Essays on Development Economics

Nicolás de Roux

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

ABSTRACT

Essays on Development Economics

Nicolás de Roux

This dissertation contains three essays in Development Economics. The first two chapters relate to the provision of credit for agricultural production in a developing country. The third chapter explores methodological issues in the measurement of risk aversion using laboratory experiments. Risk aversion has been suggested as a theoretical explanation for credit constraints in agricultural settings in developing countries. Better measures of risk aversion can be used to empirically validate these theories.

In Chapter 1 of this dissertation, I study the consequences of the use of credit scoring systems for agricultural lending in developing countries. Credit scoring has become a widespread tool to assess the creditworthiness of prospective borrowers, and has been found to increase efficiency and welfare in many settings. This chapter identifies a shortcoming in existing credit scoring systems that may lead to a market failure in agricultural lending in developing countries: Farmers' scores – and their access to credit – decline because of exogenous short-term weather shocks that do not reduce their likelihood of future repayment. I use data on the near universe of formal agricultural loans for coffee production in Colombia to show that excess rainfall shocks cause lower concurrent loan repayment, lower credit scores, and more frequent denial of subsequent loan applications. Then, I draw on the agronomic literature on coffee production and use survey data to show that productivity, income and repayment behavior recover faster from these shocks than farmers' credit histories. In the chapter I argue that these additional loan denials create costs for both farmers and the lender that could be avoided. The results presented in this chapter suggest that incorporating verifiable information on individual level shocks into credit scores would increase the efficiency of credit markets.

In Chapter 2, together with Jairo Esquivel, Margarita Gáfaro, Guillermo Otero and Moisés Mahecha, I study the determinants of repayment of loans to a public development bank. In particular, we investigate whether the public nature of a lender affects the repayment behavior of its borrowers. We conducted a field experiment where customers who receive reminder phone calls before a payment installment of loans were randomly assigned to different phone call scripts. The loans are from a public agricultural bank in Colombia. In our main treatment, we include a sentence to remind the customer of the public nature of the lender. We find strong and positive effects on repayment performance: farmers in this treatment have probabilities of ever being overdue and of entering into a period of 30 days past due that are respectively 10% and 22% lower than those of farmers treated with the traditional script. We interpret this finding as evidence that farmers are more like to repay their loans because of the public nature of the bank. Results from heterogeneity exercises show that some measures of state presence increase the magnitude of the effect of the public treatment, which suggests that state deterrence is a potential mechanism behind our findings. Furthermore, results from treatments where sentiments of altruism and peer pressure are induced by the script suggest that these motives explain part of the effect that the public nature of the lender has on repayment.

In Chapter 3, together with Juan Camilo Cárdenas, Christian Jaramillo and Luis Roberto Martínez, I address a methodological concern common in laboratory experiments. The house-money effect, understood as a person's tendency to be more daring with easily-gotten money, is a behavioral pattern that poses questions about the external validity of experiments in economics. We ran an economic experiment with 122 students, who received an amount of money with which they made risky decisions involving losses and gains; a randomly selected treatment group received the money 21 days in advance and a control group got it the day of the experiment. With our preferred specification, we find a mean CRRA risk aversion coefficient of 0.34, with a standard deviation of 0.09. Furthermore, if subjects in the treatment group spent 35% of the endowment (as they did, on average) their CRRA risk aversion coefficient is higher than that of the control group by approximately 0.3 standard deviations. We interpret this result as evidence of a small and indirect house money effect operating though the amount of the cash in advance that was actually spent. We conclude in this chapter that the house money effect may play a small role in decisions under uncertainty, especially when involving losses.

Contents

List of Tables

List of Figures

Acknowledgements

Pursuing a PhD in Economics is a unique but difficult experience. I couldn't have made it without the help and guidance of many people.

I am especially indebted to Eric Verhoogen, my principal advisor and sponsor. Eric took me on as his RA after I finished his class during my second year of the PhD, back in 2012. It was thanks to him that my interests in developments economics consolidated. His advice has guided me all these years. His support also allowed me to spend the last two years of my PhD in Bogotá, finishing this dissertation. What he thought me by example is the most important part of what I learned during my time at Columbia.

I am immensely grateful to Jonas Hjort and Suresh Naidu, whose guidance was essential to carry out the research of this dissertation. Pierre-André Chippori's advice improved substantially the first chapter, for which I am extremely grateful. I benefited from conversation and comments from other faculty at Columbia. I would like to especially thank Douglas Almond. Kiki-Pop Eleches, Supreet Kaur (now at U.C. Berkeley), Wojciech Kopczuk and Miguel Urquiola. I am also indebted to seminar participants of the Development and the Applied Microeconomics Colloquia of Columbia's Economics Department. I also thank Shane Bordeau and Amy Devine who made our lives as PhD students much easier.

My friends were a source of infinite support. Beatriz and Mario were always there for me in, during both good and bad times. My love and gratitude will be with them always. I also wish to thank Allan Aldana, Jorge Angarita, Natalia Botero, Davide Crapis, Santiago Cruz, Carlos del Toro, Gaelle Espinosa, Alejandro Fajardo, Natalie Gómez, Lina Henao, Pedro Herrera, Alejandro Hoyos, Ana Joaquina Jiménez, Cinar Kilcioglu, Ana Lucía Lalinde, Guillermo Llinás, Carolina Lozano, Lina Malagón, Sergio Martínez, David Mayo, Piero Montebruno, Santiago Melo, Francesca Moreno, David Monroy, José Antonio Ocampo, Federico Ramírez, Flor Paez, Ángela Pizarro, Daniel Rebolledo, Eduardo Restrepo, Natalia Ronderos, Carlos Rondón, Luis Felipe Saenz, Joseph Song, Natis Sotomonte, Carlos Suárez, Daniela and Juan Pablo Soulier, Víctor Saavedra, Juanita Vázquez, Camilo Vargas, Felipe Victoria and Juliana Velandia. Their friendship helped me keep my spirits high during these years.

My work and studies benefited from conversations with many friends who were pursuing a PhD at the same time as me. Among these, I wish to especially thank Anthony D'Agostino, Valentina Duque, Felipe Cortés, Pablo Egaña del Sol, Ferrán Elías, Margarita Gáfaro, Juan David Herreño, Ajin Lee, Ildiko Magyari, Luis Roberto Martínez, Juan Carlos Muñoz-Mora, Pablo Otonnello, Yogita Shamdasani, Teck Yong Tan, and Román Andrés Zárate.

The friendships of Zach Brown, Chris Hansman, Evan Riehl and Mica Sviatschi were my biggest joy and source of support during the years of the PhD. I wouldn't have made it without them and my work wouldn't be what it is without their help. I will always be indebted to them.

Diana, my love, came at the end of the PhD but held my hand through one of the hard periods of the PhD. She has been a source of infinite strength and joy that helped me carry on.

My family was also an unbounded source of support. I am indebted to my uncles Blanca, Harold, Tao, Toño, Lia, Pacho, Gustavo and María Cecilia. To my brothers Matías and Daniel, I owe the strength that comes from the thought that I will never

be by myself. My efforts come in part from my desire to set a good example for them.

Finally, I owe so much to my parents and their love that I will refrain from trying to express my gratitude with words. This dissertation is dedicated to the two of them.

Al Papá, que me dio mi pasión.

Y a la Mamá, que me enseñó a encontrar las fuerzas para seguirla.

Chapter 1

Credit Scoring Meets Agricultural Lending: Exogenous Shocks, Recovery, and Access to Formal Credit

Nicolás de Roux[†]

[†] I am especially grateful to my advisors Pierre-André Chiappori, Jonas Hjort, Suresh Naidu, and Eric Verhoogen. For useful comments I am grateful to Emily Breza, Zach Brown, Juan Camilo Cárdenas, Felipe Cortés, Xavier Durán, Andrés Escobar, Marcela Eslava, Alejandro Fajardo, Margarita Gáfaro, Chris Hansman, Juan Herreño, Marc Hofstetter, Supreet Kaur, Wojciech Kopczuk, Ajin Lee, Ana María Ibáñez, Luis R. Martínez, Josepa Miquel-Florensa, Andrés Moya, Juan Carlos Muñoz-Mora, José Antonio Ocampo, Kiki Pop-Eleches, Evan Riehl, Yogita Shamdasani, Mica Sviatschi, Miguel Urquiola and to seminar participants at Universidad de los Andes, Universidad Javeriana and Universidad del Rosario. This project could not have been carried out without the hospitality of the Banco Agrario de Colombia and the knowledge of its staff. Many thanks to all members of the Vice-presidency of Risk and the Information Management Unit, and especially to Jairo Esquivel, Moisés Mahecha and Guillermo Otero. Finally, I am grateful to the Federación Nacional de Cafeteros, and in particular to Óscar Bernal, Julián Cantor and Marcela Urueña, and to the the Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia. All errors are my own.

"If any one owe a debt for a loan, and a storm prostrates the grain, or the harvest fail, or the grain does not grow for lack of water; in that year he need not give his creditor any grain, he washes his debt-table in water and pays no rent for his year."

— Hammurabi's Code (c. 1760 B.C)

1.1 Introduction

Credit scoring has spread quickly across the banking and lending industry in both developed and developing countries (Mester, 1997; Berger, Frame, and Miller, 2005; de Janvry, McIntosh, and Sadoulet, 2010). Theoretically, this tool has the potential to lessen problems of moral hazard and adverse selection (Pagano and Jappelli, 1993; Giné, Goldberg, and Yang, 2012) and many empirical studies have found welfare and efficiency gains from its use (Einav, Jenkins, and Levin, 2012; Einav, Jenkins, and Levin, 2013; Giné, Goldberg, and Yang, 2012).

Existing credit scores are primarily calculated using fixed borrower characteristics and information on prior repayment behavior. These methods are designed for situations in which a borrower's repayment depends largely on her type or her own actions. But credit scores typically do not take into account information on exogenous shocks that affect the economic environment in which borrowers are operating. Therefore, they may not be as well suited for settings in which large, exogenous shocks that influence repayment are important.

This paper documents a market failure that results from an important shortcoming of credit scores used in agricultural lending in developing countries. In particular, I study a setting in which farmers' scores and their access to credit decline because of exogenous short-term shocks that do not reduce their likelihood of future repayment. To do this, I use a novel administrative data set with the near-universe of formal loans to small farmers in Colombia. I first show that exogenous weather shocks lead to lower repayment of concurrent loans, lower credit scores, and more frequent denial of subsequent loan applications. I then show that income recovers from these shocks faster than farmers' access to credit. Furthermore, I present evidence on recovery in loan payments in two different samples, consistent with the recovery in income. These findings imply a market failure because banks using information on weather shocks would want to lend to farmers who are currently denied credit. The implication is that the efficiency of the credit market could be enhanced if credit scores and credit histories were modified to account for these exogenous shocks.

From a theoretical perspective, the market failure arises from the fact that a signal that conveys information on the action and the type of the borrower is not incorporated in the contract with the bank. In my setting, rainfall shocks provide such information (for example, the actions of a borrower who defaults in the absence of a rainfall shock are likely different from the actions of a borrower who defaults when a rainfall shock occurs) and should be incorporated in the credit score or, more broadly, in the contract between the farmer and the bank.[1]

To document the channels described above I proceed in two steps. In the first part of the paper, I estimate the effect of rainfall shocks on concurrent loan repayment, credit scores, and future access to credit, and I establish that this effect is persistent. In the second part, I argue that, on average, farmer income from coffee production and repayment ability recover faster than farmer credit histories after a shock.

For the first step, I use administrative data on loan applications, credit scores, and repayment behavior from the Banco Agrario de Colombia (BAC) [Colombian Agrarian Bank]. The BAC is a publicly owned bank that was created in 1999 to finance rural productive activities. In 2013, it held 97% of formal agricultural loans to small farmers in Colombia (DNP and FINAGRO, 2014). I merge BAC's individual data from 2005 to 2015 with administrative data

[1]The Informativeness Principle states that "any additional information about the agent's action, however imperfect, can be used to improve the welfare of both the principal and the agent" (Holmstrom, 1979). In this regard, in the Appendix I show in a simple model of borrower screening how the the inclusion of observable exogenous shocks in the credit score can reduce the probability of both an inclusion error (lending to a un-profitable borrower) and an exclusion error (denying credit to a un-profitable one).

from the Federación Nacional de Cafeteros (FNC) [National Federation of Coffee Growers] that allows me to recover geographical coordinates for each farmer and link the loan to the closest rainfall station at the time of loan disbursement.

I estimate the effect of excessive rain shocks during loan tenure on repayment in a sample of loans destined for coffee production. Excessive rainfall is associated with lower productivity of coffee trees.[2] Therefore, it can lead to a fall in farmer income and to lower loan repayment, as long as insurance markets are incomplete.[3] The estimation specifications in these exercises include rainfall station and quarter-year of disbursement times maturity fixed effects. Thus, the identifying assumption is that the occurrence of shocks at different points in time for a given rainfall station is not systematically correlated with other time-varying factors that affect repayment of outstanding loans of farmers close to the station.

I find that excessive rainfall decreases the probability of loan repayment. In particular, loans that experience excess rainfall shocks during their first year have a probability of entering into a period of 30 days of overdues that is 22 percent larger than loans with no shocks. This leads the BAC to downgrade the scores of its clients that are reported to credit bureaus.

I then estimate the effect of rainfall shocks on subsequent loan applications. The loan application process in the BAC consists of two stages. I find that the score reported by the credit bureau is lower and that denial at both stages of the application is more likely for farmers who experienced a shock during the previous loan tenure. For example, in my preferred specification, the probability of loan denial is 13 percent larger in the first stage of the process for applications following loan tenures with rainfall shocks. In the second stage the probability of denial increases by 10 percent. Using a subset of the sample of this exercise, I show that even after two years since the maturity of the first loan, applications following loan tenures where a shock occurred have lower credit scores and are more likely to be denied.

[2] I focus on excessive rainfall since coffee production is more sensitive to it than to a lack of rain. See for example Boucher and Moya (2015).

[3] This is the case for agriculture in Colombia, and in particular for coffee production (Boucher and Moya, 2015).

Having documented the effect of rainfall shocks on repayment and future access to credit, I proceed to the second step, which shows that income recovers faster from rainfall shocks than the credit history of the farmer. Based on an extensive agronomic literature, I argue that excessive rainfall in the year before the harvest affects the productivity of coffee trees, but the next harvest is not affected after weather returns to normal.[4] This suggests that the effect of the shocks on the productivity of the coffee tree dies out after one year at most. To test this in the data, I use a representative survey of coffee farmer sales in 2005 to show that income recovers from excessive rainfall shocks on average one year after the shock. The identifying assumption for this exercise is that, in the cross-section, the occurrence of excess rainfall shocks is not correlated with unobservables that might affect coffee sales. I present various robustness checks that support it.

A question that arises at this point is whether the recovery in income translates into a recovery in repayment. One would expect this to be the case since income is the main determinant of the ability of lenders to repay their loans.[5] I present two pieces of evidence of repayment recovery after rainfall shocks. First, in a sample of loans with maturities of five years or more, I show that repayment recovers two years after a shock in the first year of the five year loan tenure. Second, in a sample of individuals with high credit scores, I show that a shock during the tenure of an initial loan causes lower repayment but has no significant effect on acceptance of subsequent loan applications. This suggests that selection driven by shocks is not an important concern in this sample. More importantly, I find no significant effect of shocks during the first loan on repayment of the subsequent loan.

The idea that a transitory shock can affect credit histories, scores, and access to credit

[4]The transitory nature of rainfall shocks is a common supposition in the development literature. See for example Kaur (2015).

[5]This idea is consistent with the fact that the BAC's repayment schedule is organized to match the harvests of its farmers. Empirical evidence on this matter is scant but still suggestive of such relation. For example, Chirwa (1997) finds that crop sales of small farmers are associated with higher loan repayment in Malawi. Acquah and Addo (2011) find that higher fishing income is associated with higher repayment of loans to fishers.

long after the recovery of productivity, income and repayment propensities can also apply to contexts other than agricultural lending in developing countries. A failure to take exogenous shocks into account could generate inefficiencies in other credit markets. If transitory shocks can be observed, and the forces driving the recovery are well understood and unrelated to lender characteristics, the costs they generate could be mitigated. As far as I am aware, this is the first paper to empirically document this mechanism. Closely related, Avery, Calem, and Canner (2004) argue that situational circumstances that temporarily affect repayment can generate problems for credit scores in consumer lending markets.[6]

If exogenous shocks generate costs for the lender, why aren't they taken into account? This fact is more puzzling for agricultural credit markets in developing countries, since shocks are in principle easy to observe. The quote from Hammurabi's Code even suggests that using this information is a matter of common sense. But below, I argue that technological constraints only recently allowed for shocks to be measured with the level of precision needed to incorporate them in credit scores.

My paper contributes to four different areas of research in development economics and finance. First, it adds to an empirical literature that estimates the impacts of credit scoring and credit histories (Einav, Jenkins, and Levin, 2012, 2013; Giné, Goldberg, and Yang, 2012; de Janvry, McIntosh, and Sadoulet, 2010). This literature has found almost uniformly that credit scoring leads to better credit market outcomes. My paper points to an inefficiency that arises in contexts where repayment is affected by orthogonal and transitory shocks and they are not recorded in credit histories or included in credit scores. These findings suggest

[6]Using data from the US, the authors study the effect of county-level unemployment rates and changes in marital status on repayment performance. They find that both variables affect loan repayment after controlling for ex-ante credit scores. Nevertheless, they mention that solving the problem that these circumstances create for credit scoring models is difficult, since information on these episodes is not available within credit reporting agencies. Furthermore, the circumstances they consider are not exogenous, in the sense that their occurrence is likely correlated with borrower characteristics. In my case, rainfall shocks are orthogonal to individual characteristics and they are in principle easy to incorporate in credit histories and scores. Furthermore, it is unclear if individuals will recover from the circumstances the authors consider. For example, a divorce can have long term consequences on borrower's income and repayment.

that the optimal information set to estimate credit scores is highly context-specific. In this sense, this paper also contributes to work that investigates how to improve credit scores.[7]

Second, this paper adds to the research that explains why individuals are credit constrained, in particular in the agricultural sector in developing countries (Conning and Udry, 2007; Giné, Goldberg, and Yang, 2012).[8] This paper shows that farmers can lose access to formal credit simply from the fact that agricultural production generates volatile income streams that lead to loan default. This loss of access can be persistent due to worse credit histories and credit scores.

Third, this paper contributes to a large literature that studies the effects of information sharing through credit bureaus and the use of credit reports (Sharpe, 1990; Pagano and Jappelli, 1993; Vercammen, 1995; Padilla and Pagano, 2000; Jappelli and Pagano, 2002; de Janvry, McIntosh, and Sadoulet, 2010; Giné, Goldberg, and Yang, 2012; González-Uribe and Osorio-Rodríguez, 2014; Dobbie, Goldsmith-Pinkham, Mahoney, and Song, 2016; Garmaise and Natividad, 2016).[9] My paper contributes to this literature by providing evidence on the costs of credit reports that do not differentiate among reasons for default. The consequences of this effect can be amplified when information is shared between lenders and the reasons for the default are not specified in the credit history.[10]

[7]See for example Rocha Sousa, Gama, and Brandao, 2016 and Wei, Yildirim, Van den Bulte, and Dellarocas, 2015. To the best of my knowledge, the credit scoring industry has not yet started to incorporate or search for information on exogenous shocks.

[8]A large agricultural-economics literature has documented the costs of credit constraints for farmers. See for example Carter and Olinto (2003), Petrick (2004), Guirkinger and Boucher (2008) and Fletschner, Guirkinger, and Boucher (2010) .

[9]Of these, Garmaise and Natividad (2016) is the most related to this paper. The authors show that random credit rating downgrades (that is, credit downgrades not related to default) in consumer credit markets in Peru cause a three-year reduction in financing. These downgrades are generated by credit rating thresholds and not by real shocks affecting repayment The main difference between the two studies is that in my case exogenous shocks do have real consequences but farmers' repayment recovers faster than the associated negative credit reports.

[10]For example, a farmer might refrain himself from applying for a production loan if he anticipates to need a loan (say a consumer loan) in a not-so-distant future. This idea is consistent with the

Fourth, this paper adds to a literature that shows that transitory shocks can have long term macroeconomic consequences (Blanchard and Summers, 1986; Ouyang, 2009; Eslava, Galindo, Hofstetter, and Izquierdo, 2010; Ball, 2014; Barrot and Sauvagnat, 2016) and also long term effects on farmer production (Rosenzweig and Wolpin, 1993).[11] After a weather shock, farmers can lose access to credit, potentially hurting productivity enhancing investments. This mechanism could operate in other credit markets besides the agricultural setting. Firms may lose access to credit due to shocks orthogonal to their characteristics and from which they recover faster than their credit histories. This in turn can lead to costs in terms of investment and total factor productivity. As far as I can tell, my paper is the first to suggest that short term real shocks can have long term consequences through use of credit scoring and credit histories.

The rest of this paper is organized as follows. Section 2 provides some background. In Section 3, I study the effect of rainfall shocks on repayment, credit scores, and future access to credit, and I document the persistence of these effects. In Section 4, I study recovery of coffee tree productivity, recovery of income after rainfall shocks, and repayment recovery. Finally, Section 5 concludes.

---

finding of Liberman (2016) that borrowers of a department store in Chile are willing to pay 11% of their monthly income for a good credit reputation.

[11]In doing so, a side contribution of my paper relates to a voluminous literature that studies the effects of weather on economic activity. Dell, Jones, and Olken (2014) provide a summary of this research. My paper documents a new margin in which weather matters: subsequent access to formal credit. Estimations of the effect of rainfall on repayment are scant. As far as I can tell, the only other work that estimates at the farmer level this effect is the paper by Pelka, Musshoff, and Weber (2015). My study is the first one to estimate the effect of rainfall shocks on repayment using coordinates of individual farms. See Castro and Garcia (2014) for an estimation of the effect of climate variations in a structural default risk model, using aggregate data from the BAC.

1.2 Agriculture in Colombia and the Banco Agrario

1.2.1 Agricultural Financing

As is frequent in developing countries, the rural sector in Colombia is under-supplied with capital. According to a national census conducted by the Colombian official statistical agency, DANE, in 2013 only 11% of agricultural producers demanded credit (either formal or informal). Of these, 89.6% got a loan. The ELCA rural panel,[12] representative of four rural regions in the country, shows better results. In 2013, 52% of households had no credit, 25% had one or more formals loans (and no informal loans), 16% had one or more informal loans (and no formal loans) and 7% had a formal and an informal loan (or more) (Cadena and Quintero, 2015). Therefore, at least according to the ELCA, formal loans are the main source of capital for rural households.

Despite these low rates of credit access in Colombia, the BAC is the main player in this market. The BAC is a publicly owned bank, created in 1999, with the mandate from the government to finance agricultural productive activities in the country.[13] In 2013 it lent 97% of formal loans to small farmers[14] (DNP and FINAGRO, 2014) and it is the only bank in the country to have presence in all of its 1123 municipalities. 89% of the banks' branch offices are in rural areas. For small farmers, collateral is guaranteed with resources from the Fondo Agropecuacio de Garantías, (FAG) [Agricultural Guarantees Fund], a public fund created for

[12]The ELCA panel is a longitudinal survey conducted by the Universidad de los Andes. It carries out surveys for both rural and urban areas. Two waves have been conducted thus far, in 2010 and 2013. For rural areas, around 4500 households where interviewed in 17 municipalities, representative of four rural regions of Colombia.

[13]According to the bank's statutes, 70% of the loan balance has to correspond to agriculture related activities.

[14]Government institutions define a small farmer as one with an amount of total assets in Colombian Pesos smaller than 120 millions (around $US 39000, at the average daily exchange rate of the first semester of 2016 equal to 3098 Colombia Pesos per US$). This includes all household assets (for example a car or a television), and not only those used for farm production.

this purpose.[15]

The fact that the government provides the collateral for loans might suggests that the BAC has no incentives to screen borrowers. This is not the case though as indicated by the following three facts. First, the process to recover the guarantee is cumbersome and costly. Second, the regulations that apply to the BAC are the same as those of other banks and the main indicators of the bank's performance are tied to borrower default. If these indicators turn out to be unfavorable, financial authorities can intervene. Third, and as discussed in more detail below, the BAC has developed its own models to screen borrowers.

1.2.2 The BAC's Loan Application Process

The typical farmer who applies for a loan from the BAC does so in the branch office closest to his farm. In the office, a loan officer does a consult with CIFIN, the credit bureau in business with the BAC. From this consult, a report is issued. It contains information on the credit history of the borrower, the CIFIN Credit Score (which results from CIFIN's own credit scoring model), and whether or not the application process can continue. At this point the application process can end with the farmer being denied credit. The immediate denial at the CIFIN stage depends on the CIFIN Credit Score, on other variables in the farmer's credit history and on the BAC's policies. For example, there is a threshold set by the BAC for the CIFIN score. If the farmer's score is below this threshold, the application is denied. There are other variables that determine if the application is denied at this stage. An example are periods with overdues during past loan tenures. The BAC's policies regarding the CIFIN consult are constantly being changed and tuned up. I refer to this stage of the application

[15] The FAG is administered by the Fondo para el Financiamiento del Sector Agropecuario (FINAGRO) [Fund for the Financing of the Agricultural Sector]. FINAGRO is a second tier bank also created with the objective of financing rural productive activities. FINAGRO lends "rediscount" funds to first tier banks. It lends at an interest rate i and first tier banks lend at a rate $i + x$ to agricultural producers. In 2013, 85% of FINAGRO's "rediscount" resources where allocated by the BAC (DNP and FINAGRO, 2014). This is another indicator of the BAC's importance for financing small farmers in the country.

process as the *CIFIN Consult Stage*.

Once the loan application makes it through the CIFIN stage, it arrives to the main offices of the BAC in Bogotá, the capital of Colombia. There, each loan application is individually revised by a credit analyst. Different inputs are taken into account by the credit analyst when he reviews the application. An important one is a score obtained from the BAC's own credit scoring model (which is different from CIFIN's model). Other examples are the projected income and cost flows of the project the farmer wants to finance. The credit analyst makes the final decision on loan approval. I refer to this stage of the application process as the *Analysis Stage*.[16]

1.3 The Effect of Rainfall Shocks on Repayment, Credit Scores and Future Access to Credit

In this section of the paper I study the effects of rainfall shocks on repayment behavior, credit scores and future access to credit and the persistence of this effect. I divide the presentation of the results in two parts. First, I study the effect of rainfall shocks on repayment behavior. Next I study their effect on subsequent access to credit and the persistence of this effect.

