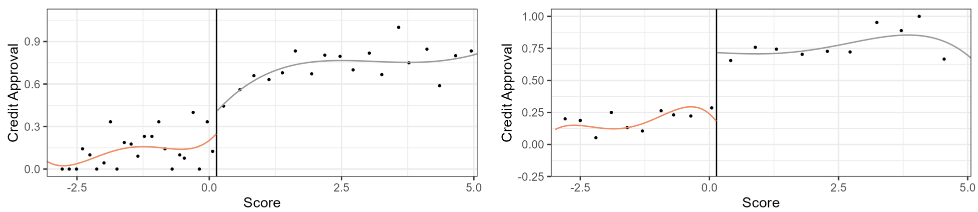


(a) TV (b) Computers

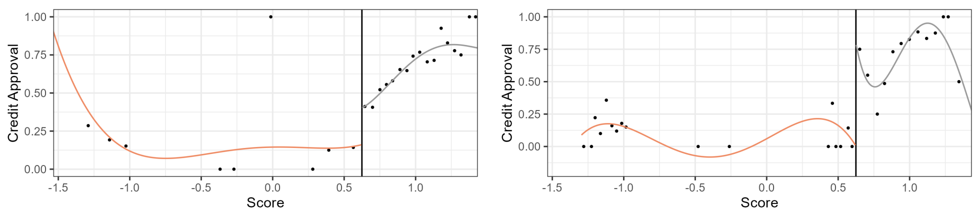
5.5 Heterogeneous effects by gender

To better understand how formal credit access affects informal credit usage, we examine how the estimated effect varies depending on the gender of the head of the household. The common procedure to obtain heterogeneous effects consists in estimating the causal effect using only the subsamples of households headed by a woman or a man, and then compare the estimators.

Figure 14: Discontinuity by gender



(a) PCA - Men (b) PCA - Women



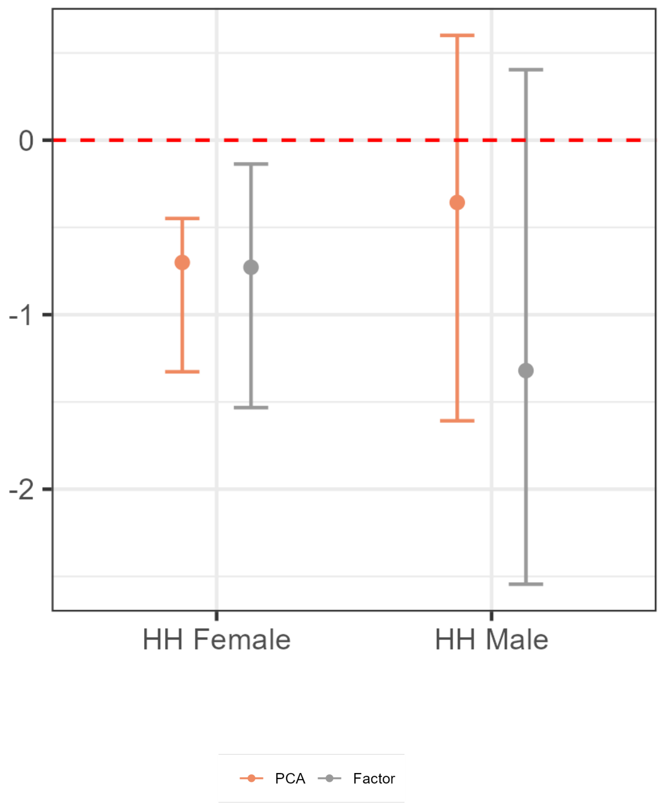
(c) Factor - Men (d) Factor - Women

Source: Author's calculations. ELCA.

Figure 15 indicates credit access generates a statistically significant decrease in informal credit use in women-headed households, while they do not generate statistically significant changes in households whose head is a man. This finding may have multiple explanations ranging from behavioural differences between genders to households' socioeconomic differences. For example, heterogeneity in the results may owe to the fact that men are more prone than women to take risks (Byrnes et al., 1999), which could result in higher levels of overall debt that the bank may not initially want to cover and therefore might not replace informal credit to the same amount. At the same time, households headed by women could face worse conditions than their counterparts, especially if the economic burden falls exclusively on them. This would indirectly imply credit access has a greater impact on the conditions of the poorest households. Indeed, it can be argued households with far fewer resources use credit more productively.

It is worth mentioning the literature has documented differences in the effect of microcredit access between genders. This heterogeneity can be explained by differences in male and female priorities and the effect of credit on women's empowerment in terms of household decision-making and economic freedom (Porter, 2016; Pitt et al., 2003; Al-shami et al., 2018). The above highlights the importance of designing financial inclusion policies in developing countries that take into account gender differences in relation to borrowing.