1.3.1 Effect on Repayment

1.3.1.1 Data and Construction of the Estimation Sample

The following are the three main data sources that I use to estimate the effects on rainfall shocks on repayment behavior. First, I use administrative records from the BAC. Second, I use data from the Sistema de Información Cafetera (SICA) [System of Coffee Information] from the FNC. The SICA contains yearly information on plot characteristics from coffee

[16] The application process has seen a series of changes in time. For example, the BAC's model used to score small farmers entered in operation in 2012.

farmers in Colombia that interact with the FNC. Finally, I use rainfall data from the Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia, (IDEAM) [Institute of Hydrology, Meteorology and Environmental Studies].

To construct the estimation sample, I first take all the farmers that are ever observed in the SICA data from 2006 to 2014. The SICA is updated every time a farmer interacts with the FNC. This interaction is frequent since coffee farmers receive social services and technical support from this institution. Also, many smaller cooperatives of coffee growers are affiliated to the FNC, and through them, farmers interact with the FNC (Muñoz-Mora, 2016).

For the list of farmers that ever appeared in the SICA data in 2006-2014, I obtain the loans that were disbursed by the BAC in the period of 2005-2011. For this set of loans, I select those of farmers with at least one farm observed in the SICA in the year of loan disbursement. The largest farm is then used to link the loan to the rainfall station. The farms in the SICA and in the IDEAM rainfall station are geo-referenced. For each farmer, I use the Euclidean distance to find the closest rainfall station to the largest farm at the time of loan disbursement. In panel A of Figure 1.1, I depict the distribution of coffee farms in the SICA across Colombia, for the years 2010-2013. Panel B shows the distribution of the rainfall station that are close to at least one farm in the SICA. Figure 1.1 shows that the distribution of the IDEAM rainfall stations is dense in coffee growing areas of Colombia.[17] Therefore, the closest rainfall station provides a good measure of the extreme weather events that the farmers faces. Finally, from the set of loans linked to a rainfall station, I take only loans whose destination is related to coffee production.[18] I consider all maturities of coffee

[17]As indicated below, the mean distance between the farm and the closest rainfall station in my main estimation sample is 6.5 kilometers.

[18]The BAC data has detailed descriptions of the destination of the loans. The following are two examples: 1) "Traditional coffee - sustainment of agricultural production": refers to loans for sustaining traditional coffee production; which include for example fertilizer purchases. 'Traditional', as opposed to 'technical' refers to coffee production with parameters such as type of coffee seed that are not as good as the ones used in 'technical' production. 2) "Coffee renovation by plantation - plantation and maintenance": broadly refers to loans destined to the plantation and maintenance of new coffee trees.

loans.[19]

Since the focus of the paper is on the effect of rainfall shocks on loan applications that follow an initial loan tenure, I construct a sample that allows me to look at this effect. That is, I select a sample of farmers with an initial loan who then applied for a subsequent loan. I refer to this sample as the sample of *Loans With Posterior Applications*. Note that this sample is different from the sample of all disbursed loans. To construct it, I start with the most recent loan originated in 2008-2011.[20] This results in a sample of one loan per farmer.[21] For each farmer, I then find the first application after the maturity of the initial loan in both the CIFIN Consult data and the Analysis Stage data.[22]

1.3.1.2 Definition of Rainfall Shocks

I now describe the definition of rainfall shocks for which I use historical data from the IDEAM. I use monthly data on precipitation available from 1982 to 2012. I construct rainfall shocks

[19] There are close to 822,000 coffee farmers ever observed in the SICA data in 2006-2014. For them there are a total of 498,000 loans in the BAC data in the period 2005-2011. Around 299,000 (60%) correspond to coffee production and close to 274,000 can be linked to a rainfall station at the quarter of loan disbursement. By selecting the sample in this manner, I ensure that the farm used to link the rainfall station to the loan is observed in the SICA data in the same year of loan disbursement and is associated to the same farmer.

[20] The reason to use as starting date 2008 is that in that year the Law 1266/ 2008 was introduced. This law required credit bureaus to erase negative information (past overdue periods for example) that was sufficiently old by 2008. For details on the Law see González-Uribe and Osorio (2016). To avoid any confounding effects of the law I start with loan that where disbursed in 2008 or afterwards. Also, I study the effect on the CIFIN Consult Stage of an application following an initial loan. Data of the CIFIN consults is only available starting in 2010. This restriction on the sample ensures that the maturity of the original loan is close enough to 2010, so I can find in the CIFIN data the subsequent loan application.

[21] For reasons explained bellow (see in particular footnote 27) I work with quarterly dates. For a few cases, there are multiple loans for the same farmer originated in the same quarter. When this is the case, I randomly select one of the loans.

[22] Note that for a given initial loan, the first application in the CIFIN consult does not need to correspond to the first application that makes it to the analysis stage. This would be the case, for example, if after the maturity of the initial loan, the application is denied at the CIFIN Stage and then, after some time, the farmer applies again and passes the CIFIN stage.

at the rainfall station level. For each station I add observed monthly precipitations to obtain a quarterly measure of precipitation. Then, I define shocks at the calendar-quarter level in the following manner: for each quarter I obtain the 80th percentile of the observed rainfall distribution from 1982 to 2012. I say that in a given quarter there was an excessive rainfall *quarter-shock* if observed rainfall was above the 80th percentile of the corresponding distribution.[23] Note that under this definition, for each calendar year I can observe up to four quarter-shocks in each station. For a given calendar year, I will say that a rainfall shock occurred (and refer to it as *Rainfall Shock*) if at leas two of these quarter-shocks occurred.[24] This definition takes into account seasonality both at the station and at the calendar-quarter level. According to this definition, a shock occurs if in a given year, rainfall is particularly high compared to the historical raininess of the rainfall station during that year.[25]

[23]I focus on excessive rainfall shocks since coffee growing in Colombia is more sensible to periods of excessive rain according to conversations with personnel from the FNC. In estimations not reported in this paper I find no effect of shocks of low rainfall. See also Gast, Benavides, Sanz, Herrera, Ramírez, Cristancho, and Marín (2013), Boucher and Moya (2015), and Turbay, Nates, Jaramillo, Vélez, and Ocampo (2014).

[24]My results are robust to considering different definitions of shocks. For example if I consider the number of quarter-shocks in a given calendar year, I find similar effects on repayment and subsequent access to credit.

[25]My definition of shocks is similar to that of Kaur (2015) and Jayachandran (2006). They define shocks at the district-year level, using the historical rainfall distribution. Excessive rainfall shocks are also defined using rainfall realizations above the 80th percentile. As Kaur (2015) points out, this definition captures the non-linear relation between rainfall and crop productivity. A similar non-linear relation likely exists between rainfall and the coffee tree's productivity which is highly sensible to extreme changes in weather (Turbay, Nates, Jaramillo, Vélez, and Ocampo, 2014). In Section 1.4.1, I describe in detail the agronomic mechanisms through which coffee production is affected by excessive rainfall.

1.3.1.3 Empirical Approach

I estimate by OLS the following linear-probability model of repayment of the initial loan in the sample of Loans With Posterior Applications:[26]

$$y_{ijm\tau} = \beta s_{j\tau} + \mu_{m\tau} + \delta_j + \epsilon_{ijm\tau} \tag{1.1}$$

where $y_{ijm\tau}$ is a dummy equal to 1 if loan i, close to rainfall station j, of maturity m and originated in quarter τ, was ever overdue by 30 days or more. $s_{j\tau}$ is a dummy equal to 1 if a rainfall shock occurred in the first year after loan disbursement.[27] I focus on the effect of rainfall shocks during the first year of loan maturity. The reason is that my sample consists of a continuum of maturities starting with maturities of less than a year. $\mu_{m\tau}$ are quarter-of-disbursement times maturity fixed effects.[28] δ_j denote rainfall station fixed effects and $\epsilon_{ij\tau}$ is a mean-zero error term. The coefficient of interest is β and it is expected to be greater than 0, indicating that rainfall shocks increase the probability of entering in a period of 30 days with overdues.[29]

---

[26]In Appendix A, I present robustness exercises where I estimate the effect in a cross-section of loans disbursed in 2005-2011. In that sample, I do not select a single loan for each farmer and instead consider all loans disbursed. The effects are very similar to the ones estimated in the main text.

[27]Since I define rainfall shocks using calendar quarters and also for simplicity, I handle other dates at the quarter level too. For example, if a loan was disbursed in March 2007, I say that the loan was disbursed in the first quarter of 2007 (denoted by 2007-1). For that particular loan, $s_{j\tau}$ will take a value of 1 if at least two of the quarters in {2007-1, 2007-2, 2007-3, 2007-4} where quarter-shocks.

[28]Although I have a continuum of maturities, its distribution is bimodal as shown in Figure **??**. The first mode corresponds to loans between one or two year maturities which are usually related to the sustainment of agricultural production (e.g. buy fertilizer). The second mode corresponds to loans of five years or more which correspond to long term investments like planting new coffee trees or investing in production infrastructure. Therefore I define my set of maturity fixed effects as consisting of a single dummy equal to one for loans of maturities of three years or more.

[29]My main outcome of interest is whether or not the loan entered in a period of 30 days past due. The reason is that this is the main indicator of default used by the BAC. According to its statistics, 80% of loans that enter in a 30 days overdue period end up in longer periods with overdues. The 30 days overdue period is also the main indicator that financial regulations in Colombia require banks to use. According to BAC officers, this is the main indicator of default used in other countries as well.

The inclusion of $\mu_{m\tau}$ and δ_j implies that the coefficient β is identified from variation in $s_{j\tau}$ across time within rainfall station, for loans of similar maturity. Therefore the coefficient β is identified and has a causal interpretation as long as that the occurrence of shocks for a given rainfall station is not systematically correlated with other time-varying factors that affect repayment of outstanding loans linked to the station. Formally, I assume that:

$$E[\epsilon_{ijm\tau}|s_{j\tau}, \mu_{m\tau}, \delta_j] = 0 \tag{1.2}$$

This assumption is reasonable given my definition of rainfall shocks. $s_{ij\tau}$ captures atypical rainfall realizations for a given year in a given station. This variation is likely to be uncorrelated with other time varying factors that affect repayment of loans linked to the rainfall station.

1.3.1.4 Results

Table 1.1 shows summary statistics of loan characteristics in my estimation sample. The average loan has a maturity close to two years and an annual interest rate of 11%. The mean distance between the farm and the closest rainfall station is 6.54 kilometers. The mean value of $s_{j\tau}$ is 0.42 and the 75th percentile is 1. This captures the fact that the period I consider was a particularly rainy one because of the occurrence of the "La Niña" climatic phenomenon in the second half of 2010 and in 2011.[30] On average, 14% of loans entered in a period of 30 days past due and 10% entered in a period of 60 days past due.

Table 1.2 presents the results from the estimation of equation 1.1 (in the sample of Loans

---

Furthermore, in the BAC data there is not a variable for default, since this is a concept that can be defined in different ways.

[30]La Niña [The Girl] is an ocean-atmospheric phenomenon. It is generated by lower than average temperatures of the Pacific Ocean, in its Eastern-Central part. La Niña is associated with higher levels of rain and cloudiness in some countries of South America like Colombia, Ecuador and Peru (Gast, Benavides, Sanz, Herrera, Ramírez, Cristancho, and Marín, 2013). Its counterpart, El Niño [The Boy] phenomenon, is associated with lover temperatures than average and has the opposite effects on rainfall than La Niña.

With Posterior Applications), and different variants of it. Each column corresponds to a different regression. All the estimations include rainfall station fixed effects. The set of quarter-of-disbursement times maturity fixed effects varies with the specification as indicated in the bottom of the table. I cluster the error terms at the rainfall station level. The first three columns of the table correspond to the baseline specification. Column (1) reports the coefficient of estimating exactly equation 1.1. It implies that a rainfall shock during the first year of loan tenure increases the probability of entering into a period with overdues in 22% compared to loans with no shock. In Column (2), I report results of an alternative definition of shocks where I use as dependent variable the number of quarter-shocks in the first year after loan disbursement. As explained before, this variable can take a value of 1, 2, 3 or 4. The coefficient implies that each additional quarter-shock in the first year increases the probability that the loan enters into a period of 30 days past due in 11.3%, compared to mean of the outcome in the entire sample.

The next four columns present robustness exercises of the previous results. Column (3) shows that a rainfall shock increases the probability of entering into a period of 60 days past due in 19%. This implies that the effect documented before is also present for longer overdue periods, consistent with the fact that most loans that enter into a period of 30 also enter longer overdue periods (see footnote 29). Column (4) shows the results for a subsample where the distance between the farm and the rainfall station is lower than the mean of the full sample (equal to 6.54 km). The estimated coefficient is again positive and significant at the 5% level. Column (5) shows results for the subsample of short terms loans. These are loans with maturities of up to three years. The coefficient is almost identical to that of the baseline scenario. Finally, column (5) reports results for long term loans. I define long term loans as loans of maturities of three years or more. Again, a rainfall shock in the first year of loan tenure increases the probability that the loan enters into a period of 30 day or more past due, relative to long term loans with no shocks in the initial year. The economic magnitude of these coefficients is sizable. For the baseline case (Column (1)) it corresponds to a 22% increase relative to loans with no shocks. This is saying that because of exogenous reasons outside of the control of the farmer, default increases by 22%.

In sum, the results presented in Table 1.1, imply a robust effect of rainfall shocks on repayment. Loans with shocks are more likely to enter into long past due periods. Now, I turn to the first piece of evidence of the effects of shocks on credit scores. In particular, I look at the effect of shocks on the scores that the BAC reports to credit bureaus and financial authorities.

1.3.2 Effect on Reported Scores

Following regulations from financial authorities, the BAC has to report each month to credit bureaus a score for each client with an outstanding loan with the BAC (hereafter *BAC Reported Score*). This score ranges from A (highest) to E (lowest). Every time a loan is generated, the loan is assigned a score of A. By rule from the financial authorities, a single score has to be reported for each client. When different loans have different scores for the same client, the lowest score is assigned to the client and is the one reported. For small farmers, the BAC reported score depends mainly on days past due and the BAC 's policies (some predetermined thresholds for example).

Table 1.3 presents results of the effect of rainfall shocks on the BAC Reported Score. Here, the sample is identical to the one I use in the baseline exercises of Table 1.2 and the estimated equation is identical to equation 1.1 but I consider two different outcomes. The first one is a dummy for whether or not the BAC Reported Score fell (to any score different than A) at any point during loan tenure. The second one is a dummy for whether or not the score fell to the worst possible score, E. The results imply that rainfall shocks lead to a worse BAC reported score. In particular, a rainfall shock during the first year of loan tenure causes the probability of the BAC reported score to fall or to fall to E in both cases by 20%, compared to loans with no rainfall shock. Therefore, rainfall shocks generate negative information for borrowers that is then reported to credit bureaus.[31] Again, the size of these

[31] In the financial literature terminology, the information a lender shares or reports to a credit

coefficients is important.

The reports from the BAC are one of the channels through which the farmer's credit history feeds into credit bureaus and the financial system in general. A negative credit history will have effects on the credit bureau's credit scores and future access to credit as I document in the next section.

1.3.3 Effect on Credit Bureau Scores and Subsequent Access to Credit

In this section, I study the effect of rainfall shocks in credit bureaus' credit scores and subsequent access to credit. The sample used is the same as the one used in the previous section, that is the sample of Loans With Subsequent Applications.

1.3.3.1 Data and Empirical Approach

I have data for both the CIFIN Consult Stage and the Analysis Stage. Regarding the CIFIN Consult Stage, data is only available for the period 2010-2015. For this period, I observe all the CIFIN consults corresponding to all loan applications received by the BAC. For the Analysis Stage, data is available from 2005 to 2015 for all loan applications that made it to this stage (and therefore, passed the CIFIN Consult Stage).

To analyze the effect of rainfall shocks on subsequent access to credit, I compare loan applications following loan tenures where a rainfall shock occurred during the first year with applications after loan tenures with no shock. More precisely, I estimate by OLS the following model:

$$x_{ijm\tau} = \alpha s_{j\tau} + \theta_{m\tau} + \gamma_j + u_{ijm\tau} \tag{1.3}$$

bureaus is divided in "negative" and "positive" information. Negative information consists mainly of defaults, while positive information includes aspects like outstanding debt. In the case of the reports to the BAC (and in general of all banks in Colombia) both negative and positive information is reported.

where $x_{ijm\tau}$ is one of three outcomes for loan application i, after the maturity of a loan of maturity m, originated in quarter τ, close to rainfall station j. The three outcomes I consider are the CIFIN Score at the CIFIN Consult Stage, a dummy for denial at this stage and a dummy for denial at the Analysis Stage. As before, $s_{j\tau}$ is a dummy equal to 1 if a rainfall shocks occurred in the first year of loan tenure (of the initial loan). $\theta_{m\tau}$ are quarter-of-disbursement times maturity fixed effects of the initial loan. γ_j are rainfall station fixed effects and $u_{ijm\tau}$ is a mean-zero error term.

Similar to the case where I studied the effect of rainfall shocks on repayment, the identifying assumption relies on $s_{j\tau}$ being uncorrelated with the error term. Again, this assumption seems reasonable given my definition of rainfall shocks.

I estimate equation 1.3 for two different samples. Recall that in the construction of the sample of Loans With a Subsequent Application I start with an initial loan and then find the most recent loan application. The selection of this initial loan is irrespective of loan maturity. The two samples I consider differ on the conditions imposed on this initial loan. The first one starts with loans irrespective of maturity, and is the one I have used thus far. The second one starts with initial loans of one year maturity. The reason to consider this second sample is that I can subset it in a way that keeps constant the amount of time elapsed between the rainfall shocks and the next application. I will provide below more details to justify the need of considering this second sample, where I argue that repayment recovers faster than credit access.

1.3.3.2 Results

Table 1.4 presents the results of the effect of rainfall shocks on subsequent access to credit. Panel A shows the results for the sample of applications that follow initial loans of all maturities while panel B shows the results for the sample that starts with one year maturity loans.

In column (1), I report results of estimating the effect of a rainfall shock during loan tenure on the probability of applying for a new loan. That is, I estimate equation 1.3 where the outcome is a dummy for whether or not the farmer applied for a loan after maturity of the original loan. In both panels A and B, I find no significant effects of rainfall shocks on the probability of applying for a new loan. In other words, there is no selection in application caused by rainfall shocks. For exposition purposes, Column (2) presents the effect of rainfall shocks on the first loan. In the case of Panel A, this coefficient is that same as the one reported in Column (1) of Table 1.2.

Now I turn to the results of the effect of shocks during the initial loan tenure on the next loan application. For the sample in panel A (the one with initial loans of all maturities), column (3) shows that the occurrence of a rainfall shock in the first year of loan tenure decreases the CIFIN score by 5.7 points. This corresponds to a decrease of around 0.05 standard deviations of the CIFIN score in the sample. The coefficient in column (4) is positive and statistically significant, indicating that a rainfall shock causes an increase in the probability of denial at the CIFIN stage for the next loan application. The magnitude is large. The size of the coefficient implies that this increase in the probability equals 12.6% of the average frequency of denial of applications following loans with no rainfall shock. Regarding the probability of denial in the Analysis Stage, the effect is also important. Applications that arrive to this stage are 9.8% more likely to be denied if a rainfall shock occurred compared to the denial rate of applications with no shocks during the previous tenure.

Regarding applications that followed loans of one year maturity, the results are very similar (and somewhat larger) to the previous ones, as can be seen in Panel B of Table 1.4. In particular, a rainfall shock during the first year of loan tenure of the initial loan leads to a decrease in the CIFIN score of 7.5 points and an increase in the probability of denial of 17.4% and 11.8% in the CIFIN Consult and the Analysis Stage respectively, compared to the average denial rate of applications following loans with no rainfall shocks.

The previous results show that rainfall shocks have a large effect on credit bureaus' scores and on subsequent access to credit. On average, the occurrence of rainfall shocks during loan tenure is associated with lower credit scores, and more frequent denial of subsequent loan applications. In the following section, I show that this result remains as the time window between the initial loan maturity and the next application grows, which indicates that the effect persists over time.

1.3.3.3 Persistence

In this section, I document the persistence of the effect of rainfall shocks on credit bureau's scores and subsequent access to credit. It is important to note at this point that in 2008 a law was introduced in Colombia that requires that negative information (past defaults) cannot remain on credit bureau registries for more than four years. Therefore, information older than this cannot be reported back to banks or used in the estimation of the credit bureaus' scoring models. This puts a time limit to the effect of negative information on future access to credit.[32]

To study the persistence of the effect, I consider a subset of the samples of the previous section. More precisely, I take the set of first applications after maturity of the original loan but add an additional requirement: that at least a given amount of time elapsed between maturity of the original loan and the first application. I consider a one and a two year time window. For these samples I estimate the same regressions of the previous section.

Table 1.5 presents the results. Columns (1) to (3) report results for the sample that starts with loans irrespective of maturity. Columns (4) to (6) correspond to the sample that starts with loans of one year. Panel A presents the results for a time window of one year. Similar to the results presented in Table 1.4, for both samples, a rainfall shock during the first year of loan tenure decreases the CIFIN score, and increases the probability of denial at both the CIFIN Consult and Analysis Stages. In all cases the coefficients are statistically significant and their size are very close to the ones reported in Table 1.4. These results imply that even after a year has elapsed between the time of the shock and the loan application, loans are still more likely to be denied. In Section 1.4, I will argue that this is enough time for the

[32]This is the same law discussed previously in Footnote 20.

income stream of the farmer to recover and that under some scenarios it is also enough for repayment to recover. Therefore the credit scoring system is leading to denial of loans that on average would be profitable to the lender.

Panel B of Table 1.5 shows the results for a time window of 2 years. The sign of the coefficients is the same as in the case of a time window of one year. For both samples of maturities, a rainfall shock during loan tenure decreases the CIFIN score. Applications after one year maturity loans are more likely to be denied at the CIFIN Consult Stage if a rainfall shock occurred. The rest of the coefficients (columns (2), (3) and (11)) are not significant at conventional statistical levels. Despite this fact, the size of the coefficients is similar, and in some cases larger, than the coefficients in Panel A. Given that the sample size drops by almost half in the regressions of panel B relative to those in Panel A, this is not surprising. In sum, I conclude that even after two years have passed since the maturity of the original loan, subsequent applications are still more likely to be denied. Again, this time window is longer than the one required for the income stream of the farmer to recover from a shock and is long enough, in some plausible scenarios, for repayment to recover.

A small caveat needs to be mentioned at this point. Rainfall shocks might affect the timing of the next loan application. So the comparison between farmers with the same time window, some of which had a shock during the initial loan tenure and some of which did not, might not be the ideal one. Despite this, the fact that the sign and size of the coefficients remains practically unchanged for different time windows provides reassurance in that the effect of shocks is persistent.

The effect of rainfall shocks on credit scores reported by the BAC, credit bureaus' scores, and future access to credit implies a cost to the farmer who is credit constrained. This cost is larger, the longer the time his credit history keeps record of his defaults. In the next section, I show that the farmer's income recovers faster than his access to credit. This implies that the lender is also incurring a cost, as long as future income is the main determinant of future loan repayment. I also present evidence on repayment recovery and argue that under some scenarios, repayment recovers faster than access to credit. The lender could have lent to farmers who experienced a shock, since on average their income stream recovers fast enough to allow for repayment of the next loan.

1.4 Recovery

In this section I argue that the income stream of the farmer recovers faster from rainfall shocks than the credit history and present empirical evidence on repayment recovery. To show that income recovers I proceed in two steps. First, I draw from a large agronomic literature to argue that rainfall shocks affect the productivity of the coffee tree if they occur at most one year before harvest. Shocks occurring before that have no effect. Second, I use data from a representative survey of coffee farmers in Colombia conducted in 2006 to show that the income stream recovers from rainfall shocks faster than credit access, consistent with the agronomic literature. Regarding repayment, I show that it recovers in a sample of long term loans and in a sample of high credit score farmers. I discuss how the results on repayment on long term loans together with the results on persistence of the effect on credit access imply that under plausible scenarios, the bank is not lending to farmers whose repayment recovers in time to repay the next loan.

1.4.1 The Productivity of the Coffee Tree and Rainfall Shocks

Coffee grows in trees and the seed of the cherry is the coffee bean. Usually, farmers harvest the cherry when it is ripe, remove the pulp and dry it to obtain "parchment" coffee. Then they sell it to cooperatives, the FNC or other buyers. The price paid to farmers basically depends on the international price (after conversion to Colombian Pesos at the ongoing exchange rate), plus a premium that recognizes the quality of Colombian coffee. For most coffee farmers, the price is fairly similar across the country.[33]

On average, the coffee tree starts producing coffee one year after its plantation. From year one to five, the productivity of the coffee tree is increasing. Then, from years 5 to 8 it is decreasing. After that age, the coffee tree can still produce coffee but renovation is highly recommended. Different factors affect this pattern including, for example, the quality of the soil (Arcila, Farfán, Moreno, Salazar, and Hincapié, 2007). When the productivity of the tree is decreasing, farmer frequently cut the tree's branches to increase its productivity in subsequent harvests.[34]

During the life time of the coffee tree, and depending on the zone of the country, there are one or two harvest each year.[35] There are two main phases that determine that amount and quality of the coffee cherries of the harvest. The first one is the flowering stage, when coffee flowers blossom and then transform into the coffee cherries. The second one is the coffee fruit development stage, when the coffee berry grows. The flowering phase lasts between 3 to 5 months and the fruit development stage lasts between six and seven months (Gast, Benavides, Sanz, Herrera, Ramírez, Cristancho, and Marín, 2013). Therefore the critical period that determines the quality of the harvest starts 12 to 9 months before.

For a given harvest, the productivity of the coffee tree is highly sensible to the weather events during this period. The agronomic literature has documented that high levels of rain during the flowering phase hinder the development of the coffee flower and are associated with lower levels of cherry production of the coffee tree. Furthermore, lack of solar radiation during the fruit development phase is also associated with lower production (Gast, Benavides, Sanz, Herrera, Ramírez, Cristancho, and Marín, 2013; Arcila, Farfán, Moreno, Salazar, and

[33]Some high quality coffees are paid higher prices.

[34]This is known as renovation by "zoca" or "zoqueo" in Colombian coffee jargon.

[35]The coffee tree is permanently producing coffee. But depending on the region, there are some times in the year where the level of production is concentrated. These are the harvest periods.

Hincapié, 2007; Boucher and Moya, 2015).[36] Therefore, periods of excessive rainfall affect production directly during the flowering phase, and indirectly during the fruit development phase since clouds diminish the amount of sun light the tree receives.[37]

In sum, periods of excessive rainfall affects the productivity of the coffee tree if they occur up to one year before the harvest. The next harvest will not be affected by these events since its flowering and fruit development phases will be under no distress after rainfall has returned to normal.[38]

Although the effect of rainfall shocks on the productivity of the coffee tree is transitory and can last up to one year, the duration of the effect on income can be different. The income stream generated by coffee growing is an economic quantity that depends on farmer's decisions and not only on the productivity of the coffee tree. In the next section, I show empirically that income recovers from excessive rainfall shocks relatively quick.

1.4.2 Income Recovery

1.4.2.1 Data and Empirical Strategy

To study income recovery I use a representative survey of small coffee farmers in Colombia, conducted by the FNC in 2006 and named "[encuesta de] Mercado Laboral Cafetero y Acceso al Crédito para Productores de Café en Colombia"(MLYCC) [Analysis of the labor market and access to credit for small Colombian coffee growers]. The survey contains information

[36] Although not as important as the reasons mentioned in the main text, excessive rainfall can also lead to excessive soil humidity which can affect coffee production, wash away fertilizers, and is associated to higher incidence of diseases in the coffee tree.

[37] Although I have no measures of solar radiation to confirm its correlation with rainfall levels, it has been documented that periods of high rainfall come in hand with lower sunshine in many coffee growing regions of Colombia (Gast, Benavides, Sanz, Herrera, Ramírez, Cristancho, and Marín, 2013; Turbay, Nates, Jaramillo, Vélez, and Ocampo, 2014).

[38] This recovery argument excludes catastrophic events. While the coffee tree is very resistant to weather events there are some that can destroy it; a landslide for example.

on coffee production from coffee farmers with less than 5 hectares cultivated with coffee. Crucially it contains the following question: *"How many 'arrobas' of parchment coffee did you sell the previous year."*[39] Coffee farmers in Colombia usually sell all the coffee they produce, so the answer to this question provides a measure both of production and income.

The survey was conducted between March and April of 2006 and asks about the decisions of the farmer in the previous year (Muñoz-Mora, 2016; Lozano, 2009). Although the question is not specific, I assume from the way it is phrased that it refers to coffee sales in 2005. Depending of the region of Colombia, there are between one and two harvests in a given calendar year.[40] Therefore, I expect shocks occurring in 2004 and 2005 to potentially affect the coffee sales of 2005. Figure 1.2 depicts a time-line where I further clarify this. From the previous section, I expect shocks to have an effect if they occur up to one year before the harvest. For example, and as shown in the figure, harvests in early 2005 are affected by weather events of early 2004. But shocks occurring before January 2004 should not affect 2005's harvest.

I then proceed to link the farms of the survey to the SICA data, allowing me to obtain its geographical coordinates and the closest rainfall station. With these data in hand I estimate by OLS the following regression:

$$r_{t,ijk} = \alpha_0 + \alpha_1 s_{t-1,j} + \alpha_2 s_{t-2,j} + \alpha_3 s_{t-3,j} + Z_i'\rho + \phi_k + u_{ijk} \tag{1.4}$$

where $r_{t,ijk}$ is the amount of coffee sold in 2005 per-hectare cultivated with coffee, by farmer i, close to rainfall station j, in coffee growing region k. $s_{t-1,j}$ is a dummy that takes a value of 1 if a rainfall shock occurred in 2005 or in 2004. Similarly, $s_{t-2,j}$ and $s_{t-3,j}$ are dummies

[39]The 'Arroba' is a traditional Spanish weight measure. The exact weight of an "arroba" varies across countries and regions. In Colombia it is equivalent to 12.5 kilos, or 27.6 pounds.

[40]See Footnote 35. Again, the coffee tree produces coffee beans permanently. In some regions of the country, production concentrates in two different times of the year leading to two different harvests.

for the occurrence of a rainfall shock in 2003-2002 and 2001-2000, respectively.[41] Z_i is a vector of controls that varies depending on the specification. u_{ijk} is a mean zero error term. ϕ_k are coffee-region fixed effects.[42] I expect α_1 to be negative since excessive rainfall shocks occurring in 2005 and 2004 affect the productivity of the coffee tree for the 2005 harvest. Consistent with the hypothesis of recovery from the shocks, I expect the α_2 and α_3 to be non-negative. When estimating equation 1.4, I cluster errors at the rainfall station level and use sampling weights provided in the survey's data.[43]

The identification assumption here is that $s_{t-1,j}$, $s_{t-2,j}$ and $s_{t-3,j}$ are uncorrelated with u_{ijk}, or in other words, that they are as good as randomly assigned. This assumption might seem strong at first. It might be possible that farmers close to rainfall stations with shocks are different in unobservable characteristics from farmers close to rainfall stations with no shocks. I cannot control for this possibility, since the variation I use here is at the rainfall station level, which excludes rainfall station fixed effects. Nevertheless, I am using variation across space in the timing of the shocks (different rainfall stations where shocked at different points in time). Again, given my definition of shocks, this variation is likely to be as good as random even in the cross section.

Figure 1.3 shows the distribution across Colombia of rainfall stations for which at least two quarter-shocks occurred in 2005 or 2004 (that is $s_{t-1,j} = 1$). Dots in green depict rainfall stations with $s_{t-1,j} = 0$ and that where close to at least one farm in the MLYCC survey (where

---

[41] To define rainfall shocks in this section I use the 90th percentile of the historical rainfall distribution instead of the 80th percentile. As discussed before, the period under analysis in the previous sections includes La Niña phenomenon, so shocks defined with the 80th percentile are strong enough to alter repayment. In the current exercise, I need to consider the 90th percentile to capture an effect on coffee production. Arguably, this is due to the fact that no extreme weather events like la Niña occurred in the period under consideration.

[42] I follow Muñoz-Mora (2016) and define four coffee regions that differ in natural conditions, like altitude, that affect coffee production.

[43] The ideal data to identify the recovery from rainfall shocks would be a long enough panel at the farm level with data on production or sales. Unfortunately such a panel does not exist for coffee production in Colombia.

close is at a distance of 6.1 km or less). Dots in purple depict rainfall stations with $s_{t-1,j} = 1$. The takeaway from the figure is that rainfall stations with $s_{t-1,j} = 1$ are distributed across the country and not concentrated in a particular region. This is consistent with the idea that shocks are as good as random in the cross section, and affect farmers in different regions of the country. In the next section, where I present the results, I show additional exercises that provide support for this claim.[44]

1.4.2.2 Results

Table 1.6 reports the results from the estimation of equation 1.4 under different specifications. The first column shows the results without the inclusion of controls. The estimated value of α_1 is presented in row "Rainfall Shock 04-05". It is negative and statistically significant. Its size implies that the occurrence of rainfall shock in 2004-2005 is associated with a reduction in the amount of coffee sold (per-hectare cultivated with coffee) equivalent to 32% of the sample mean or to 0.31 standard deviations. The median farm in the estimation sample has 67% of its area cultivated with coffee so coffee sales are likely to be the main income source of most farmers in the survey. Therefore, the reduction in sales caused by excessive rainfall shocks have an important impact on farmers' income. Regarding excessive rainfall shocks in the periods of 2002-2003 and 2002-2001 the estimated values of α_2 and α_3 are positive and non-significant at conventional statistical levels. At the bottom of the table, I report the p-value of tests where the null hypothesis is $\alpha_1 = \alpha_2$ and $\alpha_1 = \alpha_3$. I reject equality for the first test at a level of significance of 10%. For the second test, I cannot reject equality at conventional levels of confidence. Nevertheless, the p-value is low (equal to 0.12).

These results are consistent with recovery of farmer's income: excessive rainfall shocks

[44]The estimations presented here use farms at a distance of 6.15 km or less to the closest rainfall station. My estimations for the effects on repayment are irrespective of distance. Nevertheless, Boucher and Moya (2015) only find effects of rainfall on forecasts of coffee tree productivity for distances smaller than to 3 km, although they do not consider rainfall shocks but rainfall levels instead. To be conservative, I use a 6.15 km distance (which is the median distance in my sample) in the baseline specification. The results presented here are robust to considering different distances and total sales (instead of total sales per-hectare) as outcome.

during 2005-2004 significantly reduce the total amount (or weight to be more precise) of coffee sold, while shocks during the periods of 2003-2002 and 2001-2000 have no negative effect. Note that the timing of this effect matches with what is expected given the timing of the effect of rainfall shocks on the coffee tree's productivity. Recall that only shocks occurring up to one year before the harvest affect the productivity. So, as I discussed before, shocks up to 2004 should have an effect on 2005's harvest. The estimated effects come from different rainfall stations distributed across the country. For example, there were a total of 128 farms that had a rainfall shock in 2004-2005, corresponding to 26 different rainfall stations out of 246 that are close to farmers in the survey.

The second column of Table 1.6 includes a series of controls that are arguably predetermined at the start of the period under consideration (i.e. 2000-2005): the size of the farm, the size of the household and the level of education of the household head. The inclusion of this set of controls does not considerably change the estimate of α_1. This is reassuring in the sense that if rainfall shocks are as good as randomly assigned, the estimated coefficient should not be affected by the inclusion of predetermined controls. The same is true for the point estimates of α_2, and α_3, which do not change considerably with the inclusion of these controls.

Finally, in the last column I present the results of estimating equation 1.4 but with controls that are not predetermined at the time of rainfall shocks in the periods 2000-2001, and 2002-2003. In particular, I include the value reported in the survey of the density of the coffee crop (number of trees per hectare cultivated with coffee), the average age of the coffee crop, dummies for the variety of seed used, and a dummy for whether or not the crop is completely exposed to sunlight.[45] All these variables have been documented to affect the productivity of coffee crops (Muñoz-Mora, 2010; Gast, Benavides, Sanz, Herrera, Ramírez, Cristancho, and Marín, 2013). Since these variables correspond to coffee farmer's decisions, they are likely to depend on rainfall shocks during the periods of 2000-2001 and 2002-2003

[45]It is frequent for coffee crops to be intertwined with other trees that shadow coffee trees.

and therefore might bias the coefficients of interest (Angrist and Pischke, 2008). The results reported under column (3) show that even with the inclusion of "bad controls" the estimate of α_1 does not change much. This is not true though for the values of α_2 and α_3 which change considerably compared to the values under columns (1) and (2), consistent with the idea that the additional controls depend on these variables. Still, the estimated coefficient is positive and not significant suggesting that these shocks do not affect the total amount of coffee sold per-hectare in 2005.

The fact that the estimate of α_1 does not vary much across specifications is consistent with the idea that rainfall shocks are as good as random in the cross-section. As discussed previously, Figure 1.3 shows that treaded rainfall stations are distributed across the country and not concentrated in a particular region. In Table 2.2, I present a balance test where treated farms are those with a rainfall shock in the period 2004-2005. Large differences in observables across treated and control groups are not observed. The results of Figure 1.3 and Table 2.2 are consistent with the idea that rainfall shocks are as good as randomly assigned in the cross section.

To conclude, the agronomic coffee literature has documented that excessive rainfall shocks affect the productivity of the coffee if they occur within one year of the harvest. Once weather returns to normal, so does productivity. The results from the exercises with the survey data show that income follows a pattern that is consistent with the evolution of productivity. It is affected by shocks close to the current harvest but not by older shocks.

1.4.3 Repayment Recovery

A question that arises at this point is whether income recovery translates into a recovery in repayment. This is likely to be the case since one of the main determinants (if not the main determinant) of repayment is the income stream of the farmer.[46] Consistent with this idea is

[46]Empirical evidence on this matter is scant but still suggestive of such relation. For example, Chirwa (1997) finds in a cross section of loans to small farmers that crop sales are associated with higher

the fact that the BAC's repayment schedule is organized to match the harvests of its farmers. Nevertheless, it is still possible that after income recovers from a rainfall shock, the farmer does not repay his loan.

I face a fundamental problem in answering the question of whether or not a farmer would repay a loan after defaulting in a previous one, as a result of a rainfall shock. As documented previously, rainfall shocks cause on average higher rates of denial of posterior loan applications so the sample of farmers who get their next loan approved is selected. Furthermore, I do not observe repayment of farmer's who are denied a loan. To be more precise, this is a quantity that does not exist. So, without making strong assumptions, answering this question in the sample I have used thus far is hard.[47] Nonetheless, in this section I show that repayment recovers for two different subsets of loans: loans with long maturities and loans from farmers with high ex-ante credit scores. I consider that these results can be extrapolated to other samples, given the results presented previously on income recovery. Furthermore, I explain in detail how the timing of the recovery in long term loans and the persistent effect on access to credit documented in Section 1.3.3.3, imply that the bank is not lending to borrowers who could repay a second loan.

1.4.3.1 Recovery of Long Term Loans

In this section, I present evidence on repayment recovery in a sample of loans of maturities of five years or more.[48] More precisely, I estimate by OLS the following equation, for k in {1,2,3,4,5}:

---

loan repayment in Malawi. Acquah and Addo (2011) find that higher fishing income is associated with higher repayment of loans to fishers.

[47]One could postulate a Tobit selection model and estimate a first stage of loan acceptance. Nevertheless, this approach needs strong distributional assumptions. Angrist and Pischke (2008) warn against estimating Tobit models where the latent variable has no empirical counterpart.

[48]Loans of such a long maturity are used in coffee production to finance investments such as planting new coffee trees, or build warehouses to dry coffee, known as "beneficiaderos" in the Colombian coffee jargon.

$$y_{k,ij\tau} = \beta_k s_{j\tau} + \psi_\tau + \iota_j + \nu_{kij\tau} \tag{1.5}$$

This equation is identical to equation 1.1 but the outcome is a dummy equal to 1 if loan i ever entered in a period of 30 days past due at age k (where k is in given in years). As before, i indexes loans, τ indexes quarter of origination, and j rainfall stations. $s_{j\tau}$ is a dummy for rainfall shocks in the first year after the loan was disbursed. ψ_τ are quarter-of-disbursement fixed effects, δ_j are rainfall station fixed effects and $\nu_{kij\tau}$ is a mean-zero error term. I cluster errors at the rainfall station level.

In Figure 1.4, I plot the estimate of β_k for k in $\{1,2,3,4,5\}$ along with bars indicating 10% confidence intervals. I do this exercise for two different samples. Recall that I have monthly data on repayment for each loan. It is possible that I don't observe the loans until their maturity. This can either be because the loan was repaid in full before maturity or because the loan was restructured.[49] In the left panel of Figure 1.4, I don't take into account the fact that loans can be restructured and assign a 0 to loans even if I don't observe them in a given year.[50] The right panel removes from the sample loans that were restructured.[51]

The left panel of Figure 1.4 shows that the occurrence of a rainfall shock in the first year of tenure has a statistically and positive effect on the probability of entering in a period with past dues in years 2 and 3. The effect then fades out and gets close to zero for year 5. The small effect in year one can be explained by the fact that most long term maturity

[49] Loan restructuring occurs when the bank agrees with the client to start a new obligation that gathers previous obligations. The new obligation normally has different conditions than the previous one (for example a different payment schedule). This is usually done to offer alternatives of repayment to the client when he has entered overdue periods or manifested that he cannot pay.

[50] For example, consider loan A, that had zero overdues in year 1, zero overdues in year 2, a period of 30 days past due in year 3, but that was repaid in full in year 4. Then I code $y_{1,A} = 0$, $y_{2,A} = 0$, $y_{3,A} = 1$, $y_{4,A} = 0$, $y_{5,A} = 0$.

[51] I use a list provided by the BAC of restructured loans. This list is not complete though. I constructed a list of restructured loans by taking loans that end before maturity and establishing if there was a loan issued one month before, the same month or one month after for the same farmer. This list is very similar to the list provided by the BAC.

loans have a grace period (of usually one year) where farmers are not required make any capital payments. The pattern depicted in the figure suggests that repayment recovers from the shock in the first year of loan tenure. The right panel of Figure 1.4, where I remove from the estimating sample loans that where restructured, is similar to that of the left panel. Although it is less pronounced, it also shows that the effect of the shock is monotonically decreasing with the time elapsed since the shock. More importantly, in both panels there is no significant effect of the shock in year 4 or in year 5. In sum, both figures suggest that in the sample of loans of five years or more, the effect of the shock on repayment dies out in time.

1.4.3.2 The Timing of Repayment Recovery and the Persistence in Exclusion from Credit Access

Does the timing of the recovery described in the previous section imply that the bank should lend to borrowers not currently receiving loans? In short, the answer is yes. Recall from Section 1.3.3.3 that the effect of shocks on future access to credit is persistent. In particular, in that section I documented that the effect of a shock on subsequent access can last two years. That amount of time, combined with the repayment recovery presented in Figure 1.4 implies that the bank is not lending to at least some important fraction of farmers whose repayment will have recovered at the time of the first payment of a potential second loan. In Figure 1.5, I present a time schematic where I consider four plausible scenarios to illustrate this point. The idea is that the time it takes for the farmer to apply for a new loan together with the time of the bank's grace periods are enough for the farmer to recover.[52]

In this section, I work with the results on persistence obtained with the sample that starts with one year maturity loans (columns (4) to (6) of Table 1.5). The advantage of using this sample is that for a fixed time window between loan maturity and the next application, the

[52]Grace periods are usually one of the components of the loans granted by the bank. For example, for one year maturity loans, the farmer can repay the full amount he owes one year after the loan is disbursed.

time elapsed between a shock and the next application remains constant. This is not the case for the sample that starts with loans of all maturities. In that sample, the time between the first year of the initial loan and the next application varies with maturity (even if the time window between loan maturity and the next application is constant).

I consider two time windows between maturity of the first loan and the subsequent application, corresponding to those studied in Section 1.3.3.3 (Table 1.5, and in particular Column (5)). For the sample starting with one year maturity loans a shock during tenure increases the probability of denial of the subsequent loan application (at the CIFIN stage). This increase is of 22% for a time window of two years between the maturity of the first loan and the next application and of 20% for a time window of one year.

I consider four different scenarios depending on the time window and on whether or not I allow for a grace period on the second loan. To be conservative, I consider only one year grace periods. This yields four scenarios to study. Scenario 1: a two years time window and no grace period. Scenario 2: a two years time window with a grace period. Scenario 3: a one year time window and a grace period. Scenario 4: a one year time window and no grace period.

Consider Scenario 1 depicted in Figure 1.5 (top-left panel). In this Scenario, a shock occurred during tenure of the initial loan, two years elapsed until the next application and there was no grace period. In this case, the probability of denial of the next loan is 22% higher as a consequence of the occurrence of the shock. If the recovery in repayment is the same as the one depicted in Figure 1.4 (Panel B), then repayment of a second loan will start at a point in time where repayment will have already recovered from the shock. As it can be seen in the schematic, repayment recovers in years 4 and 5 and repayment of the second loan would start in year 4. Nevertheless, the applications following tenures with a shock are 22% more likely to be denied. This implies that the bank is denying loans that in principle would be repaid.

In Scenario 2 (top-right panel of Figure 1.5) I show a situation where two years elapse between maturity of the first loan an there is a grace period for the second loan. In this case, repayment of a potential second loan would start in year 5, a point in time when repayment has recovered for more that a year. The increase in the denial rate is the same as before: 22% higher for applications after a loan with a shock, even though repayment would have recovered at the time of the first payment.

In Scenario 3 and 4 (bottom-left and bottom-right panels of Figure 1.5, respectively), I consider a one year time window between maturity of the first loan and the subsequent application. Again, for Scenario 3, the grace period guarantees that repayment will have recovered at the time of the first payment of the second loan (year 4). Despite this fact, loan denial is 20% higher for applications with a shock during the previous loan tenure. Scenario 4 is the only one from those considered in this analysis where repayment of a second loans starts at year 3, a point in time where repayment has not recovered yet.

The three scenarios where repayment recovers in time for the first payment of the second loan are plausible scenarios. For example, the entire sample that starts with one year maturity loans in Table 1.5 consists of 20,549 applications. Of these, 22% (4,558) correspond to a time window of two years between maturity and the next application. As the two top panels of Figure 1.5 show, even with no grace period, for all these loans repayment will have recovered at the time of the first payment of a potential second loan. Nevertheless, denial is 22% more likely for applications following a loan with a shock. Scenario 3 is also frequent, given that all one year maturity loans have a one year grace period.

This analysis assumes though that the repayment recovery of Figure 1.5 applies to all loans (and not only long term maturity loans). But given the results on income recovery obtained with the production data (Section 1.4.2) and according to which only shocks not older than one year affect production and sales, this seems a reasonable assumption. Furthermore and as a general rule, a farmer has many plots with different characteristics within the same farm. Plots in a single coffee farm can differ in many ways, for example in the age of the coffee trees planted, the coffee variety, the density (trees per hectare) and the pattern in which they are planted. All these variables have implications in terms of the productivity of the coffee tree. Usually, farmers have plots with different characteristic to smooth production (a practice that is also recommended by the FNC). Long terms loans are most frequently used for the plantation of new coffee trees, most probably in a single plot out of the many in the farm.[53] Although the plots with the new coffee trees are not producing any coffee, the farmer can pay the loan with the income generated by the other plots. Therefore, the repayment of the long term loan reflects the repayment ability generated from the production of the whole farm (and all of its plots) and not only the plot where new trees were planted using the long term loan.

The loses for the bank of not lending to the set of borrowers who are denied a loan because of rainfall shocks are large. Under some assumptions, they amount in my sample to 7,146 millions of Colombian pesos in revenue (about 3,4 millions of US$), and about 6,066 millions of Colombian pesos in utilities (about 2.9 millions of US$). These are approximately 2.2% and 5.6% of the bank's revenues and utilities in the first quarter of 2013 respectively.[54] The

[53]For example, in the SICA data in 2011 there are about 705,000 farms observed. The average number of plots across these farms is 4.74. Averaging across farms the mean age of their plots yields 10.4 years and averaging the standard deviation of the age of their plots yields 4.5 years. This implies that within farms there is considerable variation in the age of the coffee trees across plots. For the smoothing reasons mentioned above, it is very unlikely that a farmer will renew all the trees of his farm at a single point in time.

[54]In my full sample, I observe around 274,000 loans disbursed in the period of 2005-2011 and with an average maturity of 3.4 years (these are the loans in Table A.1 in the Appendix). Suppose that the proportion of farmers who apply for a second loan is the same as the one in the estimation sample (64%). This implies 174,300 new applications under a re-application rate of 64%. For this sample, 39% of the loans had a shock in the first year. So there were around 70,000 applications for which a shock in the previous loan occurred. Now, the average time window between the end of the first loan and the second loan is 1.3 years (taken from the sample in Panel A of Table 4). From column (2) of Panel A in Table 5, the CIFIN denial rate is 0.017 points larger for applications - occurring more than a year after the maturity of the first loan - that had a shock in first the loan tenure. Therefore in the CIFIN stage, out of the 70,000 applications, 1,190 applications ($= 0.017 \times 70,000$) are denied because of the shock. Of the 70,000 applications (and assuming that the rate of loan denial in the CIFIN stage is that of the control group in the same column, and equal to 0.144), around 58,200 applications following a loan tenure with a shock make it to the analysis stage. The rate of denial in this case is 0.027 points higher for applications following a tenure with a shock. Therefore, in the analysis stage, about 1,570 ($= 0.027 \times 58,200$) additional applications are denied. This gives a total of 2,760 loans denied because of the rainfall shocks. Now, during this period the average yearly interest rate for loans to small farmers was about 14.4 percent, and the average size of the loan in my sample of coffee farmers was about 10,415 thousands of Colombian pesos. If I assume that all loans denied because of the shock would have been repaid, a simple amortization schedule for a maturity of 3.4 years with an annual interest rate of 14.4 percent, implies that the bank is losing 7,146 millions of Colombian

losses of the BAC are not restricted to the ones calculated in this sample. In August of 2014, the bank had 860,298 loans for small farmers outstanding. If the mechanisms documented in this paper apply for agricultural activities beyond coffee, it is likely that the losses of not incorporating information on exogenous shocks in credit histories and credit scores are larger. Furthermore, these calculations do not include the welfare looses coming the revenues lost by the farmers who are denied a loan they could repay.

1.4.3.3 Recovery of High Credit Score Farmers

In this section I focus on loans from farmers with high credit scores at the time of loan origination. The advantage of considering this sample is that even if these farmers get a rainfall shock during loan tenure, it is less likely that they will be denied a subsequent loan. Therefore, selection is less of a concern in this scenario and it is more likely that I can observe repayment of a second loan. I study in this sample whether a shock during the first loan tenure has an effect on repayment of the next loan.

Since I only have data on the CIFIN Stage starting in 2010 and data on rainfall shocks up to 2012, I focus on the sample of loans originated in the period 2010-2011. For these set of loans I observe the CIFIN credit score at the time of loan origination. I restrict the sample to loans of farmers with a CIFIN score in the top quartile of the original sample. As in section 1.3.3.1, I construct a sample that starts with an initial loan and then look for the subsequent loan application.

I estimate the effect of a rainfall shock on different outcomes. Table 1.8 present the results. Column (1) presents results where the outcome is dummy for whether or not the initial loan was overdue for 30 days or more.[55] In both panels, I find a significant effect of a rainfall shock on repayment of the initial loan. This result implies that rainfall shocks

pesos in interest revenues (about 3,4 millions of US$), and -subtracting the costs of operation for this amount of loans – about 6,066 millions of Colombian pesos in utilities (about 2.9 millions of US$).

[55]The estimated equation is identical to equation 1.1.

affect repayment behavior even for ex-ante high credit score farmers. Columns (2) and (3) present the results of the effect of a rainfall shock on a dummy for denial at the CIFIN Stage and the Analysis Stage respectively, for loan applications following the initial loan tenure.[56] I find no significant effect on the probability of denial, which confirms that selection is less important in the sample of high credit score loans. Finally, Column (4) presents results where the outcome is repayment of a second loan. In this case, the sample consists of subsequent loans (i.e. that were approved) after tenure of the initial loan. Again, I find no significant effect of shocks in the first loan on repayment of the second loan.

In sum, for high credit score farmers, I find that shocks in the first loan do not have a significant effect on repayment of a subsequent loan. This implies that repayment recovers from rainfall shocks for this sample of high credit risks, consistent with the findings on income recovery.