Figure 15: Effects by Gender

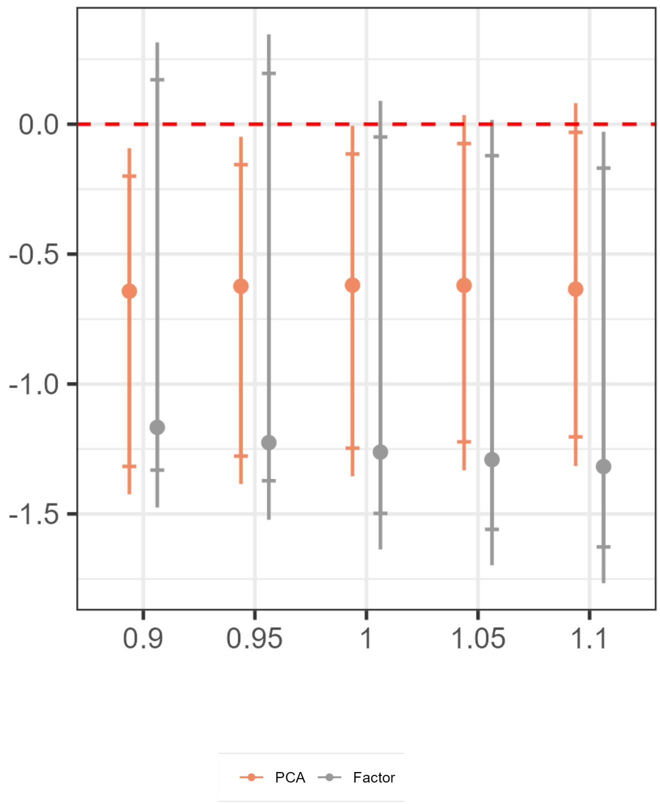


Source: Author's calculations. ELCA.

5.6 Robustness analysis

To test the results' sensitivity to the inclusion or exclusion of certain observations, two procedures were carried out as proposed by Cattaneo et al. (2019). The first one entails changing the bands where the local estimation is performed. This falsification method consists of performing the estimation with bands larger or smaller than the optimal ones. However, it only makes sense to explore robustness in bands close to the optimal ones given that widening the bands would increase the misspecification bias while narrowing them would increase the variance of the estimator. Consequently, the results are examined for bands 5% and 10% below and above the optimal ones. Figure 16 shows the results using informal credit holding as the outcome variable. Overall, coefficient estimation is robust to all the combinations examined, with minor changes in magnitude. In terms of inference, statistical significance holds in most cases.

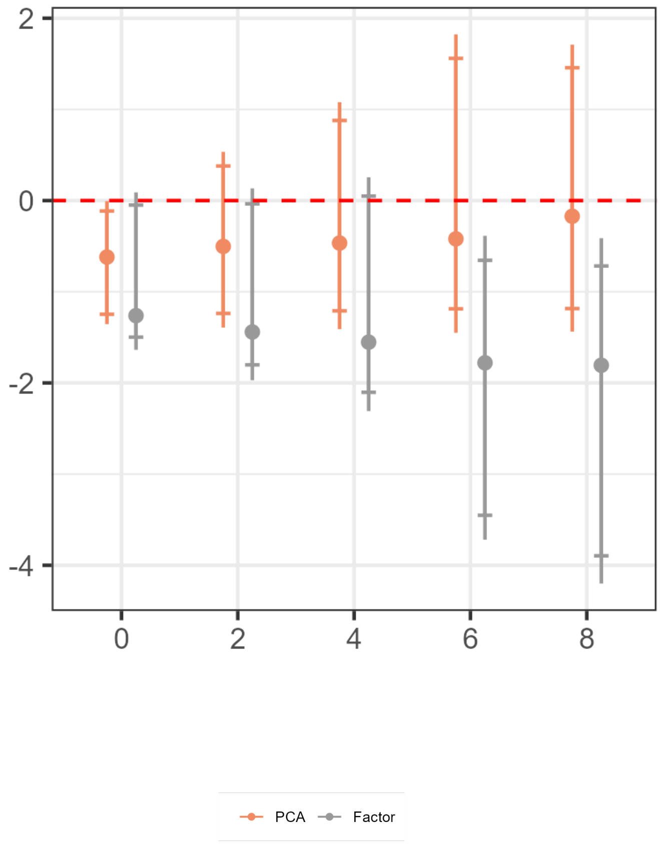
Figure 16: Informal Credit Robustness to Bandwidth Size



Source: Author's calculations. ELCA.

The second procedure consists of excluding a number of observations around the cut-off point and performing the estimation on doughnut-shaped bands. By using a triangular kernel, these observations are crucial for defining the bands and estimating the causal parameter. Omitting a certain number of observations enables checking the robustness of the results to the extrapolation of the local polynomial estimation. The results are examined in Figure 17 using informal credit holding as the outcome variable. Similar to the first procedure, the point estimates do not appear to be overly sensitive to these omitted observations, exhibiting only minor variations in size while maintaining the same sign. Inference in PCA does appear to be highly sensitive to the exclusion of observations near the cut-off point. The effect on informal credit becomes non-significant at this score in all cases examined whereas FA results remain robust.