1.4.3.4 Frequency of Default, Borrowers Quality and Exogenous Shocks

It is interesting to note in Table 1.8 that the average rate of default of initial loans with no rainfall shock is 1.3%. A rainfall shock increases this default rate to 108%. These magnitudes are different than my main estimates from Table 1.2. There, the average default rate is 15.5% and the rainfall shock causes an increase in the default probability of 22%.

Although these differences seem large at first they correspond to what is expected. First, high ex-ante credit score borrowers (those in the sample of Table 1.8) have a much lower rate of default that the rest of the population (for example, borrowers in the sample of Table 1.2). More interestingly, exogenous shocks increase by much more the rate of default (relative to the mean) in the case of high quality borrowers. The average borrower defaults for many different reasons, one of which are exogenous shocks. For high quality borrowers, there are very few causes of default, therefore exogenous shocks carry much more importance in causing a default.

[56] In this case, the estimated equation is identical to equation 1.3.

1.5 Discussion and Concluding Comments

In this paper, I documented a market failure that results from the use of traditional credit scoring and credit reports in agricultural lending. In particular, I have shown that excessive rainfall shocks affect farmers' repayment, credit scores, and future access to credit. Furthermore, I showed that income and repayment recover faster than credit access. These results imply costs both to the farmer and to the lender. Incorporating information on rainfall shocks in credit scores would likely increase efficiency and welfare in this credit market. A large recent literature has documented the benefits of using weather information to create index based insurance.[57] In a similar fashion, this paper shows that precise information on weather events can be used to obtain better credit scoring systems and to allocate credit more appropriately in developing countries.

A concern worth discussing is that my study considers a single bank in a single developing country. However, at least five countries in Latin America have banks or micro-finance institutions lending to farmers and consulting their credit histories in credit bureaus. There is additional evidence of micro-credit institutions seeking information from credit bureaus or sharing information of their clients as a discipline device.[58] Finally, since the late '90s there has been a return of Public Development Banks. A survey of 90 of these banks in 61 countries conducted in 2011 by the World Bank revealed that 92 percent target small and medium enterprises, and 83 percent target agricultural business (de Luna-Martínez and Vicente, 2012). Additionally, there is a consensus among policy makers that Public Development Banks should be regulated as private banks (de Olloqui, 2013). This usually implies maintaining and reporting credit histories of bank clients to credit bureaus and financial authorities. Therefore, it is likely that the mechanisms documented in this paper are relevant for other Public Development Banks. Also, as mentioned before, the mechanism outlined here are

[57] See for example Karlan, Osei, Osei-Akoto, and Udry (2014).

[58] See for example de Janvry, McIntosh, and Sadoulet (2010) and Giné, Goldberg, and Yang (2012)

likely to apply to credit scores in other credit markets, although in those scenarios, orthogonal shocks might be less important determinants of repayment or information on them might be hard to obtain, in different contexts.

If the mechanisms that I document in this paper exist, why haven't lending institutions taken them into account? The most plausible hypothesis is that until recently it was very costly or simply not feasible to use precise measures of weather in the computation of credit scores. It is no mystery to institutions lending to small farmers that weather affects lender's repayment and profitability. Nevertheless the exercises performed in this paper require levels of measurement and precision that historically could not be implemented because of cost or technological constraints.[59] For example, geo-referencing technologies became available at low costs only recently. In Colombia, most farmers are still not geo-referenced. Also, dense networks of weather stations with long time series of weather information remain uncommon.[60] A second possibility is that the fact that credit scoring systems are well suited for consumer lending in developed countries might lead one to conclude that they can be applied with no strings attached to agricultural lending in developing countries. In other words, there might be inertia in banking practices across banks. Furthermore, the fact that lending institutions are ignoring information on the sources of credit downgrades is not unique to my setting.[61]

[59]The BAC is aware of a theoretical relationship between weather and repayment and it invest resources in monitoring weather at an aggregate level, for example, the municipality level. But according to BAC officers, previous exercises at this level of aggregation show no effect on repayment behavior. The study by Castro and Garcia (2014) use BAC data to estimate in a structural risk model the effect of weather at an aggregate level. These levels of aggregation though cannot inform individual credit histories and credit scores.

[60]One of the advances of this paper consists in finding a scenario where such a network is available. In particular geo-referenced coffee farms combined with data from the IDEAM. It is important to note that at the time of this study the IDEAM information was not available for the bank to be used at the level of disaggregation of this paper.

[61]Garmaise and Natividad (2016), discussed in the introduction, document a similar situation in consumer lending in Peru. In their case, lending institutions have all the information needed to understand the source of an exogenous score downgrade but still do not use it. This echoes the fact that the BAC does not use such information on weather shocks.

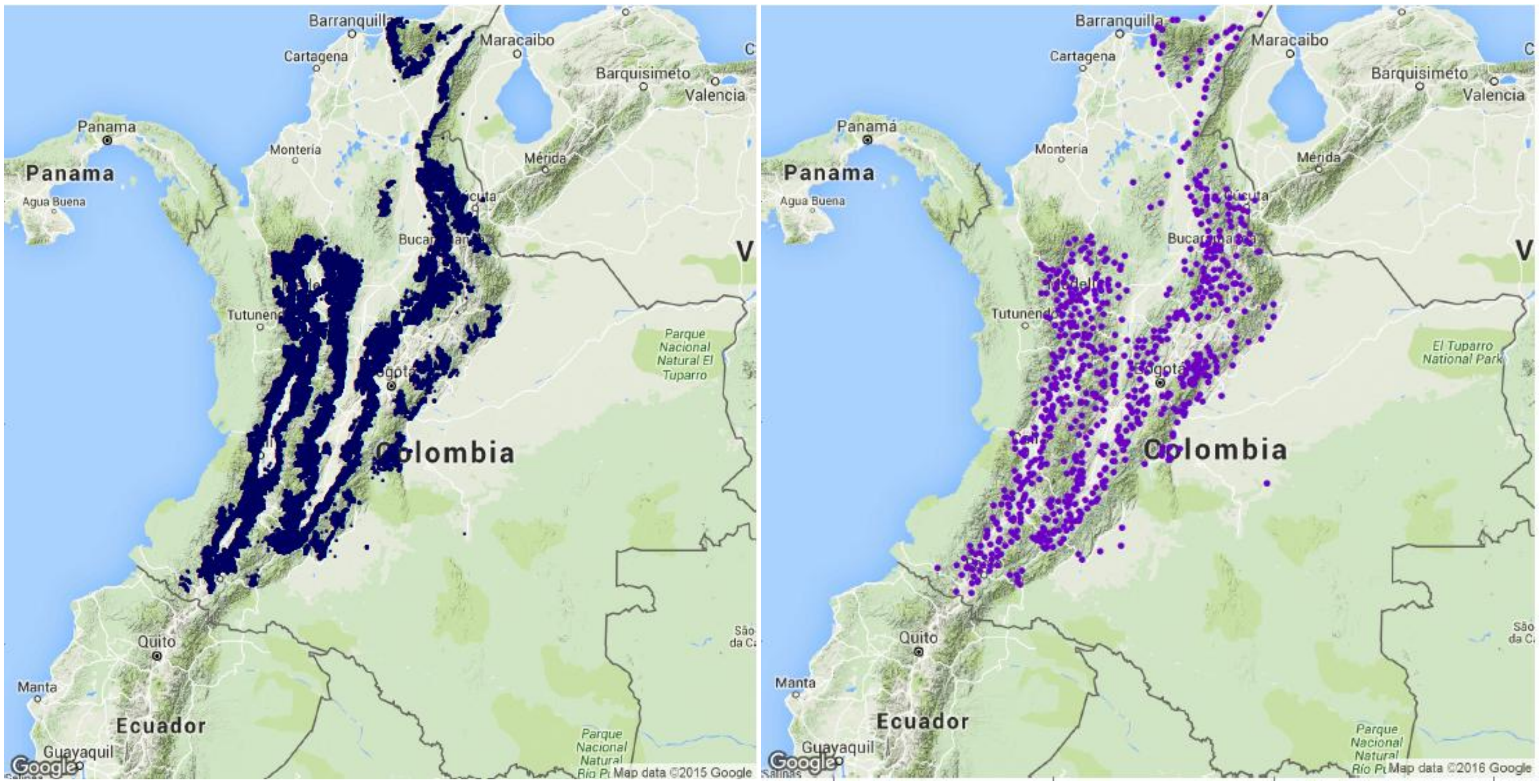
In the absence of complete insurance markets, a direct policy implication emerges from the results of this paper. Precise weather information should be used when computing credit scores. This entails geo-referencing farmers and establishing systems that accurately measure weather events near them. Once these systems are in place, detailed records should be kept. These can be included in credit reports and used when computing credit scores. Furthermore, the findings that I presented constitute an example of a situation where a policy, practice or institution used in developed countries can and should be adjusted for the particular setting of developing countries.[62]

The results of this paper open different avenues for future research. First, the welfare implications of omitting exogenous shocks from credit scores can be estimated in a structural model of lender and borrower decisions. Second, there may be other credit markets where differentiating between causes of default and credit downgrades could improve credit allocation. Third, it opens the question of how to better measure and incorporate information on exogenous shocks for credit histories and credit scores.

[62]There is a recent literature in the intersection of public finance and development that studies how the policies that are optimal in developed countries might not be optimal for developing countries. See for example Gordon and Li (2009) and Best, Brockmeyer, Kleven, Spinnewijn, and Waseem (2015). My scenario is different in that the inefficiency this paper documents from the use of traditional credit scoring likely exists in developed countries as well. The point is that for agricultural lending in developing countries we can observe the exogenous shocks that generate the inefficiency.

1.6 Figures Chapter 1

Figure 1.1: Coffee Farmers and Rainfall Stations Distribution Across Space



A. SICA farms 2010-2013

B. Rainfall Stations (close to a coffee farm)

Figure 1.2: Harvest Time Line

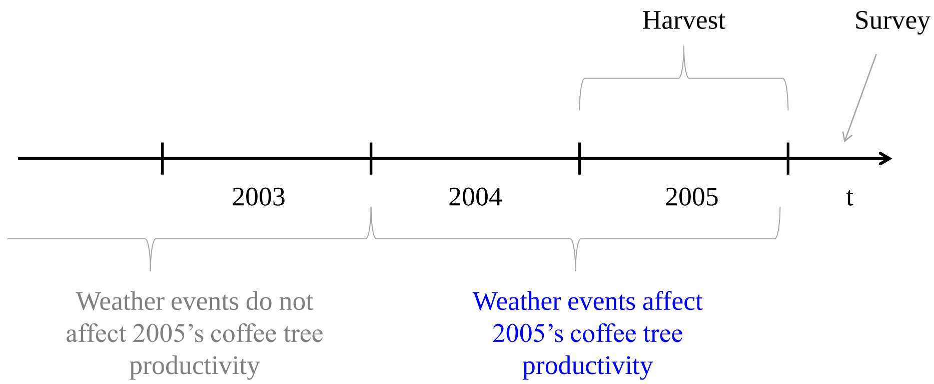


Figure 1.3: Treated Rainfall Stations

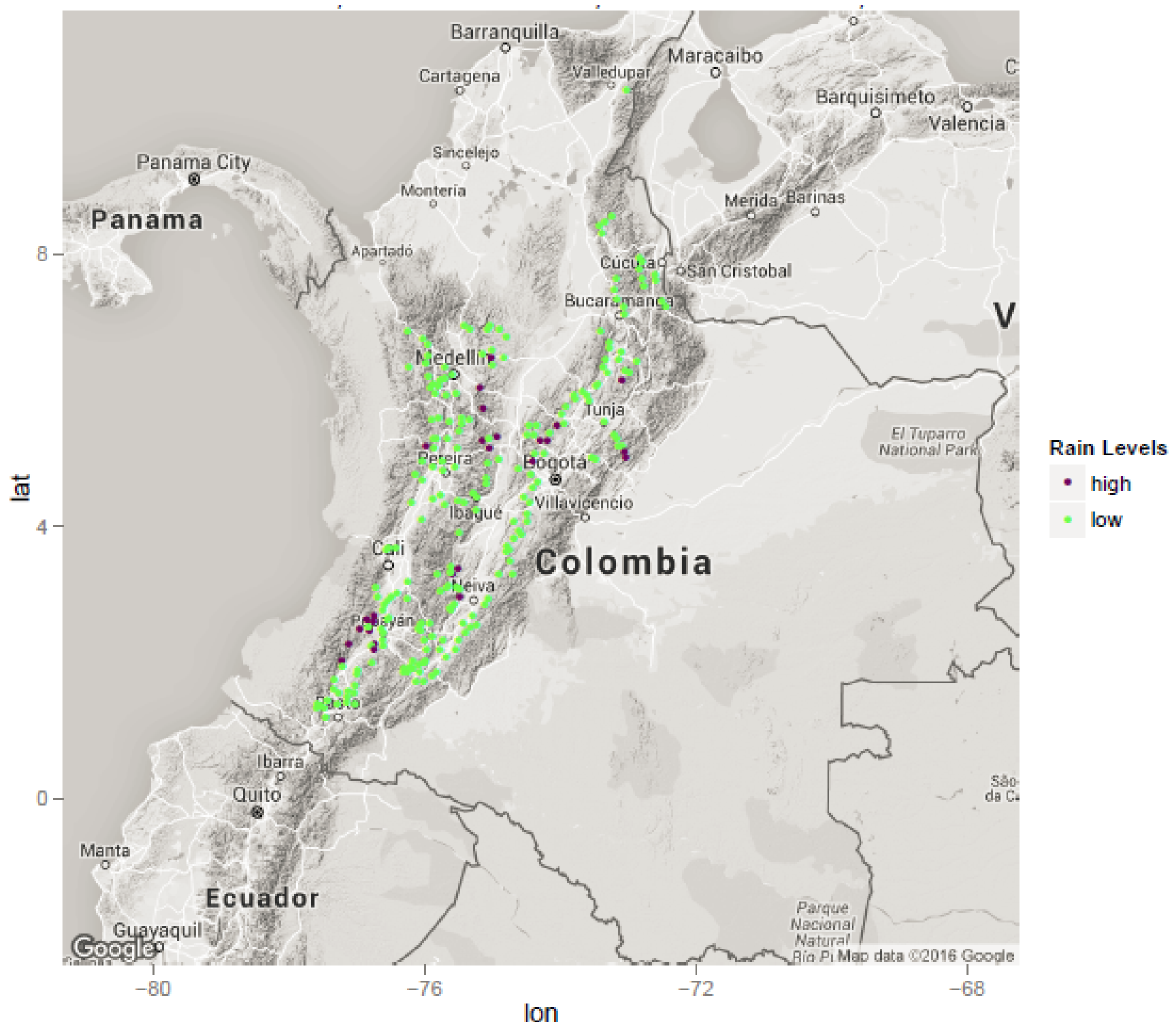
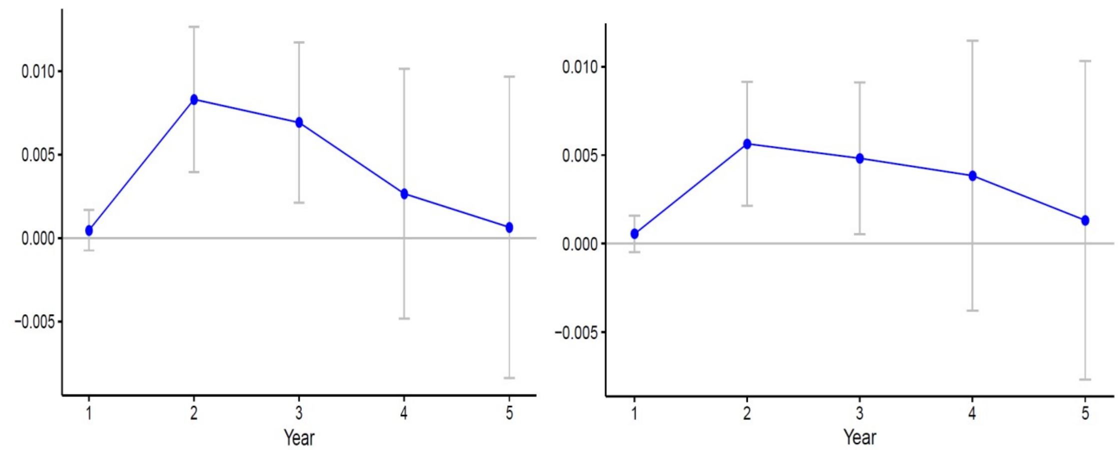


Figure 1.4: Repayment Recovery in Loans of Five Year or More Maturities

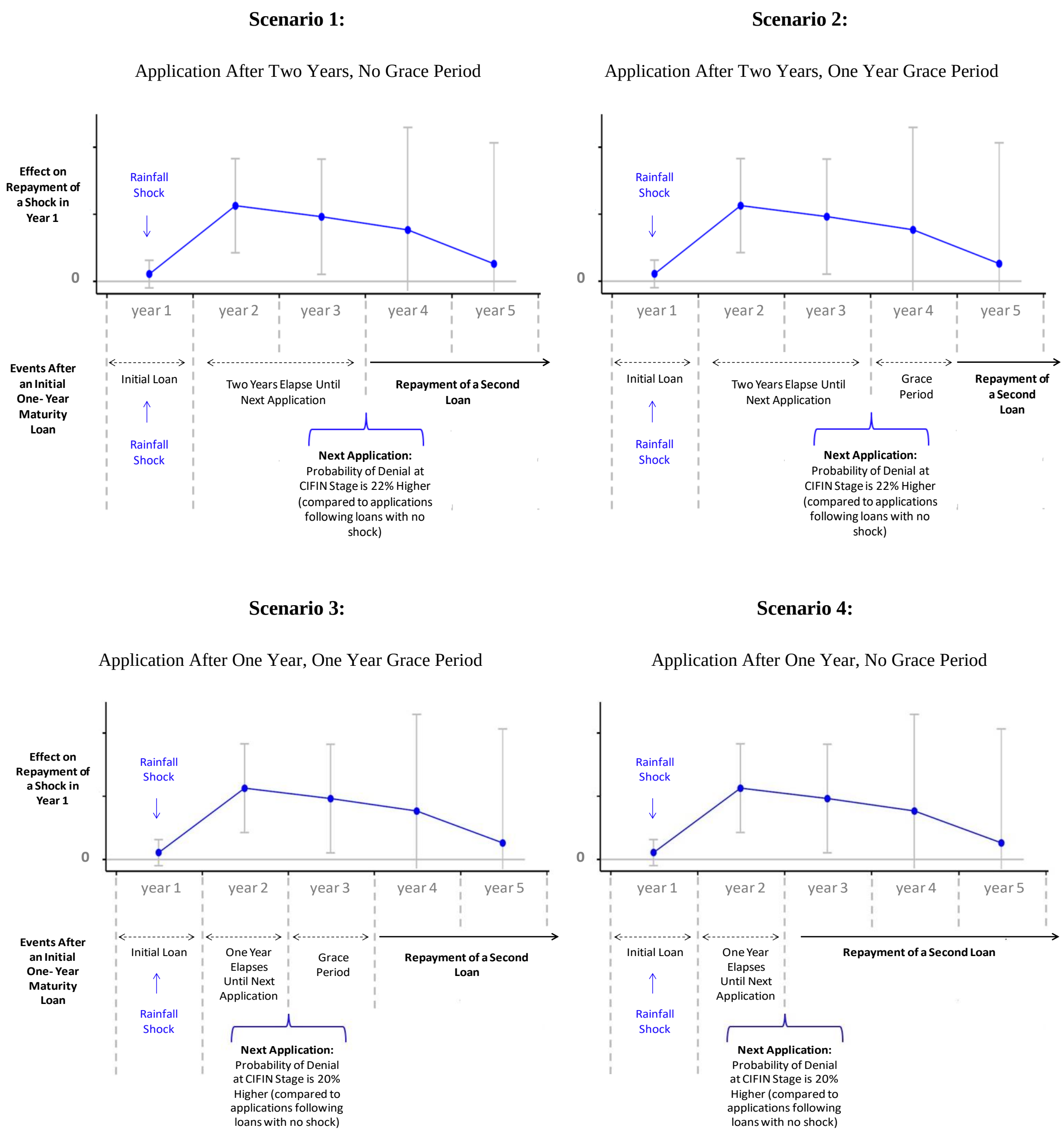


A. All Loans

B. Restructured Loans Excluded

Notes: Each panel plots coefficients for the estimated effect of a shock in the first year of loan tenure on a dummy equal to one if the loan entered in a period of 30 days past due, at any given age of the loan. The age of the loan is represented in years in the horizontal axis. Vertical bars correspond to 10% confidence intervals. The sample consists of loans of maturities of five or more years disbursed in the period 2008-2011.

Figure 1.5: Recovery Schematic



1.7 Tables Chapter 1

Table 1.1: Descriptive Statistics

	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Maturity (years)	1.65	1.40	0.17	1.00	1.00	2.00	6.58
Interest Rate (annual)	10.67	2.72	2.47	9.50	10.02	12.54	43.98
Distance to Rainfall S.	6.54	3.98	0.04	3.86	5.96	8.37	38.13
Rainfall Shock, year 1	0.42	0.49	0	0	0	1	1
# quarter-shocks, year 1	1.35	1.06	0	1	1	2	4
30 Days Overdue	0.14	0.35	0	0	0	0	1
60 Days Overdue	0.10	0.30	0	0	0	0	1
BAC Score Fell	0.14	0.35	0	0	0	0	1
BAC Score Fell to E	0.07	0.25	0	0	0	0	1

Notes: The data source is the BAC administrative data. Included loans are for coffee production, originated in the period of 2008-2011 and for which there is a subsequent application observed in the CIFIN Stage.

Table 1.2: Effect of Rainfall Shocks on Repayment

	Baseline			Heterogeneous Effects		
	30 Days Overdue		60 Days Overdue	Dist.to Stat. < Median	Maturity: Short	Maturity: Long
	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall Shock, year 1	0.034***		0.022***	0.017*	0.032***	0.041**
	(0.007)		(0.006)	(0.009)	(0.008)	(0.02)
# quarter-shocks, year 1		0.016***				
		(0.004)				
Mean (control group)	0.155		0.117	0.11	0.12	0.103
Mean (all obs.)		0.142				
Origin Date * Matu. FE	Y	Y	Y	Y	N	N
Origin Date FE	N	N	N	N	Y	Y
Rainfall St. FE	Y	Y	Y	Y	Y	Y
Rainfall St. Clst.	Y	Y	Y	Y	Y	Y
Observations	32,512	32,512	32,512	16,590	28,047	4465
Adjusted R^2	0.125	0.124	0.114	0.121	0.138	0.061

Notes: The data source is the BAC administrative data. Included loans are for coffee production, originated in the period of 2008-2011 and for which there is a subsequent application observed in the CIFIN Stage. In a given calendar year, a rainfall shock is defined by rainfall realizations in at least two quarters above the 80th percentile of the quarter-year and station specific rainfall distribution of 1982-2012. Each loan is linked to the closest rainfall station at the time of loan disbursement using the coordinate of the farmer's largest farm and the rainfall station coordinate, using the Euclidean distance. Origination date fixed effects are effects for the quarter of origination (for example 2005-2). The maturity fixed effect distinguishes between long and short term loans (equal to 1 for loans with maturities of three years or more). The outcome of columns (1), and (3) – (6) is a dummy for loans that ever entered into a period with overdues of 30 days or more. The outcome of column (2) is a dummy for loans that ever entered into a period with overdues of 60 days or more. Standard errors clustered at the rainfall station level are reported in parentheses.
*p<0.1; **p<0.05; ***p<0.01

Table 1.3: Effect of Rainfall Shocks on Reported BAC Scores

	Score Fell from A (1)	Score Fell to E (2)
Rainfall Shock, year 1	0.029***	0.017***
	(0.007)	(0.005)
Mean (control group)	0.150	0.085
Origin Date * Matu. FE	Y	Y
Rainfall St. FE	Y	Y
Rainfall St. Clst.	Y	Y
Observations	32,512	32,512
Adjusted R^2	0.16	0.075

Notes: Included loans are for coffee production, originated in the period of 2008-2011 and for which there is a subsequent application observed in the CIFIN Stage. The outcome in column (1) is a dummy for loans for which the BAC Reported Score fell from A to any other score at any point during loan tenure. The outcome in column (2) is a dummy for loans for which the BAC Reported Score fell from A to the lowest possible score, E. Standard errors clustered at the rainfall station level are reported in parentheses.
*p<0.1; **p<0.05; ***p<0.01

Table 1.4: Effect of Rainfall Shocks on Future Credit Access

	Applied New Loan (1)	Initial Loan Overdue (2)	CIFIN Score (3)	CIFIN Denial (4)	Analysis Denial (5)
	A. Initial Loan Maturity: All				
Rainfall Shock,	-0.005	0.034***	-5.747***	0.015***	0.017***
	(0.005)	(0.007)	(1.96)	(0.005)	(0.007)
Mean (control group)	0.816	0.21	925	0.119	0.173
Origin Date FE	N	N	N	N	N
Origin Date * Matu. FE	Y	Y	Y	Y	Y
Rainfall St. FE	Y	Y	Y	Y	Y
Rainfall St. Clst.	Y	Y	Y	Y	Y
Observations	51,102	32,512	31,939	32,512	24,083
Adjusted R^2	0.21	0.13	0.074	0.048	0.019
	B. Initial Loan Maturity: 1 Year				
Rainfall Shock	-0.003	0.024***	-7.148***	0.019***	0.019***
	(0.009)	(0.006)	(2.155)	(0.007)	(0.009)
Mean (control group)	0.835	0.149	941	0.109	0.16
Origin Date FE	Y	Y	Y	Y	Y
Origin Date * Matu. FE	N	N	N	N	N
Rainfall St. FE	Y	Y	Y	Y	Y
Rainfall St. Clst.	Y	Y	Y	Y	Y
Observations	28,177	20,549	20,161	20,549	16,368
Adjusted R^2	0.28	0.152	0.049	0.059	0.019

Notes: The data source is the BAC administrative data. One observation corresponds to one loan application of one individual, following loans irrespective of maturity (panel A) and applications following one year loans (panel B). The outcome of Column (1) is a dummy for individuals who applied for a new loan after the initial loan. The outcome of Column (2) is a dummy for loans that entered into a period of 30 days past due in the sample of initial loans. The outcome of Column (3) is the score reported by CIFIN when the farmer applies for a new loan and the outcome of Column (4) is a dummy for applications denied at the CIFIN Stage. The outcome of Column (5) is a dummy for denial at the Analysis Stage.
*p<0.1; **p<0.05; ***p<0.01