Figure 17: Informal Credit Robustness to Omitted Observations



Source: Author's calculations. ELCA.

6 Conclusions

Empirical evidence has consistently shown formal credit has significant positive impacts on the quality of life of households in the developing world. However, in countries like Colombia, formal credit access is limited so informal financing options proliferate which generate significant costs. Faced with this situation, authorities have promoted policies facilitating access to formal finance sources in the hope they will in turn discourage using informal sources. This study aims to estimate the causal effect of greater access to formal credit on having informal credits. To this end, a regression discontinuity design examining the credit approval rules of formal financial entities is used. These rules are based on the building of scores that measure the likelihood of payment delinquencies and defining thresholds that separate well-rated households from poorly rated ones. Our findings suggest households cease to use financing from informal sources after being approved for formal credit. Consequently, the main conclusion of this paper is that informal credit may be regarded as an inferior substitute to formal credit, and imposes significant costs both on households and society at large. Furthermore, formal credit access generates a statistically significant rise in the probability of owning quality housing and motorcycles –goods that have a crucial impact on households' quality of life.

Our findings thus support the use of various policy options aiming to increase levels of inclusion in the formal financial system through greater credit access. Considering that bad credit history is the main reason for credit rejection, it is necessary to evaluate reporting policies in credit bureaux in order to regulate minor payment default reports, such as mobile phone bills, which may have disproportionate impacts on individuals' access to credit. In this sense, Law 2157 of 2021 in Colombia, also called 'Wiping the slate clean,' seeks to clear the credit history of individuals who pay all delinquent debts within a six-month period. Although this law provides an opportunity to improve bad credit histories, its application is only valid for one year. Hence, policies of this type must be made permanent to achieve real structural changes that expand formal credit access.

Other policies aimed at increasing access include financial education programs for the general population. One of the possible reasons for the high incidence of informal credit is that its disadvantages –such as high interest rates–are not apparent to a person with no basic financial knowledge. In effect, ELCA removed the question about interest rate after the first round in 2010 as the majority of responses were in the "Don't know / No answer" categories. To a large extent, the interest rate summarises the cost of credit and without this knowledge people will not be able to fully assess various financing options adequately.

Microcredit promotion is another widely used policy. As noted previously, small-scale loans continue to be concentrated under informal credit. Relaxing conditions such as requiring collateral and eliminating restrictions on credit destination are effective strategies to increase access to this type of formal credit and reduce the use of informal sources.

References

Agasisti, T., Egorov, A. and Maximova, M. (2021). Do merger policies increase universities' efficiency? evidence from a fuzzy regression discontinuity design. *Applied Economics* 53: 185–204, doi:10.1080/00036846.2020.1803488.

Aktaruzzaman, K. and Farooq, O. (2016). Impact of microcredit on borrowers' expenditures: a fuzzy regression discontinuity design approach. *Applied Economics* 48: 3685–3694, doi:10.1080/00036846.2016.1142662.

Aktaruzzaman, K. and Farooq, O. (2017). Does microcredit increase borrowers' savings? A fuzzy regression discontinuity design approach. *International Review of Applied Economics* 31: 495–507, doi:10.1080/02692171.2016.1263607.

Al-shami, S. S. A., Razali, R. M. and Rashid, N. (2018). The effect of microcredit on women empowerment in welfare and decisions making in malaysia. *Social Indicators Research* 137: 1073–1090, doi:10.1007/s11205-017-1632-2.

Amendola, A., Boccia, M., Mele, G. and Sensini, L. (2016). Financial Access and Household Welfare: Evidence from Mauritania. *Financial Access and Household Welfare: Evidence from Mauritania* doi:10.1596/1813-9450-7533.

ANDI (2017). Las motocicletas en colombia: aliadas del desarrollo del país.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* 61: 821, doi:10.2307/2951764.

Aportela, F. (1999). Effects of Financial Access on Savings by Low-Income People.

Ashraf, N., Karlan, D. and Yin, W. (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *The Quarterly Journal of Economics* 121: 635–672.

Ashraf, N., Karlan, D. and Yin, W. (2010). Female Empowerment: Further Evidence From a Commitment Savings Product in the Philippines. *World Development* 3: 333–344.

Asobancaria (2014). Características del acceso al crédito en las zonas urbanas de colombia.

Ayyagari, M., Demirguc-Kunt, A. and Maksimovic, V. (2008). How important are financing constraints? the role of finance in the business environment. *The World Bank Economic Review* 22: 483–516.

Banerjee, A., Duflo, E., Glennerster, R. and Kinnan, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics* 7: 22–53, doi:10.1257/app.20130533.

Banerjee, A. V. and Duflo, E. (2007). The economic lives of the poor. *Journal of Economic Perspectives* 21: 141–168, doi:10.1257/jep.21.1.141.

Beck, T., Demirguc-Kunt, A. and Maksimovic, V. (2005). Financial and legal constraints to growth: Does firm size matter? *The Journal of Finance* 60: 137–177, doi:https://doi.org/10.1111/j.1540-6261.2005.00727.x.