Table 1.5: Effect of Rainfall Shocks on Future Credit Access: Persistence

	First Loan: All Maturities			First Loan: 1 Year Maturity		
	CIFIN Score (1)	CIFIN Denial (2)	Analysis Denial (3)	CIFIN Score (4)	CIFIN Denial (5)	Analysis Denial (6)
	A. Time Lapse Maturity to Application : 1 Year					
Rainfall Shock, year 1	-5.7***	0.017*	0.026**	-9.8***	0.027**	0.035**
	(3.3)	(0.009)	(0.012)	(3.8)	(0.012)	(0.015)
Mean (control group)	914	0.144	0.173	924	0.133	0.173
Observations	12,797	13,108	9,533	8,215	8,431	6,661
	B. Time Lapse Maturity to Application : 2 Years					
Rainfall Shock, year 1	-9.7*	0.2	0.017	-10.9*	0.03*	0.011
	(5.168)	(0.014)	(0.02)	(5.899)	(0.016)	(0.023)
Mean (control group)	913	0.148	0.19	917	0.138	0.127
Observations	6,380	6,457	4,790	4,519	4,558	3,548
Origin Date FE	N	N	N	Y	Y	Y
Origin Date * Matu. FE	Y	Y	Y	N	N	N
Rainfall St. FE	Y	Y	Y	Y	Y	Y
Rainfall St. Clst.	Y	Y	Y	Y	Y	Y

Notes: The data source is the BAC administrative data . One observation corresponds to one loan application of one individual. Columns (1) to (3) correspond to applications after loans irrespective of maturity whereas columns (4) to (6) correspond to applications following one year loans. Each panel corresponds to a different sample depending on the duration of the time window between maturity of the initial loan and the next loan application. Each reported coefficient corresponds to a different regression. The outcomes in Columns (1)-(2) and (4)-(5) correspond to the first application observed after the maturity of the corresponding loan, in the CIFIN Consult Stage. Columns (3) and (6) correspond to the first application that makes it to the Analysis Stage. The CIFIN Score (i.e. the outcome of Columns (1) and (4)) corresponds to the score reported by the CIFIN Credit Bureau (at the time of loan application). The outcome of Columns (2) and (5) is a dummy for denial at the CIFIN Consult Stage. The outcome of Columns (3) and (6) is a dummy for denial at the Analysis Stage. Standard errors clustered at the rainfall station level are reported in parentheses.
*p<0.1; **p<0.05; ***p<0.01

Table 1.6: Income Recovery from Rainfall Shocks

	No Controls	Controls (pred.)	Controls (all)
	(1)	(2)	(3)
Rainfall Shock 04-05	-7.67**	-7.46**	-6.43**
	(3.52)	(3.49)	(2.92)
Rainfall Shock 02-03	3.75	3.64	6.15
	(4.70)	(4.61)	(5.46)
Rainfall Shock 00-01	3.82	4.30	2.96
	(5.22)	(5.19)	(5.73)
Farm Area		0.018	-0.042
		(0.20)	(0.20)
Household Size		0.64	0.51
		(0.40)	(0.38)
Education		1.86*	1.79
		(1.13)	(1.09)
Gender		-5.39***	-4.70**
		(1.56)	(1.63)
Density			2.6***
			(.5)
Average Age			-0.16*
			(0.088)
Sun Exposed			1.20
			(1.94)
Constant	24.4***	26.6***	15.3***
	(1.61)	(3.49)	(3.70)
Mean (control group)	23.9	24.2	24.3
p-value S. 04-05 = S. 03-02	0.092	0.097	0.063
p-value S. 04-05 = S. 00-01	0.119	0.109	0.201
Observations	1,296	1,256	1,242
Coffee Region FE	Y	Y	Y
Coffee Variety FE	N	N	Y
Rainfall St. Clst.	Y	Y	Y
Adjusted R^2	0.048	0.053	0.100

Notes: The data comes from the MLYCC survey. Each column corresponds to a different regression. Only farms in a distance smaller than 6.15 km to the rainfall station are included in the sample. The outcome in all regressions is the number of arrobas (one arroba = 12.5 kg) sold per-hectare cultivated with coffee. Rainfall Shock 04-05 is a dummy for the occurrence of a rainfall shock in year 2004 or 2005. Rainfall Shock 02-03 and Shock 00-01 are defined analogously. Farm Area is the size of the farm in hectares. Household Size is the number of household members. Education is an ordered variable that increases with the level of education of the household head. Gender is a dummy for female household head. Density is trees per-hectare of the coffee plot. Average Age is average age (in years) of coffee plots. Sun exposed is dummy for farms with coffee plots with no shade. 128 farms had a rainfall shock in 2004-2005. There are 246 rainfall station clusters. Of these, 26 had a rainfall shock in 2004-2005, 7 in 2002-2003, and 13 in 2000-2001. Standard errors clustered at the rainfall station level are reported in parentheses.
*p<0.1; **p<0.05; ***p<0.01

Table 1.7: Covariate Balance

	Control	Treatment	p-value	
Household Size	3.93	4.10	0.300	
Education	0.99	0.89	0.060	*
Gender	1.18	1.16	0.597	
Coffee Area	3.1	2.8	0.204	
Density	4323	4030	0.059	*
Average Age	8.3	8.7	0.540	
Farm Area	5.8	5.7	0.840	
Sun Exposed	0.19	0.24	0.130	

Notes: The data comes from the MLYCC survey. Farms in the treatment group correspond to farms for which the closest rainfall station had a rainfall shock in 2005 or 2004. Control farms are the rest. The reported p-value corresponds to a test where the null hypothesis is equality in means across treated and control groups.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 1.8: Repayment Recovery of High Credit Score Farmers

	1st Loan Overdue (1)	CIFIN Denial (2)	Analysis Denial (3)	2nd Loan Overdue (4)
Rainfall Shock	0.014**	0.011	0.012	0.033
	(0.008)	(0.012)	(0.022)	(0.021)
Mean (control group)	0.013	0.055	0.141	0.063
Origin Date FE	Y	Y	Y	Y
Origin Date * Matu. FE	N	N	N	N
Rainfall St. FE	Y	Y	Y	Y
Rainfall St. Clst.	Y	Y	Y	Y
Observations	2,681	3,785	3,141	2,550
Adjusted R^2	0.028	0.078	0.025	0.043

Notes: The data source is the BAC administrative data. Each observation corresponds to one individual. The sample of Column (1) consists of loans originated in 2010-2011, from farmers in the top quartile of the CIFIN credit score distribution and that applied for a subsequent loan. The samples of Columns (2) and (3) correspond to one loan application following loans in the sample of Column (1). The sample in Column (4) consists of loans that were approved after the initial loan. The outcomes are: Columns (1) and (4), a dummy equal to one if the corresponding loan ever entered into a period of 30 days past due, Column (2), a dummy for denial of the subsequent application at the CIFIN Stage, and Column (3), a dummy for denial of the subsequent application at the Analysis Stage. Standard errors clustered at the rainfall station level are reported in parentheses.

*p<0.1; **p<0.05; ***p<0.01

1.8 Appendix to Chapter 1

1.8.1 Robustness Exercises

This appendix presents robustness results on the effect of rainfall shocks on repayment (Table 1.2) and on scores reported by the BAC (Table 1.3). They differ from the results presented in the main text in that this sample includes all loans disbursed between 2005 and 2011.

Table 1.9: Appendix Table: Effect of Rainfall Shocks on Repayment

	Baseline			Heterogeneous Effects		
	30 Days Overdue		60 Days Overdue	Dist.to Stat. < Median	Maturity: One Year	Maturity: Five Years
	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall Shock, year 1	0.008***		0.006**	0.009**	0.008**	0.016**
	(0.003)		(0.003)	(0.004)	(0.003)	(0.008)
# quarter-shocks, year 1		0.004***				
		(0.002)				
Mean (control group)	0.157		0.118	0.147	0.147	0.242
Mean (all obs.)		0.157				
Origin Date * Matu. FE	Y	Y	Y	Y	Y	Y
Rainfall St. FE	Y	Y	Y	Y	Y	Y
Rainfall St. Clst.	Y	Y	Y	Y	Y	Y
Observations	274,198	274,198	274,198	137,097	129,474	40,796
Adjusted R^2	0.079	0.079	0.073	0.071	0.049	0.050

Notes: The data source is the BAC administrative data. Included loans are for coffee production and originated in the period of 2005-2011. In a given calendar year, a rainfall shock is defined by rainfall realizations in at least two quarters above the 80th percentile of the quarter-year and station specific rainfall distribution of 1982-2012. Each loan is linked to the closest rainfall station at the time of loan disbursement using the coordinate of the farmer's largest farm and the rainfall station coordinate, using the Euclidean distance. Origination date fixed effects are effects for the quarter of origination (for example 2005-2). The maturity fixed effect distinguishes between long and short term loans (equal to 1 for loans with maturities of three years or more). The outcome of columns (1), and (3) – (6) is a dummy for loans that ever entered into a period with overdues of 30 days or more. The outcome of column (2) is a dummy for loans that ever entered into a period with overdues of 60 days or more. Standard errors clustered at the rainfall station level are reported in parentheses.

*p<0.1; **p<0.05; ***p<0.01

Table 1.10: Appendix Table: Effect of Rainfall Shocks on Reported BAC Scores

	Score Fell from A (1)	Score Fell to E (2)
Rainfall Shock, year 1	0.008**	0.005*
	(0.003)	(0.002)
Mean (control group)	0.150	0.085
Origin Date * Matu. FE	Y	Y
Rainfall St. FE	Y	Y
Rainfall St. Clst.	Y	Y
Observations	274,198	274,198
Adjusted R^2	0.07	0.07

Notes: The data source is the BAC administrative data. Included loans are for coffee production and originated in the period of 2005-2011. The outcome in column (1) is a dummy for loans for which the BAC Reported Score fell from A to any other score at any point of loan tenure. The outcome in column (2) is a dummy for loans for which the BAC Reported Score fell from A to the lowest possible score, E. Standard errors clustered at the rainfall station level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

1.8.2 Theory

In this Appendix, I show with a simple model of lending and credit scoring that the exclusion of exogenous shocks that affect repayment leads the lender to make mistakes more frequently. A mistake is defined as lending to a non-profitable borrower or denying credit to a profitable one. This model is related to work by de Janvry, McIntosh, and Sadoulet (2010). However, unlike previous work, I show how exogenous shocks can lead to a market failure.

Consider a scenario with two periods where a borrower and a lender interact. Suppose that the borrower was granted a loan in period $t-1$ and denote the repayment of this loan by π_{t-1}. The borrower always applies for a new loan for period t and the lender must decide if he grants this subsequent loan. The borrower is characterized by a level of profitability π_0 which is unobserved by the lender.

Repayment of the borrower in $t-1$ depends on π_0 and two random components. More precisely, I assume that:

$$\pi_{t-1} = \pi_0 + z + \epsilon \tag{A1}$$

where z is an exogenous shock (in the sense that it is independent of π_0) and that is potentially observable by the lender. I refer to z as a "rainfall shock". ϵ is another exogenous component (independent of both π_0 and z) but unobservable to the lender. I assume that $z \sim N\big(0, \sigma_z^2\big)$ and $\epsilon \sim N\big(0, \sigma_\epsilon^2\big)$. Furthermore, I assume that the lender knows the process that generates π_{t-1} (that is, it knows equation A1) but does not observe all of its components. That is, he cannot observe π_0, z, and ϵ separately.

The repayment of the subsequent loan (in the case the lender is given one) is denoted by π_t. I assume away any uncertainty in the repayment of this second loan (once the lender has made his decision) so that $\pi_t = \pi_0$. Therefore, the lender makes a positive profit in the second loan if $\pi_0 > 0$ and negative one if $\pi_0 < 0$. To make the decision of whether or not to grant a second loan, the lender makes a prediction of π_t based on past borrower behavior. In particular, he forms a "credit score" based on π_{t-1}. In the case where the rainfall shock is not observed, it is given by:

$$E[\pi_t|\pi_{t-1}] = \pi_{t-1} \tag{A2}$$

The lender grants the loan if $E[\pi_t|\pi_{t-1}] \geq 0$ and does not otherwise.

Note that this setup has various parallels with my empirical setting. First, z affects repayment of the first loan π_{t-1} but does not affect repayment of the second loan π_t. This is consistent with the effect of rainfall shocks. Since there is recovery, rainfall shocks affect repayment of the first loan but don't affect repayment of the second one. Second, the main determinant of the credit score is past repayment behavior. Individual characteristics (which are sometimes included in credit scores) could be included in this setup, but they would not modify the results of the model.

Under the stated assumptions I can compute the probability that the lender makes a mistake. In particular, I compute the probability that lender grants a loan to an un-profitable borrower and the probability that it does not grant a loan to a profitable one. I do this for two scenarios: one in which the rainfall shock z is not observed by the lender and one in which it is.

Scenario 1: z is Unobservable to the Lender

The lender grants a second if $E[\pi_t|\pi_{t-1}] \geq 0$ which is equivalent to $\pi_{t-1} = \pi_0 + z + \epsilon \geq 0$ from equation (A2) and (A1). I denote by P_u the probability that a loan is granted to an unprofitable borrower. Therefore, $P_u = P(\pi_{t-1} \geq 0) = P(\pi_0 + z + \epsilon \geq 0)$ given $\pi_0 < 0$. Note that $z + \epsilon$ is distributed $N\big(0, \sigma_z^2 + \sigma_\epsilon^2\big)$ since z and ϵ are independent so that $P(\pi_0 + z + \epsilon \geq 0) = P(z + \epsilon \geq -\pi_0) = P(z + \epsilon \leq \pi_0)$ and

$$P_u = \Phi\left(\frac{\pi_0}{\sqrt{\sigma_z^2 + \sigma_\epsilon^2}}\right) \tag{A3}$$

Note that this last expression is increasing in $\sqrt{\sigma_z^2 + \sigma_\epsilon^2}$ given that $\pi_0 < 0$. Therefore, the lender is more likely to make the mistake of lending to an unprofitable borrower the larger the variance of z or ϵ. The intuition is that a larger variance implies that the signal π_{t-1} is less informative on the profitability of the second loan, π_0.

Now I consider the probability that the lender denies a loan to a profitable borrower and denote it by P_d. P_d is given by $P(\pi_0+z+\epsilon \leq 0)$ given $\pi_0 > 0$. Therefore, $P_d = P(z+\epsilon \leq -\pi_0)$ so that, under the distributional assumptions of z and ϵ, it can be written as:

$$P_d = \Phi\left(\frac{-\pi_0}{\sqrt{\sigma_z^2 + \sigma_\epsilon^2}}\right) \tag{A4}$$

Given that $\pi_0 > 0$, this expression is also increasing in $\sqrt{\sigma_z^2 + \sigma_\epsilon^2}$. Again, the probability that lender makes the mistake (in this case of not lending to a profitable borrower) is increasing in the variance of the terms z and ϵ.

Scenario 2: z is Observed by the Lender

If z is observed by the lender, his prediction of π_t changes. In particular, since he knows the process generating π_{t-1} (given by equation A1) he discounts π_{t-1} with the observed value of z. In other words, the credit score is now $E[\pi_t|\pi_{t-1}, z] = \pi_{t-1} - z$. Substituting equation A1 yields $E[\pi_t|\pi_{t-1}, z] = \pi_0 + \epsilon$. As before, the lender grants the second loan if $E[\pi_t|\pi_{t-1}, z] \geq 0$ and does not otherwise.

In this case, the probability of granting a loan to an unprofitable borrower is given by $P_u = P(E[\pi_t|\pi_{t-1}, z] \geq 0) = P(\pi_0 + \epsilon \geq 0)$ with $\pi_0 < 0$. Therefore,

$$P_u = \Phi\left(\frac{\pi_0}{\sigma_\epsilon}\right) \tag{A5}$$

with $\pi_0 < 0$. The probability of denying a loan to a profitable borrower, P_d in this scenario is given by $P(\pi_0 + \epsilon \leq 0)$ with $\pi_0 > 0$, that is:

$$P_u = \Phi\left(\frac{-\pi_0}{\sigma_\epsilon}\right) \tag{A6}$$

with $\pi_0 > 0$.

The probabilities of lender mistakes of both scenarios can be easily compared from equations A3 to A6. Clearly, the probability of lending a loan to an un-profitable borrower, P_u, is larger in Scenario 1 (equation A3) than in Scenario 2 (equation A5), given that both expressions are increasing in the term in the denominator and $\sigma_z^2 > 0$. Similarly, the probability of denying a loan to a profitable borrower, P_u, is larger in Scenario 1 (equation A4) than in Scenario 2 (equation A6). The difference in the mistakes probabilities is larger the larger is the variance of the "rainfall shock", σ_z^2. The intuition is that if the variance of z is larger, the precision of the signal π_{t-1} increases more when the z is included in the credit score.

These results imply that the inclusion of exogenous shocks in the credit score increase its precision and reduce the probability of lender mistakes. The gains from this inclusion are larger if the variance of the excluded shocks is large. This is particularly true in the context of agricultural lending in developing countries where exogenous shocks can have large effects on production and income, as shown in the main text.

Chapter 2

Does Public Make a Difference? Experimental Evidence from Agricultural Credit in Colombia

Nicolás de Roux, Jairo Esquivel, Margarita Gáfaro, Moisés Mahecha and Guillermo Otero[†]

[†] This paper benefited from comments of Pedro Dal Bó, Andrew Foster, Jonas Hjort, Suresh Naidu, Kiki Pop-Eleches, Anja Sautmann, Miguel Urquiola and Eric Verhoogen. We thank Manuel Buriticá, Fabio Díaz, Judy Laiton and Carlos Ferro from the Banco Agrario de Colombia. Andrés Felipe Fajardo provided excellent research assistance. All errors are our own.

2.1 Introduction

Small farmers contribute a large share of agricultural production in developing countries. Access to credit is crucial to improve their productivity since it facilitates investments in modern inputs and technologies. But information asymmetries and high levels of risk in agricultural production cause private banks to ration small farmers out of credit markets (Carter, 1988). Specifically, small scales of production, informality in land tenure, and bans to land foreclosures preclude some farmers to post collateral for private banks (Feder, Onchan, and Raparla, 1988; Escobal and Mediano, 2006; Fafchamps, 2013). Furthermore, monitoring requirements and a deficient transport infrastructure increase the costs of providing credit in rural areas (Carter, 1988; Barham, Boucher, and Carter, 1996). In response, governmental programs to facilitate credit access have arisen.

Some governments post the collateral for loans granted by private banks to small farmers.[1] Others allocate credit through state-owned financial institutions, which offer subsidized interest rates and low collateral requirements (Klein, Meyer, Hannig, Burnett, Fiebig, et al., 1999). However, these schemes can exacerbate moral hazard in repayment effort. Theoretical models emphasize the importance of collateral to generate incentives for loan repayment. If borrowers are not required to post collateral, the cost of default will be lower and may discourage timely repayment and increase default rates (Bardhan and Udry, 1999; Ghatak and Guinnane, 1999; Kurosaki and Khan, 2012). Moreover, there is a generalized belief that borrowers may not have strong incentives to repay loans from state-owned banks because these institutions often face solvency problems associated with mismanagement, corruption and weak deterrence (Snowden, 1991; De Aghion, 1999; Scott, 2007; Hainz and Hakenes, 2012).

In this paper, we study the motives behind the repayment decision of small farmers when credit is provided by a public institution and collateral losses are not involved. We focus on

[1] See, for example, the case of Mexico and Nigeria in FAO (2013).

three issues that we consider critical about agricultural credit markets and public programs that allocate credit to farmers. First, we explore if moral hazard determines repayment behavior in this context. Second, and more importantly, we study to what extend the public nature of the lender determines moral hazard and strategic default. Finally, we build on the literature studying the determinants of contributions to public funds and tax compliance to explore the mechanisms through which the public nature of the lender affects moral hazard.

To study these issues we implemented an experiment with the Banco Agrario de Colombia (BAC) [Colombian Agrarian Bank], a state-owned financial institution that provides more than 90 percent of the credit granted to small farmers in Colombia (Ocampo, 2014). Most BAC clients are not required to post collateral for their loans, but their repayment behavior is reported to credit scoring agencies. As part of its normal functioning, the BAC works with call centers that reach clients to remind them that a payment installment is upcoming.

The experiment consists of different treatment arms that modify the traditional script that the call center operators follow during the phone call. Each treatment makes salient a particular feature of the environment where the farmer makes his decision to repay. The motivation for this design is as follows. Suppose that bringing to mind a particular feature of the environment modifies the decision of the borrower. Then, we will interpret this result as evidence that the decision is affected by this feature, in line with behavioral theories of salience and consumer choice.[2] In other words, the feature contained in the reminder is a motive that drives a particular behavior. For example, if reminding farmers that the bank is public affects their repayment decision, then we will interpret this as evidence that farmers pay (more or less) *because* the bank is public. This is the same type of interpretation of Bursztyn, Fiorin, Gottlieb, and Kanz (2016), who argue that morale is a cause of repayment. Their conclusion is based on experimental evidence that reminders that make salient a moral

[2]See Taylor and Thompson (1982) for a discussion in the psychological literature. Chetty, Looney, and Kroft (2009) propose a model of tax salience and consumer choice. Bordalo, Gennaioli, and Shleifer (2015) elaborate on their model and refer to the consumers of Chetty, Looney, and Kroft (2009) as FBOR (forgetful but otherwise rational). These individuals forget some aspects of a problem but when they are reminded of them, they take them into account in their decision process.

or religious cost of not paying credit cards of an Islamic bank increase repayment.

The main result of this paper is that farmers treated with calls that include a reminder of the public character of the bank substantially improve repayment under different default measures, compared to farmers treated with the traditional script. More precisely, compared to the control group, farmers in this treatment had a probability of ever being past due that is 10% lower, and probabilities of entering into periods of 15 days and 30 days past due that are, respectively, 18% and 22% lower. The responsiveness of repayment performance to our main treatment suggests that moral hazard plays a role in this context, as borrowers can exercise some discretion about the timing of their payments. More importantly, we interpret these results are suggesting that farmers repay their loans, at least in part, because of the public nature of the bank. This result is surprising. As discussed above, there are reasons to believe that moral hazard is less costly when the lender is a public bank because collateral loses are not involved and debts are more likely to be forgiven. In fact, among officers of the BAC, a widely held belief is that customer default on their loans because of the public nature of the bank.[3]

To test monetary determinants of debt repayment, we estimate the effects of a treatment in which we exclude the information about the public nature of the bank and we remind the borrower about the consequences of debt default on future access to credit. This treatment decreases by 17 percent the probability of entering into a period of 15 days past due. With other measures of repayment, we do not find statistically significant effects of this treatment although all the point estimates are negative. The null effects we find are imprecise, and we are not able to rule out economically significant effects or to distinguish these coefficients from the coefficient estimates of our main treatment. Although not strongly, these results suggest that monetary reasons play a role in motivating repayment.

---

[3]Since we compare the effect of our treatments relative to a control group that received a call with the traditional script, our results partial out the effect of mitigating limited attention on repayment behavior. Also, it is not likely that the effects we identify are driven by borrowers updating their expectations about stronger collection enforcement from the bank, as these phone calls are part of the traditional debt collection strategy of the BAC.

We explore different sources of heterogeneity to inform the previous results. We find heterogeneous effects of the public treatment for certain measures of state presence (or state capacity). In particular, the number of homicides and the number of hectares cultivated with coca in the municipality of the investment mitigate the effect of the public treatment on repayment. This result is consistent with a state deterrence mechanism underlying (at least in part) the effect of the public treatment: when the public nature of the BAC is mentioned in the script, it cues awareness of the long arm of the state, thus increasing repayment rates. Regarding the future access to credit treatment, we find that it is stronger for farmers with higher education levels (in particular for those who finished high school). We find no heterogeneous effects by amount of debt outstanding, or, at the municipality level, the size of the municipality (in squared kilometers), the size of its rural population, or the number of subversive actions carried by illegal armed groups.

To disentangle some behavioral features of the public nature of the bank that might induce farmers to repay, we designed other four treatments in our experiment. In each one of them, we kept the sentence in the script where we state that the bank is state owned but we added an additional sentence intended to induce feelings of altruism, reciprocity or peer pressure. In particular, in the first one of these treatments, we tell clients that timely repayments allow the bank to extend its benefits to more farmers (thus evoking altruism). In the second one, we state that timely repayment is a way of showing gratitude to the bank (thus evoking reciprocity). The third and fourth of these treatments give clients information about the repayment behavior of other farmers in their municipality or engaged in the same productive activity as theirs. We find that the treatments evoking altruism and peer pressure from producers of the same product lead to better repayment outcomes compared to the control group. For example, the probability of ever being overdue is lower in 10% and 14% (in the altruism and the activity peer treatments respectively), compared to farmers treated with the traditional script. These effects hold for our other measures of repayment performance. Nevertheless, we are not able to statistically differentiate these effects from the effects of the treatment with only information about the public nature of the bank.

We consider two alternative explanations for this result. The first one is that altruism and peer effects might be irrelevant in determining repayment performance. In this case, the effect we find of the altruism and activity peer treatments would be entirely driven by the behavioral response to the reminder that the bank is a state-owned institution, which in this case would not depend on altruism or peer effects. The second explanation is that the reminder of to the public treatment alone already encompasses, or brings to mind, feelings of altruism and peer pressure. Therefore, the additional information does not add anything to the information of the public treatment. In this case, we say that the public nature of the bank influences behavior, in part, because it relates to altruism and peer pressure.

Although it is difficult to distinguish between these two explanations, we discuss evidence that supports (although not strongly) the assumption that altruism and peer effects have an effect on repayment. Under this interpretation, our results suggest that repayment to state-owned banks is mediated (at least in part) by altruism towards other potential borrowers and peer pressure. This finding is consistent with strategies that state-owned banks have implemented to improve repayment in other contexts.[4]

Our paper contributes to different literatures. First, it adds to studies of moral hazard in credit markets and non-monetary determinants of pecuniary payments. Closely related to our work, Karlan, Morten, and Zinman (2015) show, with experimental evidence for the Philippines, that sending text messages to microcredit clients with reminders of their due date, only increases timely repayment if the message includes the name of the loan officer. The authors argue that feelings of reciprocity towards the bank officer increase repayment effort.[5] Similarly, Bursztyn, Fiorin, Gottlieb, and Kanz (2016) use text messages to appeal to moral considerations about repayment among delinquent credit card users in Indonesia. They show

[4]For example, in 2016 a financial institution owned by the Indian government launched a television advertisement in which a little girl appeals to feelings of reciprocity towards the bank to convince her father to repay his debt (Gopakumar, 2016).