Bruhn, M. and Love, I. (2014). The real impact of improved access to finance: Evidence from mexico. *The Journal of Finance* 69: 1347–1376, doi:https://doi.org/10.1111/jofi.12091.

Burgess, R. and Pande, R. (2005a). Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review* 95: 780–795, doi:10.1257/0002828054201242.

Burgess, R. and Pande, R. (2005b). Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review* 95: 780–795, doi:10.1257/0002828054201242.

Burgess, R., Wong, G. and Pande, R. (2005). Banking for the poor: Evidence from india. *Journal of the European Economic Association* 3: 268–278.

Byrnes, J. P., Miller, D. C. and Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin* 125: 367–383, doi:10.1037/0033-2909.125.3.367.

Calonico, S., Cattaneo, M. D. and Farrell, M. H. (2022). Coverage error optimal confidence intervals for local polynomial regression. *Bernoulli* 28: 2998–3022, doi:10.3150/21-BEJ1445.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* 82: 2295–2326, doi:https://doi.org/10.3982/ECTA11757.

Card, D., Mas, A. and Rothstein, J. (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics* 123: 177–218, doi:10.1162/qjec.2008.123.1.177.

Carneiro, P., Galasso, E. and Ginja, R. (2019). Tackling social exclusion: Evidence from chile. *Economic Journal* 129: 172–208, doi:10.1111/ecoj.12594.

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs*. Cambridge University Press, doi:10.1017/9781108684606.

Demirgüç-Kunt Asli, D. S. S. A., Leora Klapper and Hess, J. (2018). *The Global Findex Database 2017: Measuring Financial Inclusion and the Fin-tech Revolution*. Washington, DC: World Bank, License: Creative Commons Attribution CC BY 3.0 IGO.

Departamento Nacional de Planeación (2020). *Política Nacional de Inclusión y Educación Económica y Financiera*. https://colaboracion.dnp.gov.co/CDT/Conpes/Económicos/4005.pdf [Accessed: 01.31.2022].

Doan, T., Gibson, J. and Holmes, M. (2014). Impact of Household Credit on Education and Healthcare Spending by the Poor in Peri-Urban Areas, Vietnam. *Southeast Asian Economies* 31: 87, doi:10.1355/ae31-1f.

Dupas, P. and Robinson, J. (2013a). Savings constraints and microenterprise development: Evidence from a field experiment in kenya. *American Economic Journal: Applied Economics* 5: 163–92, doi:10.1257/app.5.1.163.

Dupas, P. and Robinson, J. (2013b). Why don't the poor save more? evidence from health savings experiments. *American Economic Review* 103: 1138–71, doi:10.1257/aer.103.4.1138.

Fan, Y. and Li, Q. (1996). Consistent model specification tests: Omitted variables and semiparametric functional forms. *Econometrica* 64: 865, doi:10.2307/2171848.

Gobierno de México (2020). *Política Nacional de Inclusión Financiera*. https://www.gob.mx/cms/uploads/attachment/file/643214/PNIF__2020.pdf [Accessed: 01.31.2022].

Hahn, J., Todd, P. and Van Der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica* 35: 189–190.

Hansen, B. E. (2000). Testing for structural change in conditional models. *Journal of Econometrics* 97: 93–115, doi:10.1016/S0304-4076(99)00068-8.

Hansen, B. E. (2017). Regression kink with an unknown threshold. *Journal of Business and Economic Statistics* 35: 228–240, doi:10.1080/07350015.2015.1073595.

Herlands, W., Wilson, A. G., McFowland, E. and Neill, D. B. (2018). Automated local regression discontinuity design discovery. Association for Computing Machinery, 1512–1520, doi:10.1145/3219819.3219982.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies* 79: 933–959.

Iregui, A., Melo, L., Ramirez, M. and Tribin, A. (2018). Crédito formal e informal de los hogares en colombia. In Roa, M. and Mejía, D. (eds), *Decisiones financieras de los hogares e inclusión financiera: evidencia para América Latina y el Caribe*. CAF.

Kaboski, J. P. and Townsend, R. M. (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica* 79: 1357–1406, doi:https://doi.org/10.3982/ECTA7079.

Kaboski, J. P. and Townsend, R. M. (2012). The impact of credit on village economies. *American Economic Journal: Applied Economics* 4: 98–133.

Karlan, D. and Zinman, J. (2010). Expanding credit access: Using randomized supply decisions to estimate the impacts. *The Review of Financial Studies* 23: 433–464.

Khan, T. A. (2020). Regression Discontinuity Design with Unknown Cutoff: Cutoff Detection Effect Estimation. Master's thesis, University of Victoria.

Khandker, S. R. (2005). Microfinance and poverty: Evidence using panel data from Bangladesh. *World Bank Economic Review* 19: 263–286, doi:10.1093/wber/lhi008.

Lopez, C. (2015). El dilema de los créditos formales e informales en colombia.

Lyons, A., Kass-Hanna, J. and Greenlee, A. (2020). Impacts of financial and digital inclusion on poverty in south asia and sub-saharan africa. *Available at SSRN 3684265* .

Masino, S. and Niño-Zarazúa, M. (2020). Improving financial inclusion through the delivery of cash transfer programmes: The case of mexico's progresa-oportunidades-prospera programme. *The Journal of Development Studies* 56: 151–168, doi:10.1080/00220388.2018.1546845.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142: 698–714.

Mel, S. de, McKenzie, D. and Woodruff, C. (2012). Business training and female enterprise start-up, growth, and dynamics : experimental evidence from Sri Lanka. Policy Research Working Paper Series 6145, The World Bank.

Niño, M. (2013). Microcredit, Labor, and Poverty Impacts in Urban Mexico. *Review of Development Economics* 17: 747–762, doi:10.1111/rode.12063.

ONU-Habitat (2003). The challenge of slums: Global report on human settlements.

Ozier, O. (2018). The impact of secondary schooling in kenya: A regression discontinuity analysis. *Journal of Human Resources* 53: 157–188, doi:10.3368/jhr.53.1.0915-7407R.

Pitt, M. M. and Khandker, S. R. (1998). The impact of group-based credit programs on poor households in bangladesh: Does the gender of participants matter? *Journal of Political Economy* 106: 958–996.

Pitt, M. M., Khandker, S. R. and Cartwright, J. (2003). Does micro-credit empower women? evidence from bangladesh.

Planeación, D. N. de (2014). Índice de pobreza multidimensional para colombia (ipm-colombia) 1997-2010.

Porter, J. and Yu, P. (2015). Regression discontinuity designs with unknown discontinuity points: Testing and estimation. *Journal of Econometrics* 189: 132–147, doi:https://doi.org/10.1016/j.jeconom.2015.06.002.

Porter, M. (2016). Effects of microcredit and other loans on female empowerment in bangladesh: the borrower's gender influences intra-household resource allocation. *Agricultural Economics* 47: 235–245, doi:10.1111/agec.12225.

Pretis, F. (2022). Does a carbon tax reduce co2 emissions? evidence from british columbia. *Environmental and Resource Economics* doi:10.1007/s10640-022-00679-w.

Pretis, F., Reade, J. J. and Sucarrat, G. (2018). Automated general-to-specific (gets) regression modeling and indicator saturation for outliers and structural breaks. *Journal of Statistical Software* 86, doi:10.18637/jss.v086.i03.

Qiu, P., Asano, C. and Li, X. (1991). Estimation of jump regression function. *Bulletin of informatics and cybernetics* 24: 197–212, doi:10.5109/13419.

Rodríguez, D., Santana, M. and Pardo, C. F. (2015). *La motocicleta en America Latina: caracterización de su uso e impactos en la movilidad en cinco ciudades de la región.*

Ruiz, C. (2013). From Pawn Shops to Banks: The Impact of Formal Credit on Informal Households. *The World Bank Development Research Group* : 1–35.

Universidad de los Andes (2016). *Encuesta Longitudinal de Hogares Colombianos*. https://encuestalongitudinal.uniandes.edu.co/es/ [Accessed: 01.31.2022].

Vázquez, A. (2015). Determinantes Para Participar En El Mercado Formal De Productos Financieros: El Caso Del Ahorro Y Del Crédito En México. *Estudios Económicos CNBV* 3.

World Bank (2014). *Global Financial Development Report 2014 : Financial Inclusion*. https://openknowledge.worldbank.org/bitstream/handle/10986/16238/9780821399859.pdf?sequence=4&isAllowed=y [Accessed: 23.09.2022].

World Bank (2022). *The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19*. https://www.worldbank.org/en/publication/globalfindex [Accessed: 01.31.2022].

Yang, L. (2019). Regression discontinuity designs with unknown state-dependent discontinuity points: Estimation and testing. *Studies in Nonlinear Dynamics and Econometrics* 23, doi:10.1515/snde-2017-0059.

Zheng, J. X. (1996). A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics* 75: 263–289, doi:10.1016/0304-4076(95)01760-7.