[5]In a related study (Cadena and Schoar, 2011) find that reminders in the form of text messages do increase timely repayment of micro loans in Uganda.

that receiving these messages increases the share of clients meeting their repayment obligation by 15 percent. We provide further experimental evidence on the existence of moral hazard in loan repayment and on the effects of non-monetary motives of repayment effort among farmers with access to credit from a public bank in Colombia.

Our results add new empirical evidence on the relation between the government and beneficiaries of public credit programs. Contrary to the common notion of poor repayment performance to state-owned banks, we show that reminding the client that the BAC is a public institution improves repayment outcomes. Moreover, our evidence suggests that both altruism and peer effects explain part of the repayment response we observe. It has been widely documented that under joint liability schemes, peer effects can improve repayment through screening, monitoring and risk sharing among group members (Ahlin and Townsend, 2007; Karlan, 2007; Giné, Krishnaswamy, and Ponce, 2011; Carpena, Cole, Shapiro, and Zia, 2012; Breza, 2015). Our focus on peer effects under individual liability is related to work by Giné and Karlan (2014), who show with experimental evidence that microcredit borrowers under individual liability schemes present similar repayment rates to those in joint liability groups. Similarly, Guiso, Sapienza, and Zingales (2013) show that individuals in the U.S are more likely default in their mortgages if they know other people who have defaulted. We argue that repayment to a state-owned institution is partly determined by peer pressure.

This reminder of the paper is organized as follows. The next section presents the experimental design and describes the implementation. In section 2.3 we discuss our results and in section 2.4 we provide some concluding remarks.

2.2 Experimental Design and Implementation

The BAC partners with call centers to make regular phone calls to its customers before a due date and remind them that a payment is due. Customers are randomly assigned to different call centers. We partnered with one of these call centers to implement the experiment and focused on loans with a payment due in August 2016. From these we select loans for agricultural production to small farmers.[6] This amounted to a total of 15,000 loans.

Our experiment consists of seven different treatments. We randomly assigned 1200 loans (out of the 15000 loans) to each one of them. As discussed previously, the treatments were designed to test whether the public nature of the BAC influences the repayment of its customers. When the call center operator calls the customer, he reads a script to each client. Each treatment adds a sentence at the end of this script.[7] Table 2.1 lists the different treatments and shows the sentence added to the script. Treatment 0 (t0) is the control treatment, where no sentence is added to the usual script. Treatment 1 (t1) is the main treatment of interest. We refer to it as the "Public Treatment." When assigned to it, the customer is told: "We remind you that BAC is a financial institution, in its mayor part property of the Colombian State." Treatments 3 to 5 test different mechanism through which we hypothesize that the Public Treatment affects repayment. For example, Treatment 3 contains the sentence: "We remind you that BAC is a financial institution, in its main part property of the Colombian State. Your timely payment will help us to continue financing other agricultural producers." Finally, Treatment 6 test whether a motive of loan repayment is related with future credit access with the following sentence: "We remind you that keeping the payments of your loans up to date is the best commercial reference and facilitates future access to loans from the BAC and from other financial institutions."

The call center assigns loans quasi-randomly to its operators.[8] Any given operator could be assigned any loan in the 8400 of the experiment, so that, within operator, the treatments

[6]A small farmer is person engaged in agricultural productive activities with a value of assets smaller than 120 millions of Colombian pesos (which equal about US$ 39000, at the average daily exchange rate of the first semester of 2016, of 3098 Colombia Pesos per US$)

[7]Appendix A shows a translation from Spanish of the script and the place where the sentence is read.

[8]The call center uses a software to make a list of the customers to be called in a particular day. Customers are assigned to operators depending on the time that the operator starts working and on the speed in which the phone calls of a particular day are made. As the operator finishes with a set of loans, new loans in the queue are assigned to her.

are random. We instructed the call center to make the phone calls no more than a week before the payment was due.[9]

At the beginning of the experiment, the call center operators were told that an exercise was taking place and that for some loans, the script would change. A given call center operator makes around 170 calls in a day and, during the time of the experiment, she made calls for loans in the experiment intertwined with calls for loans that were not. When a given call corresponded to a loan in the experiment, the operator was warned in her computer screen. Figure 2.1 shows the typical screen of an operator during the experiment. In the example of the figure, a loan in Treatment 5 is depicted. The operator can see in the screen the sentence she has to read at the end of the call. Operators were warned that a loan was part of the experiment even in the case of the control group (note the line "Pro–Agra–T1"), even if they just had to read in those cases the traditional script. This controls for the fact that even if the script is the traditional one, operators could have treated loans differently if they knew they were part of the experiment.

Finally, some obligations where not treated. These correspond to cases where the customer made the payment before the programmed treatment date. Also, in some cases the customer was not reached by voice. For example, there were cases where the customer never picked up the phone. In other cases, a message was left for him (usually with a family member) to return the call.

Table 2.2 shows pretreatment summary statistics. For each treatment, the table shows the mean of the variable and, in parentheses, its standard deviation. We ran a separate regression of each variable on a set of dummies for each treatment (excluding the control dummy) and report the p-value of F-test for joint significance. The table shows that there are no large

[9]For most of the loans this was the case. Nevertheless, in some occasions the bank instructed the call center to treat some groups of loans in advance, for example in a particular region of the country. In this case though, our treatments where used, and within these groups of loans, which script was used was random. Nevertheless, this implied that the loans were treated in a date possibly before a week of the due date. Our results are robust to controlling for the time window (in days) between the day of the treatment and the due date, as shown in Appendix B.

differences across treatments in pretreatment characteristics of the loans. For example, the size of the loans is very similar across treatments and close to 7,500 thousand Colombian pesos (about 2,400 US$). The p-value of a test of joint significance is 0.4. Therefore we cannot reject the null hypothesis that the treatments are jointly uncorrelated with loan size.

2.3 Results

2.3.1 Simple Mean Differences

We focus on four outcomes of delay in repayment or default. The first one is dummy for whether the loan was ever past due in a window of 120 days after the payment installment date in august. The second one is the maximum number of days past due of the loan in the same 120 days window. The third outcome is a dummy for whether the loan entered into a period of 15 days past due in the first month after the date of the payment installment. Finally, the fourth outcome is a dummy for whether the loan entered into a period of 30 days past due in a window of 60 days after the date of the payment installment.

Table 2.3 shows the simple mean differences across treatments for our four measures of default. Statistically significant difference are denoted with stars. The table shows that 31 percent of loans in the control group were ever overdue. Albeit the differences are not statistically significant, for all outcomes, farmers in the public treatment (t1) show better repayment performance under all the measures of default than the control group. The same is true for treatments t2 (public & altruism), t3 (public & reciprocity), t4 (public & activity peer), and t6 (public & future access) with some of these differences being statistically significant. In contrast, giving information about repayment rates in the municipality of the borrower seems to worsen repayment performance: the mean of all default outcomes is higher for farmers in this treatment compared to the control group.

2.3.2 Regression Estimates

In other to improve the precision of the estimated effect we do a regression analysis where each outcome of interest is regressed on treatment dummies and a set of predetermined loan characteristics. As discussed in Section 2.2, since the call center operators were not able to reach all borrowers assigned to each treatment arm, our estimates are of Intention-to-Treat (ITT) effects, which are obtained with the whole sample of borrowers in the randomization.

Specifically, we estimate by OLS the following equation

$$y_i = \alpha + T_i'\beta + X_i'\gamma + u_i, \tag{2.1}$$

where y_i represents one of our four measures of default and T_i is vector of indicators for the treatment assignment. We include a set of loan and borrower characteristics, X_i, to improve the precision of our estimates. These control variables include the size of the loan, a dummy variable indicating if the loan is the result of restructured debts, the age of the loan at the time of the due date, the value of household's assets, a dummy for borrower's sex, a dummy for college degree, BAC office fixed effects, and the BAC's ex-ante credit score.[10] u_i is a mean zero error term. The identifying assumption is that $E[u_i T_i] = 0$, which follows from the random assignment of the elements in T_i.

Table 2.4 presents the results of an estimation including our main treatment (t1 - public) and the treatment where customers are reminded of the negative consequences of delinquency for future access to credit (t6 - future access). The table shows that for all measures of performance, giving information about the public nature of the BAC leads to better repayment outcomes. In particular, the probability of ever being overdue decreases by 3.4 percentage points, which represents a 10 percent reduction with respect to the control group. Furthermore, there is a significant effect on the number of days past due, which fall by 0.7 in the public treatment. The probability of entering into a period of 15 days past due decreases in

[10]We use the score of the BAC's own credit scoring model.

18 percent and the probability of entering into one of 30 days decreases in 22 percent with respect to the control group. These results provide evidence of the existence of strategic default and moral hazard among BAC borrowers. Furthermore, they show that the public nature of the bank does not imply worse repayment outcomes. On the contrary it implies better repayment behavior.

It is interesting that the public treatment has not only short run effects (for example, it reduces the probability of entering into short overdue periods, captured by the first outcome) but also longer term outcomes (the probability of entering into periods of 15 or 30 days past due). The fact that this last outcome is affected is important since this is the main measure of default used by the BAC and financial regulators.

The treatment in which the negative consequences of debt delinquency on future credit access are reminded (t6 - future access) has in general imprecise effects on loan repayments. We only find that this treatment is statistically significant for repayment delays longer than 15 days, with a 17 percent decrease with respect to the control group. Nevertheless, the negative point estimate across outcomes suggests that an important monetary reason for farmers to repay their loans in this context (so as not to lose future access to credit) plays a role, although not as important as the public nature of the bank.[11]

2.3.3 Heterogeneity

In this section we study the heterogeneity of the effect of the public and future access to credit treatments in a regression framework. In particular, we estimate by OLS the following equation:

[11] It is also possible that lost of future credit access is always in the top of the mind of farmers and that the reminder does not make this feature of the environment more salient. Under this scenario we would also find weak effects of this treatment. This interpretation, though, is not consistent with the heterogeneity results by education level presented below.

$$y_{im} = \mu + T_i'\eta + \delta z_{im} + (T_i' z_{im})\theta + X_i'\gamma + C_m'\rho + e_i, \tag{2.2}$$

where i denotes a loan and m a municipality.[12] z_{im} is the heterogeneity variable, which can vary at the loan or the municipality level. As before X_i is a vector of individual controls. y_{im} is a dummy equal to one if the loan entered into a period of 15 days past due. T_i is a vector of dummies for the treatment. C_m is a set of controls that vary at the municipality level.[13] The coefficient of interest is that of the interaction term, θ. e_i is a mean zero error term.

Table 2.5 shows the results of the heterogeneity analysis. Each column corresponds to one source of heterogeneity. This table is equivalent to Table 2.4 (for the 15 days past due outcome) but it includes the corresponding interaction terms. In all specifications, the main effect of the public treatment is statistically significant and of negative sign. Its magnitude is similar to the one reported in Table 2.4. The first column considers as source of heterogeneity the amount of debt outstanding at the moment of the payment installment. We find no heterogeneous effect on either the public or the future access to credit treatments by this variable (neither the coefficients of $t1 \times z_{im}$ or $t6 \times z_{im}$ are statistically significant). In the second column, we study heterogeneity by educational attainment by interacting the treatment dummies with a dummy equal to one for farmers that finished high school. The interaction term with the public treatment is not statistically significant. The interaction with the future access to credit treatment is negative and significant at the 10% level. This implies that the effect reported in Table 2.4 for the future access to credit treatment is driven

[12]The municipality of a given loan corresponds to the municipality where the investment for agricultural production of the loan is done. All municipality data comes from the Panel Municipal del CEDE, of the Universidad de los Andes. This data set contains municipality level information at a yearly frequency.

[13]These controls are the rural population of the municipality and the total income of the municipality's government (which proxies for the size of the municipality's government).

by farmers with high levels of education.[14] A possible interpretation for this result is that only for educated farmers future access to credit is relevant or meaningful. The results reported under columns "Municipality Area" and "Rural Population" show that the treatments are not heterogeneous by these measures of municipality size (the interaction terms are non-significant).

Finally, we study the heterogeneity of the effect by variables that proxy for state presence or state capacity. We consider the number of homicides, the number of hectares cultivated with coca and the number of subversive actions conducted by illegal armed groups. In the three cases, an increase in the variable implies a lower level of state presence or state capacity. We find that for the number of homicides and hectares cultivated with coca, the interaction term with the treatment effect is positive and statistically significant. Therefore, the higher the number of homicides or the number of hectares cultivated with coca, the lower the effect of the public treatment on repayment. This results is consistent with a state deterrence mechanism. In municipalities were the state is less strong or has less enforcing power, repayment is less affected by the public cue in the script. Nevertheless, we find no heterogeneous effects by the number of subversive actions since the interaction terms in the last column of Table 2.5 are not statistically significant

2.3.4 Behavioral Mechanisms

Our results so far provide evidence suggesting that, contrary to common beliefs, the public character of the lender leads to better repayment performance. The heterogeneity exercises documented above suggest that there is some room for a state deterrence mechanism behind this effect. We now turn to study behavioral mechanisms that could potentially explain the effect of the public treatment. Table 2.6 shows the coefficient estimates from a specification in which we include the full set of treatment indicators. As suggested by the result of the

[14] In our sample of farmers only 10% finished high school. Therefore, our dummy captures high educational levels relative to the average education level in the sample.

simple mean comparison (Table 2.3), treatments t1 (public treatment), t2 (public & altruism) and t5 (public & activity peer) lead to better repayment outcomes (in three out of our four performance measures). In contrast, the coefficient estimates of treatment t4 (public & municipality peer) are positive, close to zero and not statistically significant.

For interpretation purposes, it is important to explore if the effects of t2 (public & altruism) and t5 (public & activity peer) differer from those of t1 (public). Table 2.7 reports the p-values from a test where the null hypothesis is that the effect of each treatment (t2 - t6) does not differ from the treatment where we only remind the farmer of the public character of the bank (t1). The table shows that the magnitude of the effects of t2 (public & altruism) and t5 (public & activity peer) cannot be distinguished from that of t1 (public).

The only significant differences with the magnitude of the effect of t1 (public) are relative to t4 (public & municipality peer). These differences are significant at the 5% level in three out of four outcomes and at the 10% level in one. These results imply that telling farmers that others in the municipality repay their loans in time has a strong negative effect on repayment that counteracts the positive effects of the public nature of the bank in repayment performance. There are two interpretations for this result. First, it is possible that treatment induces a free-rider effect. Under this interpretation, farmers think that because others in their municipality are repaying in time (and hence contributing to the public fund that is the BAC) the marginal benefit of their contribution is lower. Second, it is also possible that they simply do not relate in terms of identity to their municipality. Feigenberg, Field, and Pande (2010) show that more frequent interactions across borrowers from a microcredit institution build social networks that induce informal risk-sharing and reduce default rates. The fact that treatment t4 (public & municipality peer) does not induce differential repayment with respect to the control group suggests that the monitoring and risk sharing do not to play an important role operating through geographical proximity.

More importantly, the fact that we cannot distinguish the magnitudes of the effects of treatment t1 (public) and from those of treatment t2 (public & altruism) and treatment t5 (public & activity peer) reveals information on the mechanisms behind the effects on repayment of the pure public reminder of t1. We consider two alternative hypothesis about the way in which these elements determine repayment behavior. Their plausibility depends on whether altruism and peer pressure affect repayment directly or through the relation between the borrower and the public institution. On the one hand, it is possible that both altruism and activity peers have zero effects on repayment and that the relation of the borrower with the public bank is independent of these two elements. In that case, the treatment effects we estimate capture the effect of other mechanisms that determine the behavioral responses to the acknowledgment of the bank as a state-owned institution and that we consider in the section below (or that we do not consider in this paper). On the other hand, it is possible that altruism and peer pressure have an effect on repayment performance. In that case, the coefficients estimates across treatments would not change because the response to the public reminder alone already encompasses the response to feelings of altruism and peer effects.

Under the assumption that altruism and peer pressure have an effect on repayment, the second hypothesis is more plausible. Our evidence and previous work from other authors favors, although not strongly, this assumption. First, peer effects at the municipality level have a strong effect counteracting the potential benefits of the public information treatment. This suggests that peer effects indeed have a role determining strategic repayment delays. Second, in theory, altruistic and reciprocity considerations can influence behavior (see for example Sen (1977); Benabou and Tirole (2006)) and empirically, several studies have documented that considerations of a moral kind are important to the choices of economic agents and can even trump selfish motivations. For example, Pruckner and Sausgruber (2013) argue that a moral reminder increases the level of honesty in the context of payments to an honor system cash box for newspapers in Austria, and Hallsworth, List, Metcalfe, and Vlaev (2015) find that moral suasion considerably enhances compliance with pecuniary obligations to tax authorities in the UK. Similarly, Dal Bó and Dal Bó (2014) show that exposure to messages with moral standards results in a significant increase in contributions in a voluntary contribution game. Moreover, there is empirical evidence supporting the idea of agents with an altruistic utility function in different contexts (Andreoni, 1990; Goeree, Holt, and

Laury, 2002; Korenok, Millner, and Razzolini, 2013). Also, evidence from public good games suggests that agents choose their level of contribution according to their beliefs about the contributions of others (Fischbacher, Gächter, and Fehr, 2001; Chaudhuri, Paichayontvijit, et al., 2006; Croson, 2007), and from the public finance literature Dwenger, Kleven, Rasul, and Rincke (2006) show that intrinsic motivations such as altruism and reciprocity induce tax compliance in a context with weak deterrence.

These results support - although not definitively - the assumption that both altruism and peer effects by themselves have an effect on repayment. Therefore, the effect of the public treatment (alone) could be mediated in part by altruism and peer effects, since we do not find statistical difference in the effects of t2 (public & altruism) and treatment t5 (public & activity peer) relative to t1 (public).

2.4 Conclusions

In this paper we implemented a randomized experiment to explore the motives of loan repayment to state-owned banks. We showed that farmers are more likely to repay their loans if they are reminded of the fact that the bank is public. We argued that these results can be interpreted as evidence that people repay their loans, at least in part, because of the public nature of the bank. We studied the heterogeneity of the effect of the public and future access to credit treatments according to different variables. We find that the effect of the future access to credit treatment is driven by farmers with relatively high level of education. Furthermore, we find heterogeneous effects by variables that proxy for state presence or state capacity. This result suggests that state deterrence can explain in part the fact that repayment increases when awareness of the state is cued by the script. Results obtained with treatments where sentences designed to induce feelings of altruism and peer pressure suggest that these motives could be behind the effect of the public treatment.

Our results have two implications for policy. First, the way in which reminders are framed can have positive consequences in terms of repayment, lessen moral hazard and therefore lead to welfare improvements in credit markets. Second, and perhaps contrary to popular belief in developing countries, they show that relations with public institutions do not necessarily lead to more costly behavior from individuals.

2.5 Figures Chapter 2

Figure 2.1: Call Center Operator Screen

2.6 Tables Chapter 2

Table 2.1: Treatments

Treatment	Name	Sentence in Script
t0	Control	[No sentence added]
t1	Public	We remind you that BAC is a financial institution, in its main part property of the Colombian State.
t2	Public & Altruism	We remind you that BAC is a financial institution, in its main part property of the Colombian State. Your timely payment will help us to continue financing other agricultural producers.
t3	Public & Reciprocity	We remind you that BAC is a financial institution, in its main part property of the Colombian State. With you timely payment, you fulfill the obligation you acquired and, above all, you thank the Bank and the State for financing your agricultural activities.
t4	Public & Municipality Peer	We remind you that BAC is a financial institution, in its main part property of the Colombian State. Furthermore, we inform you that an important part of our clients in your municipality makes their payments on time. We invite you to make part of this group.
t5	Public & Activity Peer	We remind you that BAC is a financial institution, in its main part property of the Colombian State. Furthermore, we inform you that an important part of our clients engaged in [*Activity*] pays on time. We invite you to make part of this group.
t6	Future Access	We remind you that keeping the payments of your loans up to date is the best commercial reference and facilitates access to future loans from the BAC and from other financial institutions.

Note: In t5 [*Activity*] depends on the productive activity of the farmer. For example, if the farmer is a coffee grower, the sentence read was: "We remind you that BAC is a financial institution, in its mayor part property of the Colombian State. Furthermore, we inform you that an important part of our clients engaged in coffee growing make their payments on time."

Table 2.2: Baseline Differences Across Treatment Arms

	t0	t1	t2	t3	t4	t5	t6	t7	p-value
loan size	7525.3	7381.5	7480.2	7514.2	7436.6	7704.8	7392.1	7691.5	0.400
	(4269)	(4125)	(4014)	(4058)	(4036)	(4396)	(4067)	(4464)	
principal balance	6589	6548	6618	6543	6537	6782	6451	6749	0.530
	(4226)	(4097)	(3951)	(4017)	(3977)	(4410)	(4012)	(4465)	
age (days)	503	491	483	505	498	496	488	514	0.15
	(305)	(304)	(306)	(315)	(304)	(310)	(301)	(315)	
assets	47718	48479	47821	48193	48529	48217	48722	48896	0.42
	(18771)	(19148)	(18454)	(18352)	(19031)	(19761)	(19344)	(25464)	
female	0.342	0.333	0.334	0.348	0.320	0.320	0.323	0.348	0.481
	(0.47)	(0.47)	(0.47)	(0.48)	(0.47)	(0.47)	(0.47)	(0.48)	
not professional	0.844	0.826	0.840	0.842	0.832	0.833	0.839	0.834	0.787
	(0.36)	(0.38)	(0.37)	(0.37)	(0.37)	(0.37)	(0.37)	(0.37)	

Notes: Loan size, principal balance and assets are denoted in thousands of Colombian pesos. Female is a dummy variable that take a value of 1 for females. Not professional is dummy for farmers who do not have a college degree. The reported p-values corresponds to an F-test of joint significance of a regression of the control of interest in a vector of dummy variables for each one of the six treatments.

Table 2.3: Baseline Differences Across Treatment Arms

	Past Due	Days Past Due	15 Days Past Due	30 Days Past Due
t0 - control	0.321	5.200	0.164	0.102
t1 - public	0.297	4.754	0.147	0.089
t2 - public & altruism	0.289*	4.609	0.142	0.088
t3 - public & reciprocity	0.298	4.728	0.143	0.092
t4 - public & municipality peer	0.327	5.418	0.164	0.117
t5 - public & activity peer	0.275**	4.625	0.14*	0.092
t6 - future access	0.298	4.816	0.147	0.094

Notes: Each number corresponds to the mean of the outcome of in the corresponding treatment. It was calculated using the 1,200 observations for each treatment. Past Due is dummy for loans that were ever past due in a window of 120 days after the payment installment date. Days Past Due is the maximum number of days past due in the same 120 days window. 15 Days Past due is a dummy for loans that entered into a period of 15 days past due in the first month after the date scheduled for payment. 30 days past due is a dummy for loans that entered into a period of 30 days past due in a window of 60 days after the date of the payment installment. The starts denote the significance of a t-test of difference in means between the mean in treatment t1 (public) and the corresponding treatment. * denotes a p-value smaller than 0.1, ** smaller than 0.05 and *** smaller than 0.01.

Table 2.4: Impact of Information about State-Owned Lender and Future Credit Access

	Past Due	Days Past Due	15 Days Past Due	30 Days Past Due
t1 - public	-0.034*	-0.77*	-0.03**	-0.023*
	(0.02)	(0.39)	(0.02)	(0.012)
t6 - future access	-0.018	-0.53	-0.027*	-0.014
	(0.02)	(0.39)	(0.015)	(0.012)
obs.	3,600	3,600	3,600	3,600
mean control	0.321	5.2	0.164	0.102
r2 adj	0.104	0.165	0.131	0.141
F-test p-value	0.23	0.139	0.091	0.176

Notes: Standard errors in parentheses. Controls included are size of loan, dummy for restructured loans, age of the loan at time of due date (in days), household assets, a dummy for females, a dummy for no college graduate, BAC office fixed effects, and BAC ex-ante credit score. The F-test p-value corresponds to a test of joint significance of the coefficients of the treatment dummies (that is, excluding the controls from the test). *p<0.1; **p<0.05; ***p<0.01.

Table 2.5: Heterogeneity Analysis

	15 Days Past Due						
z_{im}	Amount Outstanding	High School	Municipality Area	Rural Population	Homicides	Hectares with Coca	Subversive Actions
t1 - public	-0.06**	-0.04*	-0.038**	-0.059**	-0.05***	-0.038**	-0.036**
	(0.029)	(0.02)	(0.017)	(0.023)	(0.017)	(0.015)	(0.017)
t6 - future access	-0.0429	-0.0153	-0.0267	-0.0219	-0.0371**	-0.0352**	-0.0263
	(0.029)	(0.02)	(0.017)	(0.023)	(0.017)	(0.015)	(0.017)
z_{im}	0.012**	0.062	0.0087	-0.089	-0.0004	-0.089*	0.0026
	(0.0053)	(0.044)	(0.018)	(0.13)	(0.0009)	(0.051)	(0.0028)
t1 × z_{im}	0.0048	0.071	0.01	0.19	0.0018***	0.1*	0.0017
	(0.0038)	(0.065)	(0.01)	(0.11)	(0.00063)	(0.055)	(0.0021)
t6 × z_{im}	0.0024	-0.11*	-0.0021	-0.039	0.0007	0.076**	-0.0006
	(0.0037)	(0.06)	(0.0097)	(0.11)	(0.00055)	(0.038)	(0.002)
obs	3595	2340	3595	3595	3574	3595	3595
mean control	0.164	0.164	0.164	0.164	0.164	0.164	0.164
r2 adj	0.133	0.137	0.13	0.131	0.132	0.132	0.13

Notes: Standard errors in parentheses. Each column corresponds to a different regression according to the heterogeneity variable used (z_{im}). In all columns, the outcomes is a dummy for loans that entered into a period of 15 days past due in the first month after the date scheduled for payment. Amount Outstanding is the debt outstanding for the loan (in millions of current Colombian pesos). High School is a dummy for loans whose farmer finished high school. Municipality Area is the area of the municipality in thousands of squared kilometers. Rural Population is the number of inhabitants in the municipality rural area in 2012. Homicides is the number of homicides in the municipality in 2010. Hectares with Coca is the number of hectares cultivated with coca crops in 2010. Subversive Actions is the number of confrontations between state military forces and illegal armed groups in 2010. Controls included at the loan level are: loan size, dummy for restructured loans, age of the loan at time of due date (in days), household assets, a dummy for females, a dummy for no college graduate, BAC office fixed effects, and BAC ex-ante credit score. Controls included at the municipality level are: rural population in 2012, total income of the municipality government in 2012.
*p<0.1; **p<0.05; ***p<0.01.