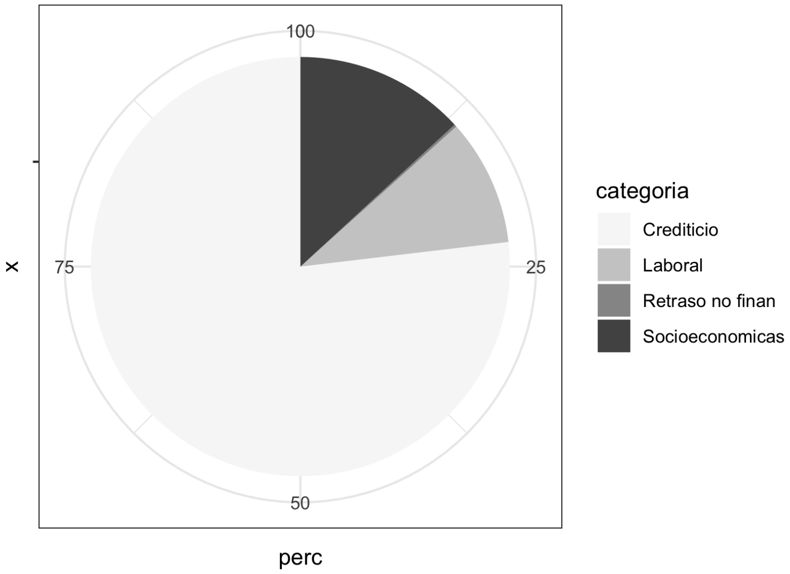
7 Appendix

A Dimensionality reduction methods

PCA is a multivariate analysis method used to reduce the dimensionality of data with a large number of variables while retaining as much as possible the variability of the original data set. Dimensionality reduction is achieved by creating a new set of variables known as components, each reproducing a part of the original variation. The components are arranged in descending order of their contribution to the initial set's variance, with the first component reproducing the larger part of the variability. It should be noted that each unit of observation, in this case a household, has a specific value for each principal component, known as a score. The standard principal component analysis is mostly used for a set of numerical variables, but in our exercise many of the variables considered are categorical (such as job quality, employment status, etc.). For this reason, we conduct categorical principal component analysis, which enables to quantify the categorical variables while reducing data dimension. The principal components are obtained through the spectral decomposition of the original data's correlation matrix, from which the eigenvalues and eigenvectors of the matrix are derived. Using the eigenvectors, the principal components are obtained as linear combinations of the original variables, while the eigenvalues represent the contribution of each component to the variability of the original data. Table 3 shows the eigenvalues obtained together with their contributions to the total variance of the data, while Figure 3 shows how the different categories of variables contribute to the first principal component This graph shows that credit-related variables are the main contributors to the first component, followed by socio-economic variables and job variables, respectively.

Factor analysis reduces the dimension of the data by explaining the correlation structure between the observed variables from a set of unobservable variables commonly known as factors. Here, the observed variables are linear combinations of the factors and error terms. The method serves to identify a set of unobservable latent variables underlying the observed data set. More precisely, observable variables are expressed in terms of common factors, unique factors and measurement errors. Common factors impact more than one observed variable and "factor loadings" are measures of the influence of a common factor on an observable variable. The maximum likelihood or principal axis factorisation method is commonly used to estimate factor loadings which are the regression coefficients between observed variables and factors. The factor loadings in Table 4 reveal how the variables contribute to the factors; the variables that contribute most to the first factor are the socio-economic variables and the features describing household credit.

Figure 18: PCA score and contributions by categories



(a) PCA Contribution

Source: Author's calculations. ELCA.

Table 3: PCA decomposition

Dimension	Eigenvalue	Proportion	Cumulative
1	4.18	16.06	16.06
2	2.42	9.3	25.36
3	1.97	7.58	32.94
4	1.71	6.56	39.5
5	1.22	4.71	44.21
6	1.1	4.24	48.44
7	1.05	4.02	52.46
8	1.01	3.89	56.35
9	1	3.86	60.21
10	0.99	3.81	64.02
11	0.96	3.71	67.73
12	0.94	3.61	71.35
13	0.93	3.57	74.91
14	0.88	3.38	78.29
15	0.83	3.21	81.5
16	0.78	2.99	84.49
17	0.72	2.79	87.28
18	0.68	2.61	89.89
19	0.65	2.49	92.38
20	0.56	2.16	94.54
21	0.42	1.63	96.17
22	0.38	1.46	97.62
23	0.33	1.26	98.88
24	0.19	0.74	99.62
25	0.05	0.2	99.82
26	0.05	0.18	100

Table 4: Factor loading

Factor1	Factor2	Factor3	Variable
0.2	0.718		Income
0.128	0.791		Inc percap
	0.458	0.127	Strata
	0.3		Adults
			sex
0.145	0.43		Education
	0.161		Age
	-0.196		Employment
	0.141	-0.104	Tenure
	0.136		active
0.971			Fee
0.971			Debt
0.324			Cons debt
0.69	0.252		Form credits
0.327	0.157		Term
0.788	0.209		Debt size
	0.103	0.992	Default bills
		0.776	Default

B Independent variables

Variable	Structure
Income	Numeric
Income Percap	Numeric
Strata	Categoric
Adults	Numeric
Gender	Categoric
Education	Numeric
Age	Numeric
Job status	Categoric
Tenant	Categoric
Assets	Numeric
Loan payments	Numeric
Balance Owed	Numeric
Consumer credit	Numeric
Number of Credits	Numeric
Max Loan Term	Numeric
Max Loan Amount	Numeric
Default	Categoric
Tenant Default	Categoric
Healthcare Default	Categoric
Banks Default	Categoric
Education Default	Categoric
Default	Categoric