Table 2.6: Effect of Non-monetary Mechanisms

	Past Due	Days Past Due	15 Days Past Due	30 Days Past Due
t1 - public	-0.033*	-0.72*	-0.028**	-0.02*
	(0.018)	(0.38)	(0.014)	(0.01)
t2 - public & altruism	-0.037**	-0.74*	-0.03**	-0.02
	(0.019)	(0.38)	(0.014)	(0.012)
t3 - public & reciprocity	-0.0251	-0.5823	-0.0266*	-0.0109
	(0.018)	(0.37)	(0.014)	(0.012)
t4 - public & municipality peer	0.0052	0.17	-0.0028	0.013
	(0.012)	(0.38)	(0.014)	(0.012)
t5 - public & activity peer	-0.0462**	-0.7009*	-0.0315**	-0.0125
	(0.018)	(0.38)	(0.014)	(0.013)
t6 - future access	-0.026	-0.51	-0.025*	-0.0091
	(0.018)	(0.38)	(0.014)	(0.012)
obs	8,400	8,400	8,400	8,400
mean control	0.321	5.2	0.164	0.102
r2 adj	0.103	0.154	0.134	0.128
F-test p-value	0.05	0.071	0.112	0.086

Notes: Standard errors in parentheses. Controls included are size of loan, dummy for restructured loans, age of the loan at time of due date (in days), household assets, a dummy for females, a dummy for no college graduate, BAC office fixed effects, and BAC ex-ante credit score. The F-test p-value corresponds to a test of joint significance of the coefficients of the treatment dummies (that is, excluding the controls from the test).
*p<0.1; **p<0.05; ***p<0.01.

Table 2.7: Hypothesis Testing: Equality of Coefficients

	Past Due	Days Past Due	15 Days Past Due	30 Days Past Due
t2 - public & altruism	0.823	0.954	0.973	0.804
t3 - public & reciprocity	0.677	0.722	0.896	0.442
t4 - public & municipality peer	0.040	0.018	0.071	0.005
t5 - public & activity peer	0.473	0.969	0.827	0.530
t6 - future access	0.726	0.581	0.828	0.357

Note: p-values of t-test where the null hypothesis is that the effect of the corresponding treatment is equal to the effect of the t1 (public treatment).

2.7 Appendix to Chapter 2

2.7.1 Call Script

Good day sir ____, you are speaking with ____ from the Banco Agrario de Colombia, our call is of a preemptive character to remind you that your obligation will be due the ____ for an approximate value of ____. Mr/Ms (first and last name of the client), do you have anything to manifest regarding this matter?

If the client manifest he can pay: write down date of payment and register as "commitment to pay."

If the client manifest he will have problems with the payment: Direct the client to the closest bank office to review payment alternatives.

[*Insert treatment sentence here*]

Mr/Ms (first and last name of the client), thank you for your time, remember that you spoke with ____ of the Banco Agrario de Colombia, have a nice day/afternoon/night.

2.7.2 Robustness Exercises

This appendix presents robustness results on the effect of the treatments in the different outcomes of repayment. In particular, it replicates tables 2.4) and 2.6 but with the inclusion of a control for the days between the treatment and the day of the payment installment. Not all of the loans in the randomization received the treatment. For example, some loans paid long before the payment was due. For these cases the control variable is not defined and these observations are not included in the estimations.

Table 2.8: Impact of Information about State-Owned Lender and Future Credit Access
(with days between treatment and payment installment)

	Past Due	Days Past Due	15 Days Past Due	30 Days Past Due
t1- public	-0.039*	-0.84**	-0.032**	-0.023*
	(0.02)	(0.411)	(0.016)	(0.013)
t6- future access	-0.032	-0.76*	-0.035**	-0.019
	(0.02)	(0.41)	(0.016)	(0.013)
obs.	3,397	3,397	3,397	3,397
mean control	0.321	5.2	0.164	0.102
r2 adj	0.138	0.188	0.149	0.15
F-test, p-value	0.127	0.08	0.054	0.162

Notes: Standard errors in parentheses. Controls included are size of loan, dummy for restructured loans, age of the loan at time of due date (in days), household assets, a dummy for females, a dummy for no college graduate, BAC office fixed effects, BAC ex-ante credit score, and the number of days between the treatment and day of the payment installment. The F-test p-value corresponds to a test of joint significance of the coefficients of the treatment dummies (that is, excluding the controls from the test).
*p<0.1; **p<0.05; ***p<0.01.

Table 2.9: Effect of Non-monetary Mechanisms
(with days between treatment and payment installment)

	Past Due	Days Past Due	15 Days Past Due	30 Days Past Due
t1 - public	-0.037**	-0.81**	-0.031**	-0.022*
	(0.019)	(0.39)	(0.015)	(0.012)
t2 - public & altruism	-0.047**	-0.88**	-0.032**	-0.019
	(0.019)	(0.39)	(0.014)	(0.012)
t3 - public & reciprocity	-0.036*	-0.69*	-0.03**	-0.012
	(0.019)	(0.39)	(0.015)	(0.012)
t4 - public & municipality peer	-0.0077	-0.0087	-0.0085	0.011
	(0.019)	(0.39)	(0.015)	(0.012)
t5 - public & activity peer	-0.05***	-0.73*	-0.033**	-0.013
	(0.019)	(0.39)	(0.015)	(0.012)
t6 - future access	-0.038**	-0.70*	-0.03**	-0.013
	(0.01903)	(0.39)	(0.015)	(0.012)
obs.	7,920	7,920	7,920	7,920
mean control	0.321	5.2	0.164	0.102
r2 adj	0.137	0.176	0.149	0.136
F-test p-value	0.055	0.089	0.13	0.117

Notes: Standard errors in parentheses. Controls included are size of loan, dummy for restructured loans, age of the loan at time of due date (in days), household assets, a dummy for females, a dummy for no college graduate, BAC office fixed effects, BAC ex-ante credit score, and the number of days between the treatment and day of the payment installment. The F-test p-value corresponds to a test of joint significance of the coefficients of the treatment dummies (that is, excluding the controls from the test).
*p<0.1; **p<0.05; ***p<0.01.

Chapter 3

Is it my Money or Not? An Experiment on Risk Aversion and the House-Money Effect

Juan Camilo Cárdenas, Nicolás de Roux, Christian Jaramillo and Luis Roberto Martínez†

Published as Cárdenas, De Roux, Jaramillo, and Martínez (2014)

† We thank Glenn Harrison who commented on a previous version and greatly enriched the analysis. Also we thank two anonymous reviewers and the editor of *Experimental Economics* for their comments.

3.1 Introduction

The house-money effect, understood as people's tendency to be more daring with easily-gotten money, is a behavioral pattern which finds support in incentivized experiments using real money by Thaler and Johnson (1990). Since experiments in economics usually start by handing out money to the subjects so that they never stand to suffer any net monetary losses, the participants' behavior could be modified as a result of the house-money effect. This poses questions about the external validity of experiments in economics (Guala, 2005, p. 231), and particular questions about the incentives used: to what extent do people behave in the experiment like they would have in a real-life situation, given that they play with easily-gotten house money (Levitt and List, 2007)?

The experimental literature has addressed this question in the context of altruism (Cherry, Frykblom, and Shogren, 2002), public goods (Clark, 2002), auctions (Ackert, Charupat, Church, and Deaves, 2006) and capital expenditure (Keasey and Moon, 1996). The general idea of windfall gains has been also explored in the psychology and economics literature (Arkes, Joyner, Pezzo, Nash, Siegel-Jacobs, and Stone, 1994; Keeler, James, and Abdel-Ghany, 1985). Most of these papers deal with the issues arising from having people play with their own money by having participants earn money in an initial stage and then making choices with their earnings.

This paper studies the effect of house money on the risk preferences of a group of 122 undergraduate students within an age range of 16 to 28. The students were randomly assigned to a control or a treatment group and given money to participate in the experiment, which they were told involved risky choices and possibly losses. As usual, the money handed out for participating was enough to cover the potential losses. However, while the control group received this initial money just before they made their choices, the treatment group received the money three weeks in advance so that they had time to spend it before making their choices. (A back-of-the-envelope calculation suggests that on average 35% of the cash in advance was spent.) This experimental design, inspired by (Bosch-Domènech and Silvestre, 2010), is as close as we can get to having them gamble with their own money.

We find evidence of an indirect house money effect operating through the money that participants had with them at the time of choosing between lotteries. More specifically, we find that for the treatment group, each additional thousand Colombian pesos (COP) spent (USD$0.50) leads to an increase of 0.0019 in their CRRA risk aversion coefficient. We interpret this finding as evidence of a house money effect on those subjects of the treatment group who actually spent some of the cash provided to them in advance. In our preferred specification, the mean relative risk aversion coefficient equals 0.34 with a standard deviation of 0.09. Therefore, our estimated 35% expenditure of the endowment would lead to a reduction of 0.3 standard deviations in the risk aversion coefficient. This interpretation rests on two assumptions. First, that the money participants had with them at the time of the experiment is a good proxy for endowment not spent, if compared to the same measure in the control group. Second, and more importantly, we assume that the house money effect only operates for those people who actually spent some of their endowment. We will have more to say about this assumption below.

The results that we report here add to a vast literature documenting risk aversion and suggest that it would be advisable to include credible controls for the house-money effect in experimental work in economics.[1]

---

[1]Although only partially comparable, empirical evidence from survey work from a developing country suggests risk aversion coefficients between 0 and 5 (Azam, Biais, and Kamionka, 2002). Meanwhile, a survey on experimental studies in developed and developing countries reports estimated coefficients for the CRRA that range from the lowest estimate of 0.05 in Ethiopia to 2.57 in Paraguay (Cárdenas and Carpenter, 2008); Harrison, Humphrey, and Verschoor (2010) also report coefficients of these magnitudes. However, Harrison and Rutström (2008) use a method quite similar to ours on the data of Hey and Orme (1994) and find a CRRA of 0.66 with a standard error of 0.04. In general, the estimated ranges found show also a non-negligible sensitivity to the type of experimental procedure used.

3.2 The Experiment

Our experimental design is based on dividing the subject pool randomly in half and giving the treatment group an endowment of cash three weeks in advance of the actual decision-making experiment. The control group receives the same amount of money but on the day of the experiment as is usually done in lab experiments that involve potential losses. With that time period in between we expect to balance between giving sufficient time for them to incorporate the cash as part of their pocket money and not allowing for some discounting of the endowment between the treatment and control groups.

The subjects were volunteers from an undergraduate psychology course at the Universidad de los Andes in Bogotá (Colombia), recruited in two different semesters of the same course (one in 2009 and another in 2012). The students in the class were randomly assigned to a treatment (cash in advance) or control group (cash experiment day) and then asked to consent to participate in an economic experiment that involved risky choices. Of a total of 122 students who accepted to participate in the two sessions, 61 were assigned to the control group and 61 to the treatment group. Within each session the random splits between treatment and controls were 51/49 and 48/52. Students in the sessions belonged to more than 20 different minors and majors from social, medical, natural sciences, medicine and engineering (no more than 14% of the participants in any of the groups belonged to any particular major). Table 3.1 shows the average characteristics of each group.

Treatment subjects were then given COP 40,000 in small change (roughly USD 20 given an exchange rate of COP 1,971 on the initial day of the first round. The minimum monthly wage in Colombia at the time was COP 497,000). Three weeks later, again in class, the decision-making session took place. The control group was given their respective COP 40,000 and everybody proceeded then to make their choices under uncertainty.

Notice in particular the averages for the available pocket money of the subjects in the treatment and control groups. We asked everyone at the entrance to the room and before the control group received their endowment, how much money they had in their pockets. Our treatment group had significantly more cash in their pockets than the control, as expected, but the difference was smaller than the endowment of 40,000 COP (67,000 COP — 41,000 COP $\approx$ 26,000 COP).[2] If we assume that the money brought to the session by the control group is representative of what members of this student community carry in their pockets, we can think that the treatment group spent on average 14,000 COP (or 35%) of their cash in advance. For those in the control group, we added the 40,000 COP to their pocket money and therefore we have now a comparable variable, Adjusted pocket money, which will turn out to be an important part of our analysis.[3]

Following Binswanger (1980) Ordered Lottery Selection (OLS) design (Harrison and Rutström, 2008, pp. 52-56), all participants were handed a piece of paper with six different uniform-probability lotteries involving possible losses (Figure 3.1a) depending on a coin toss. They were then asked to choose one lottery to play. All 122 made their choice at once. At that point they did not know they would have further choices to make.

After collecting their choices, they were handed a second set of six lotteries (Figure 3.1b). None of these involved losses and they were told that the outcome would depend on another coin toss and that their payments would be computed using the sum of results of both lotteries. After collecting their new choices, they were asked to fill out a brief socioeconomic survey. Only then did both coin tosses take place. The first coin toss determined the outcome

[2] A word of caution is due at this point. As suggested by an anonymous referee, in societies where students pay much of their expenses using debit or credit cards, the question "how much money did you have in your pocket when you entered the class room" might be blurred. We are confident, however, that this should not be of concern as Colombian students rarely use electronic payments for their daily expenses in food, transportation or entertainment, among others because most establishments have a minimum amount for allowing such transactions, and the access to banking in general is more limited than in industrialized countries; also, not all establishments take electronic payments around and on campus.

[3] We thank two reviewers for this suggestion.

in the first game (possible losses) and the second coin toss determined the outcome for the gains only lottery.

3.3 Results

Figure 3.2 shows the distribution of choices in two different ways. In Panels (a) and (b) we show the distributions of choices for Game 1 and Game 2 respectively. In panels (c) and (d) we compare the same data but splitting in control and treatment respectively. From a first look at the distributions one can infer that prospect theory is alive and well and that in general people made riskier choices in Game 1 where losses were possible. However, there seems to be no major difference between the treatment and control groups and therefore a more rigorous statistical analysis is needed.

These results can be compared to data from a more comprehensive study (Cárdenas and Carpenter, 2013) that included more than 3,000 subjects representative of several Latin American cities using this same design of the potential losses and gains with these six lotteries. In the case of the lottery with potential losses, the variation found in that large sample is higher with more people choosing the more conservative lotteries than here and a smaller fraction (17.2%) of subjects in that sample choosing the riskiest lottery E than in our students sample (23.8%). In the case of the second game with gains only, again our students showed a slightly higher level of risk tolerance with more students choosing lottery E than in the adults sample and fewer students choosing the safe ones.

Table 3.2 shows the means of choices in each game for the treatment (cash in advance) and control (cash experiment day) groups. Game 1 and Game 2 indicate the choice in each game. In both games, lotteries A through F of Figure 3.1 are coded 1 through 6: Game 1 = 1 means the subject chose lottery A in Game 1, and a larger value indicates the choice of a riskier lottery. As expected through prospect theory, the average player moved from riskier lotteries in Game 1 to safer ones in Game 2 creating two different distributions of choices when comparing within subjects (Wilcoxon signed-rank test, p-value=0.0001). However, this result seems driven mostly by the control group and rather minor for the treatment group both for men and women. This could mean that, if there was a house money effect operating, it could be leading to attenuation in loss aversion. Within games we only find a significant difference between treatment and control groups when comparing the choices of men in Game 1 (see Table 3.2). That difference vanishes for Game 2, suggesting that the effect is exacerbated when involving the possibility of losses.

These differences however open up more questions than answers. To assess in more detail the effect of being treated on risk aversion, we estimate for each game a series of structural models of choice under uncertainty, using the survey data as explanatory variables and following closely Harrison and Rutström (2008, pp. 69-74).

In these exercises, each subject i is assumed to have a CRRA utility function

$$u_i(I) = \frac{I^{1-\gamma_i}}{1-\gamma_i} \tag{3.1}$$

where I denotes total wealth and γ_i is the relative risk aversion parameter of that individual - a higher γ_i is associated with a lower level of risk tolerance, $\gamma_i < 0$ corresponds to risk loving, $\gamma_i > 0$ to risk aversion and $\gamma_i = 0$ to risk neutrality.[4]

For both games 1 and 2, Let $\mathrm{EU}_i(j)$ denote the expected utility for subject i of choosing lottery j in that game, where $j \in \{A, B, C, D, E, F\}$ according to Figure 3.1. Let L_j and R_j denote the payoff if the left or right outcomes of lottery j are realized. The expected utility of choosing this lottery, with the CRRA utility function and a probability 1/2 for each

[4]Total wealth I is defined as initial wealth (w) plus the payoff of the realized outcome of the game. For Game 1 we set w=40,000 so that there is no negative total wealth I in any of the outcomes (note that the utility function is well defined for non negative values of I). This assumption is grounded on the fact that all individuals were given an initial endowment of 40,000, the only difference being one of timing. For Game 2, we set w=0.

outcome, is given by:

$$EU_{i,j} = \frac{1}{2}\frac{(L_j + w)^{1-\gamma_i}}{1-\gamma_i} + \frac{1}{2}\frac{(R_j + w)^{1-\gamma_i}}{1-\gamma_i} \tag{3.2}$$

Using this formula for the expected utility, for each individual we compute a probability of observing the choice the individual actually made. In order to do so and following Harrison and Rutström (2008), we assume a multinomial logit probability specification. Let $h \in \{A, B, C, D, E, F\}$ denote the lottery the individual actually chose. The probability of individual i choosing lottery h is given by:

$$P_i(h) = \frac{e^{\mathrm{EU}_i(h)}}{\sum_j e^{\mathrm{EU}_i(j)}} \tag{3.3}$$

where again $j \in \{A, B, C, D, E, F\}$.[5]

We further assume that the risk aversion coefficient γ_i is a linear function of observed characteristics X_i, i.e. $\gamma_i = \alpha + X_i \cdot \beta$, where α is a constant and β is a vector of size $k \times 1$, k being the number of variables included in the model. Our objective is to estimate the values of α and β. The maximum likelihood (MLE) routine that we implement finds the values of α and β that maximize the following log likelihood function (i.e. that maximize the probability of observing our sample of choices assuming a multinomial logit probability specification):

$$\ln L = \sum_i \ln(P_i(h_i)) \tag{3.4}$$

Note that once the estimated values $\hat{\alpha}$ and $\hat{\beta}$ are obtained, we can use the characteristics

[5]This statistical assumption implies a possibility of decision error, since an individual may not choose with certainty a lottery that has a higher expected utility than all the others. For example, among lotteries with expected payoffs [1, 1, 1, 1, 1, 10], a risk neutral person will choose the one with expected payoff equal to 10 only with probability 2/3, even though the expected utility of this lottery is higher than that of all others. We thank an anonymous referee for pointing this out and suggesting the above example.

of the individual i, namely X_i, to obtain a linear prediction of γ_i, $\hat{\gamma_i} = \hat{\alpha} + X_i \cdot \hat{\beta}$. The value of $\hat{\gamma_i}$ will depend on the model being estimated (i.e. on the individual characteristics that we include in the linear function of γ_i).

We estimate four different specifications of the structural model for each one of the games. Tables 3.3 and 3.4 report the results for Game 1 and Game 2 respectively.[6] Each column corresponds to one specification. We report the estimated coefficients and their respective standard errors. For example, column (1) of Table 3.3 corresponds to a linear specification of γ_i given by $\gamma_i = \alpha + \beta_1 \times \text{Treatment} + \beta_2 \times \text{Session}$, where Session is a dummy that takes a value of 0 if the experimental session is the one conducted in 2009, and a value of 1 if it is the one conducted in 2012.

Consider Game 1 which involves the possibility of losses (Table 3.3). We confirm our previous finding that males are more tolerant to taking risks than females, and find that people of higher socio-economic status measured by the variable Stratum also choose riskier lotteries.

We do not find that the treatment in itself has an effect on the risk coefficient of the subjects (across columns the coefficient of Treatment is not statistically significant). However, as illustrated by column (4) the interaction between the available pocket money at the start of the experiment and the treatment does tell a story: the less money an individual in the treatment group had in her pocket, the more conservative her decision was. No such effect is found for the controls.[7]

Our interpretation of this result rests on the following assumption: it is necessary that

---

[6] All the monetary variables enter the estimations in thousands of COP.

[7] To see this, denote by π_1 the estimated coefficient of Pocket Money (adj) and by π_2 that of Pocket Money(adj)*Treatment. For the cash-in-advance treatment, an increase of one thousand COP in Pocket Money (adj) implies a change in γ of $\pi_1 + \pi_2$. For the control group, it implies a change in γ of π_1 (recall that the Treatment dummy takes a value of 1 for the cash-in-advance treatment). Nevertheless, since π_1 is not statistically different from 0, the effect of Pocket money (adj) for the control group is $\pi_1 = 0$, and that of the cash-in-advance treatment is $\pi_1 + \pi_2 = \pi_2 = -0.0019$.

subjects in the treatment group actually spent part of the endowment for them to consider that they are actually playing with their own money. In other words, receiving money in advance is not a sufficient condition for the house money effect to operate. It could be the case that the treatment subjects who kept their money over this time felt an obligation to bring the money to the decision stage session, but as a reviewer noted, we would need to ask directly the participants about their reasons for spending or keeping their money over the period of time. This assumption implies that in the extreme case of a participant in the treatment group who did not spend any of his endowment before the decision-making session, we should not observe any difference in his behaviour relative to the control group.

If we additionally assume that the money that the controls brought into the room is a good proxy for the average pocket money in the population, then the difference in means of Table 3.1 of the variable Money in Pocket (67,000 COP - 41,000 COP = 26,000 COP) gives us an idea of how much the treatment subjects actually spent on average of their cash in advance (approximately 40,000 COP – 26,000 COP = 14,000 COP, which is 35% of the cash in advance). This already suggests that any house money effect found should not be large. Under the assumptions just mentioned, the coefficient of the interaction between Treatment and Pocket money (adj) can be seen as the effect of cash in advance that is actually spent. The higher the amount of the endowment spent by the subjects in the treatment (i.e. the more they are "playing with their own money"), the lower the adjusted pocket money will be. Since the coefficient is negative, this implies a higher estimated value of γ, or in other words, a higher level of risk aversion as participants in the treatment group played with more of their own money.

Phrased differently, we can think of the cash in advance actually spent as having some distribution across individuals in the treatment group, where some of them spent all the endowment, some of them did not spend at all and on average they spent 14,000 COP. For those that spent all the money we would expect more risk averse behavior during the decision stage of the experiment when compared to the controls. For those that did not spend any amount we would expect no difference with the controls. To put numbers to this interpretation, the linear prediction of γ_i using the model of column (4) in Table 3.3 implies an average estimated γ_i of 0.34 with a standard deviation of 0.09.[8] If the participants in the cash-in-advance treatment had spent all the endowment they would have values of γ_i larger on average than those of the controls by an amount of (-0.0019)*(-40) = 0.076 which is almost one standard deviation.[9] From our rough approximation of the average money that was actually spent by participants in the cash-in-advance treatment, i.e. 14,000 COP, we can infer that their γ_i is on average greater by an amount of only (-0.0019)*(-14) = 0.026 which is approximately 0.3 standard deviations. We can summarize our finding by saying that the evidence suggests a small house money effect driven by the fact that members of the treatment group spent less than half of the cash in advance provided. Further, the fact that the effect does not happen among the control group rules out the explanation of more risk aversion caused by diminishing marginal utility of money.[10]

Let us now turn to the analysis of the second game, our control for risk under uncertainty but with no potential losses involved. Although we already reported that most individuals did switch from riskier to more conservative choices, the results in Table 3.4 show similar patterns to those reported for Game 1. Males, although not significantly now, show less risk aversion, those with higher socio-economic levels (expenses) also show more tolerance to risk and once again the available pocket money makes a difference but only for the treatment

[8]Harrison and Rutström (2008, pp. 69-71) assume the same utility function and apply a similar version of the MLE procedure we employ to the data of Hey and Orme (1994). In one of the versions of this exercise they find a pooled value of γ of 0.66 with a standard error of 0.04.

[9]As mentioned in Table 3.1, Pocket money (adj) is defined as money at time of play (i.e. the money brought in plus 40,000 COP for the control group). If subjects in the treatment group spent all their cash in advance, money brought in would have been equal for the two groups and the adjustment would leave the treatments at -40,000 COP.

[10]We thank Reviewer 2 for highlighting this.

group and in the same direction as before.

A few notes are worth mentioning here. Recall that Game 2 took place after the students had made their choice for Game 1 but before the toss of the coin for Game 1 was made. Also, they were not told in Game 1 that a second game was going to be played later on. One could argue that the choice made for the second game involved some kind of risk hedging between games since they did not know the outcome of the coin toss. To control for this possibility we ran a separate regression not reported here where the choice in Game 1 was used as a control for the choice in Game 2 and no effect was found.

3.4 Conclusions

The use of monetary incentives is central to the experimental methods in economics. The code of ethics among experimentalists continues to suggest that we refrain from using the disposable income of our experimental subjects, that is, from having participants walk out of the lab with negative earnings and instead it requires that we provide them with an endowment they can use to allow for decisions involving losses. This has caused concern among skeptics of experiments because of the so called house-money effect and the implications it would have for external validity of laboratory or field lab experiments.

To get at this debate, some labs have introduced the notion that the endowment is earned during a task performed at the experiment, partially correcting for the problem of subjects thinking of the endowment as a windfall gain. We have, however, taken a different approach, by giving the endowment well in advance (21 days) to half of our sample and the endowment to the other half at the day of the experiment. Further, they had to make decisions about risk involving losses and gains. We asked everyone at the day of the experiment what cash they had available in their pockets and confirmed that the treatment group had in fact spent part of the endowment they had received and kept another part, suggesting the money was incorporated as part of their disposable income. On average there is no major statistical difference in the distributions of the observed coefficient of risk behavior across the two groups. However, when controlling for the available cash they had in their pockets at the time of the experiment, we find that those in the treatment group who had more money with them on the day of the experiment tended to be more risk tolerant while those who had less were more risk averse during the experiment. If we interpret the spending of the endowed money as a signal of considering it as one's own, our findings suggest a small house money effect.

By providing the endowment in advance we have both complied with the ethical code of experimental economics but also introduced more realism and external validity as the subjects seem to have incorporated some of the mental accounting processes of their daily life into the experiment. In other words, the more I spent part of my endowment the more it felt like "it's my money". The data suggest that those who spent more of their endowment arrived facing the experiment much like a risky decision involving losses but constrained by their pocket money whereas those with more cash –provided by the experimenter, felt like taking riskier decisions, in other words "it's not my money".