C Structural change detection methods

One of the first works to estimate an unknown cut-off point in an RDD was Card et al. (2008). In their study on racial segregation, the authors examine the mass migration of white people (white flight) out of a neighbourhood when minority percentage in the neighbourhood exceeds a tipping point. This tipping point is unknown to the researcher, and it may vary depending on the racial tolerance of the population. Card et al. (2008) use a structural change search technique based on Hansen (2000), an extension of Andrews (1993). Similarly, recent but scarce literature has proposed adaptations of structural change methodologies to detect cut-off points from which regression discontinuity can be performed. Porter and Yu (2015) made the first proposal specifically designed for RDD, estimating the discontinuity point through the difference kernel estimator (DKE), defined as:

$$\widehat{\pi} = \ arg\ max\ \widehat{\alpha}^2(\pi) \tag{3}$$

where $\widehat{\pi}$ is the discontinuity point and $\hat{\alpha}$ is the treatment effect. The DKE was originally proposed by Qiu et al. (1991) to estimate 'jumps' in regressions. Additionally, the existence of the treatment effect is pre-established with a test statistic that checks for both the presence of the treatment effect first and then a possible individual-selection effect. This statistic is also based on structural change tests, particularly the non-parametric tests of Fan and Li (1996) y Zheng (1996). Crucially, Porter and Yu (2015) proved that $\widehat{\alpha}$ is asymptotically independent of $\widehat{\pi}$, demonstrating that the cut-off point estimation does not affect the efficiency of the treatment effect estimator and thus, given a $\widehat{\pi}$, the estimation of the regression discontinuity can be carried out as if the cut-off point were known from the start.

In contrast to the structural change approach, Herlands et al. (2018) propose a method from statistical learning to detect discontinuities. This method exploits the fact that RDD is a multi-point pattern and employs an anomalous pattern detection model to search for discontinuities. The search is performed in local neighbourhoods with a comparison of log-likelihood ratios between a null model that assumes no discontinuity and an alternative model that does assume it. Neighbourhoods are constructed from the assignment variable, and for each neighbourhood, points are repeatedly partitioned between two groups, which are then used to compute the alternative model. Given a set of validated discontinuities, the estimation of the effect is performed with a standard RDD using two-stage least squares or by a non-parametric estimator.

Other proposals for cut-point detection include Hansen (2017) for the case where the discontinuity is in the slope of the treatment function (the case of kink regression) and Yang (2019) for cut-points that vary depending on another variable.

D Khan's (2020) method

In Khan's method, first the Andrews (1993) test is implemented to detect unknown breaks in the probability of credit approval. Then, bands are estimated that define a neighbourhood of the assignment variable around the detected cut-off point. Finally, using only the sample within the neighbourhood, the discontinuity is again searched for using Pretis et al. (2018) indicator saturation method.

The Andrews' test determines the best discontinuity point candidate by identifying the largest structural change among other possible structural changes or outliers. In our case, this is convenient given that we expect the distribution of the probability of receiving credit will not have major jumps outside the cut-off point. The Andrews' test for a single structural change is defined as follows:

$$\begin{aligned} &H_0 : \beta_t = \beta_0, \forall t \geq 1 \\ &H_{1T} : \beta_t = \begin{Bmatrix} \beta_1(\pi) \\ \beta_2(\pi) \end{Bmatrix} \begin{matrix} \text{for t = 1 to } T_\pi \\ \text{for t = } T_\pi + 1 \text{ to T} \end{matrix} \end{aligned} \tag{4}$$

where π is the structural change and T_π indicates the index where that change occurs. This hypothesis test is carried out with the Andrews Sup F statistic, which is nothing else than the maximum of Chow's statistic for linear regression. The bands are then estimated with the MSE-optimal selectors proposed by Calonico et al. (2014). This algorithm seeks to minimise the Mean Squared Error (MSE) of the local polynomial estimator, given a polynomial order and a kernel.

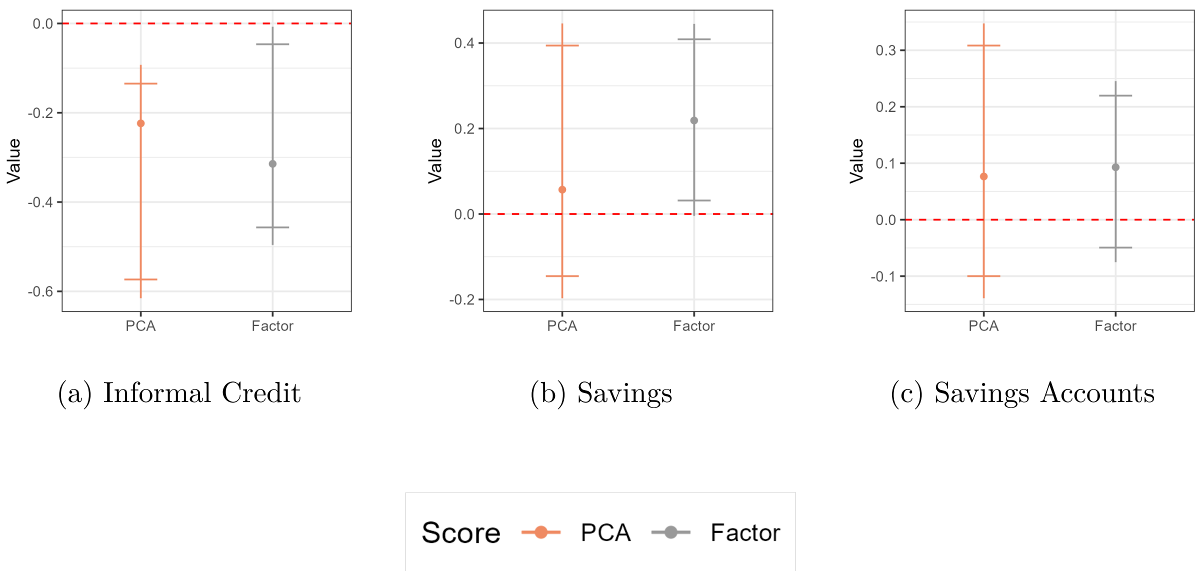
The indicator saturation method detects structural changes and outliers by starting with a general model allowing a jump at each point and then removing insignificant ones using selection criteria in an automatic search algorithm. Selection criteria include backward elimination, hypothesis testing, goodness-of-fit measures and diagnostics. The initial general model is as follows:

$$y_t = \mu + \sum_{j=1}^{n} l_j 1_{(t \geq j)} + \sum_{j=2}^{n} d_j 1_{(t \geq j)} + \sum_{j=2}^{n} k_j m_{(t \geq j)} + u_t \tag{5}$$

where μ is the expected value of y_t, the first summation corresponds to indicators to detect and control for outliers, the second gives indicators to detect and control for structural changes and the third represents other user-defined indicators which, in this case, are used to detect discontinuities in the slope. Since this method can detect multiple jumps, restricting the sample in the second step aims to reduce the number of potential points of interest. This method performs best when there is some prior knowledge about a potential cut-off point.

E Intention-to-Treat Effect

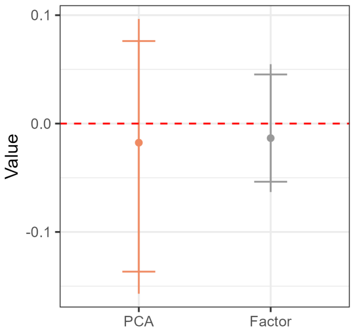
Figure 19: ITT on Financial Behavior



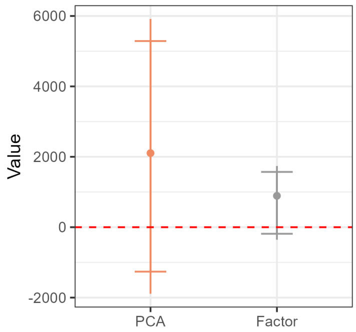
(a) Informal Credit (b) Savings (c) Savings Accounts

Source: Author's calculations. ELCA.

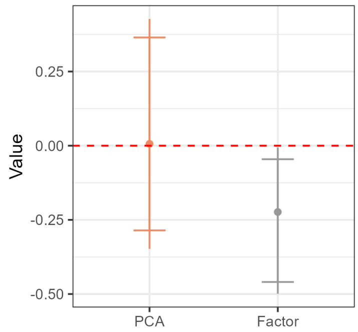
Figure 20: ITT on QoL



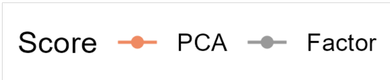
(a) Poverty



(b) Spending on Education

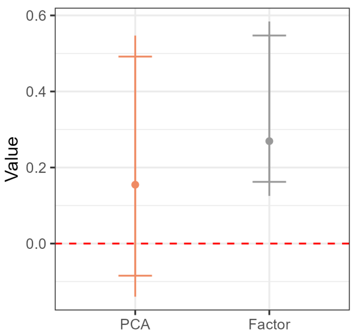


(c) Informal Housing

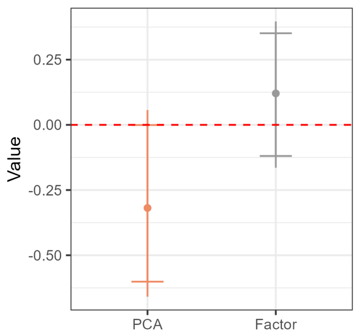


Source: Author's calculations. ELCA.

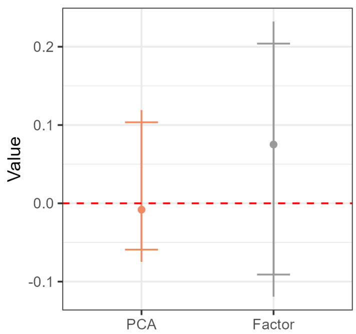
Figure 21: ITT on durables



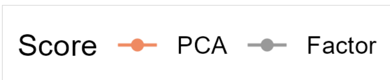
(a) Motorcycles



(b) Computers



(c) TV



Source: Author's calculations. ELCA.