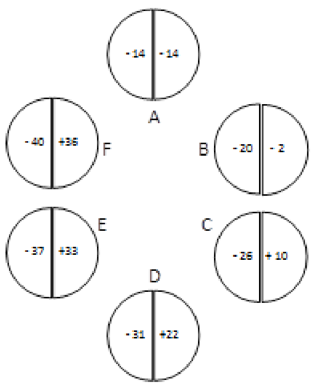
Experiments that involve studying strategic behavior with possible losses should take into account that when subjects receive an endowment they might not treat it as part of their real income. Our results would suggest that experimenters could control for available cash in the pockets of the subjects at the time of the experiment, even if the experiment provides an endowment to cover for losses as this would help explain variation in behavior. These factors should be tested, using a similar design of giving an endowment to subjects well in advance, but for other domains of interactions such as fairness and bargaining games (ultimatum, dictator, Coase bargaining), cooperation games (trust an public goods) or labor relations (effort, gift exchange) where a subject must decide over the allocation of her own resources and test for robustness and potential house-money effects.

Our research opens other new questions for further experimental tests on decisions under uncertainty.[11] We are well aware that this design is based on the same probability of 0.50 over all possible lotteries and this might impose a strong assumption about the application of expected utility theory, although it minimizes the potential problems of humans handling probabilities (Kahneman, Slovic, and Tversky, 1982). Nevertheless, further tests with variable probabilities would enrich this finding, using other risk experiments available. On the one hand we could estimate this effect in other samples with different demographics including age, education level, financial literacy or income. On the other hand one could explore how the magnitude of the house-money bias changes with the time delay between the transfer of the endowment and the experimental decision. These could all deepen our understanding of how incentives work in the laboratory and of how income shocks may interact with behavior under uncertainty.

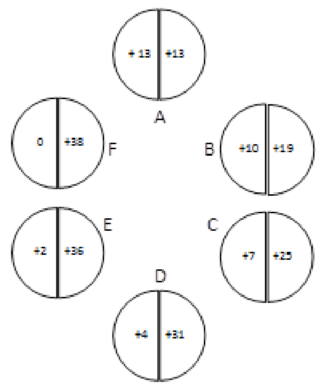
[11]A natural test of our findings could be conducted with occasional tourist casino players. Imagine a random group of tourists that receive a voucher-like gift in cash well in advance before their visit to the casino and another group that receives the voucher in the day of the visit. If our hypothesis holds, the latter group would make riskier decisions in the casino.

3.5 Figures Chapter 3

Figure 3.1: Games and payoffs



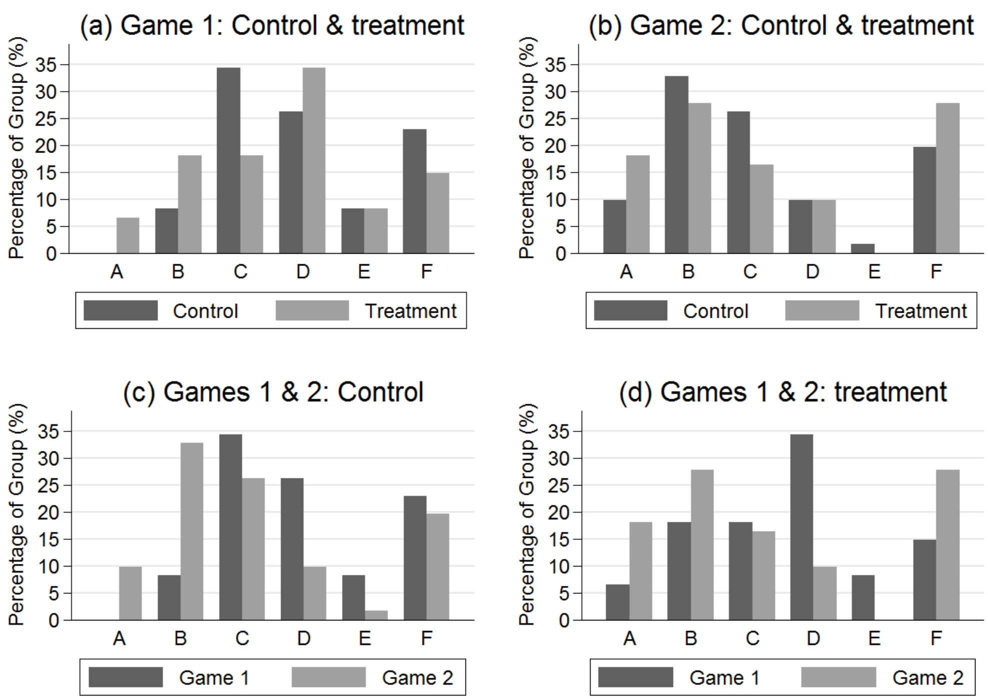
(a) Game 1: Possible losses



(b) Game 2: Gains only

Notes: In each game, both risk and expected return increase clockwise from the top. However, lotteries E and F have the same expected return. Payoffs in thousands of Colombian Pesos (COP) (Exchange rate (USD): COP 1,971).

Figure 3.2: Decisions in Game 1 and Game 2



3.6 Tables Chapter 3

Table 3.1: Demographic Characteristics of Treatment and Control Groups

	Mean Value		p-value	
	Control	Treatment	Rank-sum test	t-test
Female	0.361	0.574	0.019	0.018
Age	19.6	18.8	0.170	0.034
Single	0.967	1.00	0.156	0.156
Siblings	1.4	1.5	0.405	0.291
Semesters at university	3.3	3.0	0.261	0.371
Monthly expenses	445,080	443,440	0.840	0.968
Housing stratum	4.72	4.8	0.898	0.671
Money in pocket	41,132	67,098	0.000	0.001
Adj. Money in pocket	81,132	67,098	0.014	0.057

Notes: All money variables in Colombian pesos (COP) (1) Only two participants (both in the control group) reported "other" as marital status. (2) Using mid-point of reported range. (3) Housing strata in Colombia range from 1 (lowest) to 6 (highest). (4) Amount of money at time of making decisions (pocket + 40,000 for participants in control group).

Table 3.2: Experimental Results by Treatment and Gender

	All		
	Treatment	Control	p-value
Game 1 (losses)	3.64	4.03	0.206
Game 2 (gains)	3.30	3.20	0.956
p-value	0.130	0.000***	-
	Male		
	Treatment	Control	p-value
Game 1 (losses)	3.62	4.31	0.095*
Game 2 (gains)	3.27	3.31	0.929
p-value	0.360	0.004***	-
	Female		
	Treatment	Control	p-value
Game 1 (losses)	3.66	3.55	0.659
Game 2 (gains)	3.31	3.00	0.886
p-value	0.242	0.059*	-

Notes: Lotteries are coded 1-6 for each game, with a higher number indicating a riskier choice. p-values from a Rank-sum (Mann-Whitney) test for differences in distributions. *** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Maximum Likelihood Estimation of γ (Game 1)

	(1)	(2)	(3)	(4)
Treatment	0.070	-0.028	-0.023	0.138
	(0.045)	(0.078)	(0.080)	(0.120)
D(Session 2)	0.019	0.042	0.034	0.038
	(0.044)	(0.048)	(0.048)	(0.05)
D(Male)		-0.13*	-0.16**	-0.15**
		(0.070)	(0.072)	(0.070)
D(Male)×Treatment		0.164	0.149	0.131
		(0.107)	(0.108)	(0.109)
Expenses			-0.0003	-0.006
			(0.009)	(0.011)
Stratum			-0.051*	-0.05*
			(0.026)	(0.030)
Pocket Money (adj.)				0.0012
				(0.0009)
Pocket Money (adj.)×Treatment				-0.0019*
				(0.0011)
Constant	0.29***	0.37***	0.63***	0.54***
	(0.034)	(0.063)	(0.15)	(0.16)
Observations	122	122	122	122

Notes: Standard errors in parentheses. Treatment equals one for participants who received cash in advance. *** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Maximum Likelihood Estimation of γ (Game 2)

	(1)	(2)	(3)	(4)
Treatment	-0.038	-0.129	-0.223*	0.194
	(0.075)	(0.125)	(0.121)	(0.223)
D(Session 2)	0.090	0.104	0.118	0.106
	(0.077)	(0.079)	(0.076)	(0.086)
D(Male)		-0.097	-0.104	-0.098
		(0.125)	(0.125)	(0.128)
D(Male)×Treatment		0.159	0.213	0.168
		(0.164)	(0.161)	(0.177)
Expenses			-0.057***	-0.069***
			(0.0164)	(0.0180)
Stratum			-0.027	-0.037
			(0.0446)	(0.041)
Pocket Money (adj.)				0.003*
				(0.002)
Pocket Money (adj.)×Treatment				-0.005**
				(0.002)
Constant	0.41***	0.47***	0.94***	0.82***
	(0.063)	(0.111)	(0.251)	(0.254)
Observations	122	122	122	122

Notes: Standard errors in parentheses. Treatment equals one for participants who received cash in advance. *** p<0.01, ** p<0.05, * p<0.1

Bibliography

Ackert, L. F., N. Charupat, B. K. Church, and R. Deaves (2006): "An Experimental Examination of the House Money Effect in a Multi-Period Setting," *Experimental Economics*, 9(1), 5–16.

Acquah, H., and J. Addo (2011): "Determinants of loan repayment performance of fishermen: Empirical evidence from Ghana," *Cercetari Agronomice in Moldova*, 44(4).

Ahlin, C., and R. M. Townsend (2007): "Using repayment data to test across models of joint liability lending," *The Economic Journal*, 117(517), F11–F51.

Andreoni, J. (1990): "Impure altruism and donations to public goods: A theory of warm-glow giving," *The Economic Journal*, 100(401), 464–477.

Angrist, J. D., and J.-S. Pischke (2008): *Mostly Harmless Econometrics: An empiricist's Companion*. Princeton University Press.

Arcila, J. P., J. Farfán, A. B. Moreno, L. F. Salazar, and E. Hincapié (2007): "Sistemas de Producción de Café en Colombia," Discussion paper.

Arkes, H. R., C. A. Joyner, M. V. Pezzo, J. G. Nash, K. Siegel-Jacobs, and E. Stone (1994): "The Psychology of Windfall Gains," *Organizational Behavior and Human Decision Processes*, 59(3), 331 – 347.

Avery, R. B., P. S. Calem, and G. B. Canner (2004): "Consumer Credit Scoring: Do Situational Circumstances Matter," *Journal of Banking and Finance*, 28, 835–856.

Azam, J. P., B. Biais, and T. Kamionka (2002): "Risk Aversion and Entrepreneurship in Cote d'Ivoire: A Structural Econometric Approach," Manuscript.

Ball, L. M. (2014): "Long-Term Damage from the Great Recession in OECD Countries," NBER working paper no. 20185.

Bardhan, P., and C. Udry (1999): *Development Microeconomics*. Oxfor University Press.

Barham, B. L., S. Boucher, and M. R. Carter (1996): "Credit constraints, credit unions, and small-scale producers in Guatemala," *World Development*, 24(5), 793–806.

Barrot, J.-N., and J. Sauvagnat (2016): "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks," *The Quarterly Journal of Economics*, 131(3), 1543–1592.

Benabou, R., and J. Tirole (2006): "Incentives and Prosocial Behavior," *American Economic Review*, 96(5), 1652–1678.

Berger, A. N., W. S. Frame, and N. H. Miller (2005): "Credit Scoring and the Availability, Price, and Risk of Small Business Credit," *Journal of Money, Credit and Banking*, pp. 191–222.

Best, M. C., A. Brockmeyer, H. Kleven, J. Spinnewijn, and M. Waseem (2015): "Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence from Pakistan," *Journal of Political Economy*, 136(6), 1311–1355.

Binswanger, H. P. (1980): "Attitudes Toward Risk: Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, 62(3), 395–407.

Blanchard, O. J., and L. H. Summers (1986): "Hysteresis and the European Unemployment Problem," NBER working paper no. 1950.

Bordalo, P., N. Gennaioli, and A. Shleifer (2015): "Memory, Attention and Choice," Working Paper.

Bosch-Domènech, A., and J. Silvestre (2010): "Averting Risk in the Face of Large Losses: Bernoulli vs. Tversky and Kahneman," *Economics Letters*, 107(2), 180 – 182.

Boucher, S., and A. Moya (2015): "Gestión de Riesgo: Recomendaciones para el Diseño e Implementación de un Seguro Indexado para Pequeños Productores Cafeteros," in *Misión de Estudios para la Competitividad de la Caficultura en Colombia*, ed. by J. J. Echavarría, P. Esguerra, D. McAllister, and C. F. Robayo. Fondo Nacional del Café, Bogotá.

Breza, E. (2015): "Peer Effects and Loan Repayment: Evidence from the Krishna Default Crisis," Working Paper.

Bursztyn, L., S. Fiorin, D. Gottlieb, and M. Kanz (2016): "Moral incentives in credit card debt repayment: evidence from a field experiment," World Bank working paper.

Cadena, X., and C. Quintero (2015): *Evolución de los Servicios Financieros en Colombia, 2010 - 2013*. Universidad de los Andes - CEDE.

Cadena, X., and A. Schoar (2011): "Remembering to pay? Reminders vs. financial incentives for loan payments," Working Paper No. 17020, National Bureau of Economic Research.

Carpena, F., S. Cole, J. Shapiro, and B. Zia (2012): "Liability structure in small-scale finance: evidence from a natural experiment," *The World Bank Economic Review*.

Carter, M. R. (1988): "Equilibrium credit rationing of small farm agriculture," *Journal of Development Economics*, 28(1), 83–103.

Carter, M. R., and P. Olinto (2003): "Getting Institutions "right" for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment," *American Journal of Agricultural Economics*, 85(1), 173–186.

Castro, C., and K. Garcia (2014): "Default Risk in Agricultural Lending, the Effects of commodity price volatility and climate," *Agricultural Finance Review*, 74(4), 501–521.

Chaudhuri, A., T. Paichayontvijit, et al. (2006): "Conditional cooperation and voluntary contributions to a public good," *Economics Bulletin*, 3(8), 1–14.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002): "Hardnose the Dictator," *American Economic Review*, 92(4), 1218–1221.

Chetty, R., A. Looney, and K. Kroft (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99(4), 1145–1177.

Chirwa, E. W. (1997): "An Econometric Analysis of the Determinants of Agricultural Credit Repayment in Malawi," *African Review of Money Finance and Banking*, 2(1), 107–122.

Clark, J. (2002): "House Money Effects in Public Good Experiments," *Experimental Economics*, 5(3), 223–231.

Conning, J., and C. Udry (2007): "Rural Financial Markets in Developing Countries," *Handbook of Agricultural Economics*, 3, 2857–2908.

Cárdenas, J. C., and J. Carpenter (2008): "Behavioural Development Economics: Lessons from Field Labs in the Developing World," *The Journal of Development Studies*, 44(3), 311–338.

——— (2013): "Risk Attitudes and Economic Well-Being in Latin America," *Journal of Development Economics*, 103, 52 – 61.

Cárdenas, J. C., N. De Roux, C. Jaramillo, and L. R. Martínez (2014): "Is It my Money or Not? An Experiment on Risk-Aversion and the House-Money Effect," *Experimental Economics*, 17(1), 47–60.

Croson, R. T. (2007): "Theories of commitment, altruism and reciprocity: Evidence from linear public goods games," *Economic Inquiry*, 45(2), 199–216.

Dal Bó, E., and P. Dal Bó (2014): ""Do the right thing:" The effects of moral suasion on cooperation," *Journal of Public Economics*, 117, 28–38.

De Aghion, B. A. (1999): "Development banking," *Journal of Development Economics*, 58(1), 83–100.

de Janvry, A., C. McIntosh, and E. Sadoulet (2010): "The Supply-and Semand-Side Impacts of Credit Market Information," *Journal of development Economics*, 93(2), 173–188.

de Luna-Martínez, J. d., and C. L. Vicente (2012): "Global Survey of Development Banks," World Bank, Policy Research Working Paper, 5969.

de Olloqui, F. e. (2013): *Bancos Públicos de Desarrollo, Hacia un Nuevo Paradigma?* Banco Interamericano de Desarrollo.

Dell, M., B. F. Jones, and B. A. Olken (2014): "What do we Learn from the Weather? The New Climate Economy Literature," *Journal of Economic Literature*, 52(3), 740–798.

DNP, and FINAGRO (2014): "Sistema Nacional de Crédito Agropecuario, Propuesta de Reforma," .

Dobbie, W., P. Goldsmith-Pinkham, N. Mahoney, and J. Song (2016): "Bad Credit, No Problem? Credit and Labor Market Consequences of Bad Credit Reports," NBER working paper no. 22711.

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2006): "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany," *American Economic Journal: Economic Policy*, 8(3), 203–232.

Einav, L., M. Jenkins, and J. Levin (2012): "Contract Pricing in Consumer Credit Markets," *Econometrica*, 80(4), 1387–1432.

——— (2013): "The Impact of Credit Scoring on Consumer Lending," *The RAND Journal of Economics*, 44(2), 249–274.

Escobal, J. A., and P. Mediano (2006): "Cómo lograr que la titulación de predios sirva para dinamizar la actividad agropecuaria en el sector rural," *CIES, Lima, Perú*.

Eslava, M., A. J. Galindo, M. Hofstetter, and A. Izquierdo (2010): "Scarring Recessions and Credit Constraints: Evidence from Colombian Plant Dynamics," Documento CEDE, 2010-27.

Fafchamps, M. (2013): "Credit Constraints, Collateral, and Lending to the Poor," *Revue d'économie du développement*, 21(2), 79–100.

FAO (2013): *Four case studies on credit guarantee funds for agriculture*. Food and Agriculture Organization of the United Nations (FAO). Rome.

Feder, G., T. Onchan, and T. Raparla (1988): "Collateral, guaranties and rural credit in developing countries: evidence from Asia," *Agricultural Economics*, 2(3), 231–245.

Feigenberg, B., E. M. Field, and R. Pande (2010): "Building social capital through microfinance," Working Paper No. 16018, National Bureau of Economic Research.

Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 71(3), 397–404.

Fletschner, D., C. Guirkinger, and S. Boucher (2010): "Risk, Credit Constraints and Financial Efficiency in Peruvian Agriculture," *The Journal of Development Studies*, 46(6), 981–1002.

Garmaise, M. J., and G. Natividad (2016): "Consumer Default, Credit Reporting and Borrowing Constraints," Working Paper.

Gast, F., P. Benavides, J. R. Sanz, J. C. Herrera, V. H. Ramírez, M. A. Cristancho, and S. M. Marín (2013): *Manual del Cafetero Colombiano, Investigación y Tecnología para la Sostenibilidad de la Caficultura*. Federación Nacional de Cafeteros, Cenicafé.

Ghatak, M., and T. W. Guinnane (1999): "The economics of lending with joint liability: theory and practice," *Journal of Development Economics*, 60(1), 195–228.

Giné, X., J. Goldberg, and D. Yang (2012): "Credit Market Consequences of Improved Personal Identification: Field experimental Evidence from Malawi," *The American Economic Review*, 102(6), 2923–2954.

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

Giné, X., K. Krishnaswamy, and A. Ponce (2011): "Strategic Default in joint liability groups: Evidence from a natural experiment in India," Work in progress.

Goeree, J. K., C. A. Holt, and S. K. Laury (2002): "Private costs and public benefits: Unraveling the effects of altruism and noisy behavior," *Journal of Public Economics*, 83(2), 255–276.

González-Uribe, J., and D. Osorio-Rodríguez (2014): "Information Sharing and Credit Outcomes: Evidence from a Natural Experiment," Working Paper.

Gopakumar, G. (2016): "Banks Make Emotional Appeals to Get Borrowers to Repay Loans," Live Mint (October 12). Accessed March 9, 2017, http://www.livemint.com/Industry/1kDe2a4XJwUixBmzBYSXLK/Banks-make-emotional-appeals-to-get-borrowers-to-repay-loans.html.

Gordon, R., and W. Li (2009): "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 93(3), 855–66.

Guala, F. (2005): *The Methodology of Experimental Economics*. Cambridge University Press, New York.

Guirkinger, C., and S. R. Boucher (2008): "Credit Constraints and Productivity in Peruvian Agriculture," *Agricultural Economics*, 39(3), 295–308.

Guiso, L., P. Sapienza, and L. Zingales (2013): "The determinants of attitudes toward strategic default on mortgages," *The Journal of Finance*, 68(4), 1473–1515.

Hainz, C., and H. Hakenes (2012): "The politician and his banker — How to efficiently grant state aid," *Journal of Public Economics*, 96(1), 218–225.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2015): "The making of homo honoratus: From omission to commission," Working Paper No. 21210, National Bureau of Economic Research.

Harrison, G. W., S. J. Humphrey, and A. Verschoor (2010): "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda," *Economic Journal*, 120(543), 80–104.

Harrison, G. W., and E. E. Rutström (2008): "Risk Aversion in the Laboratory," in *Risk Aversion in Experiments*, ed. by J. C. Cox, and G. W. Harrison, pp. 41–196. Emerald.

Hey, J. D., and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), 1291–1326.

Holmstrom, D. (1979): "Moral Hazard and Observability," *The Bell Journal of Economics*, 10(1), 74–91.

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

Jayachandran, S. (2006): "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of political Economy*, 114(3), 538–575.

Kahneman, D., P. Slovic, and A. Tversky (1982): *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge.

Karlan, D., M. Morten, and J. Zinman (2015): "A personal touch: text messaging for loan repayment," Working Paper No. 17952, National Bureau of Economic Research.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry (2014): "Agricultural Decisions after Relaxing Credit and Risk Constraints," *Quarterly Journal of Economics*, 129(2).

Karlan, D. S. (2007): "Social connections and group banking," *The Economic Journal*, 117(517), F52–F84.

Kaur, S. (2015): "Nominal Wage Rigidity in Village Labor Markets," Working Paper.

Keasey, K., and P. Moon (1996): "Gambling with the House Money in Capital Expenditure Decisions: An Experimental Analysis," *Economics Letters*, 50(1), 105 – 110.

Keeler, J. P., W. L. James, and M. Abdel-Ghany (1985): "The Relative Size of Windfall Income and the Permanent Income Hypothesis," *Journal of Business & Economic Statistics*, 3(3), 209–215.

Klein, B., R. Meyer, A. Hannig, J. Burnett, M. Fiebig, et al. (1999): *Better practices in agricultural lending*. Food and Agriculture Organization of the United Nations (FAO).

Korenok, O., E. L. Millner, and L. Razzolini (2013): "Impure altruism in dictators' giving," *Journal of Public Economics*, 97, 1–8.

Kurosaki, T., and H. U. Khan (2012): "Vulnerability of microfinance to strategic default and covariate shocks: evidence from Pakistan," *The Developing Economies*, 50(2), 81–115.

Levitt, S. D., and J. A. List (2007): "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?," *Journal of Economic Perspectives*, 21(2), 153–174.

Liberman, A. (2016): "The Value of a Good Credit Reputation: Evidence from Credit Card Renegotiations," *Journal of Financial Economics*, 120(3), 644–660.

Lozano, A. (2009): "Acceso al crédito en el sector cafetero colombiano," *Ensayos sobre Economía Cafetera*, 25(95), 121.

Mester, L. J. (1997): "What's the Point of Credit Scoring?," *Business Review*, 3, 3–16.

Muñoz-Mora, J. C. (2010): "Los caminos del café: Aproximación a la Eelación entre el Conflicto Armado Rural y la Producción Cafetera Colombiana," *Ensayos sobre Política Económica*, 28(63), 14–65.

——— (2016): "Fighting while 'Talking': Assessing the Impact of Civil War on Agriculture," Doctoral Dissertation Chapter.

Ocampo, J. A. (2014): "Misión para la Transformación del Campo," Departamento Nacional de Planeación.

Ouyang, M. (2009): "The Scarring Effect of Recessions," *Journal of Monetary Economics*, 56(2), 184–199.

Padilla, A. J., and M. Pagano (2000): "Sharing Default Information as a Borrower Discipline Device," *European Economic Review*, 44(10), 1951–1980.

Pagano, M., and T. Jappelli (1993): "Information Sharing in Credit Markets," *The Journal of Finance*, 48(5), 1693–1718.

Pelka, N., O. Musshoff, and R. Weber (2015): "Does Weather Matter? How Rainfall Affects Credit Risk in Agricultural Microfinance," *Agricultural Finance Review*, 75(2), 194–212.

Petrick, M. (2004): "A Microeconometric Analysis of Credit Rationing in the Polish Farm Sector," *European Review of Agricultural Economics*, 31(1), 77–101.

Pruckner, G. J., and R. Sausgruber (2013): "Honesty on the streets: A field study on newspaper purchasing," *Journal of the European Economic Association*, 11(3), 661–679.

Rocha Sousa, M., J. Gama, and E. Brandao (2016): "A New Dynamic Modeling Framework for Credit Risk Assessment," *Expert Systems With Applications*, 45, 341–351.

Rosenzweig, M. R., and K. I. Wolpin (1993): "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India," *Journal of political economy*, pp. 223–244.

Scott, D. F. (2007): "Strengthening the governance and performance of state-owned financial institutions," World Bank Policy Research Working Paper No. 4321.

Sen, A. K. (1977): "Rational fools: A critique of the behavioral foundations of economic theory," *Philosophy & Public Affairs*, pp. 317–344.

Sharpe, S. A. (1990): "Asymmetric Information, Bank Lending, and Implicit Contracts: A stylized Model of Customer Relationships," *The journal of finance*, 45(4), 1069–1087.

Snowden, P. (1991): "World development report 1989: Financial systems and development. Oxford University Press, New York for The World Bank 1989; 155 pp. plus tables," .

Taylor, S., and S. Thompson (1982): "Stalking the Elusive Vividness Effect," *Psychological Review*, 89(2), 155–181.

Thaler, R. H., and E. J. Johnson (1990): "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice," *Management Science*, 36(6), 643–660.

Turbay, S., B. Nates, F. Jaramillo, J. J. Vélez, and O. L. Ocampo (2014): "Adaptación a la Variabilidad Climática entre los Caficultores de las Cuencas de los Ríos Porce y Chinchiná, Colombia," *Investigaciones Geográficas, Boletín del Instituto de Geografía*, 2014(85), 95–112.

Vercammen, J. A. (1995): "Credit Bureau Policy and Sustainable Reputation Rffects in Credit Markets," *Economica*, pp. 461–478.

Wei, Y., P. Yildirim, C. Van den Bulte, and C. Dellarocas (2015): "Credit Scoring with Social Network Data," *Marketing Science*, 35(2), 234–258.