

What limits bank lending in emerging markets?

An experiment testing informational and contractual frictions

M. Ali Choudhary*

State Bank of Pakistan &

Center for Economic Performance, London School of Economics

Anil K. Jain

Federal Reserve Board

June 21, 2021

Abstract

Credit access is limited in rural areas, especially in developing economies. Using a novel three-stage experimental design in Pakistan, first, we document that bank lending only serves a small fraction of rural credit demand. Second, we test the importance of information and enforcement technology frictions for limiting bank lending

*This research would not have been possible without the substantial ground support of Shafqat Ali Shah. We would like to thank Atif Bajwa, Abhijit Banerjee, Bastian von Beschwitz, Shahid Amjad Chaudhry, Samar Hasnain, Naveed Hamid, Mushtaq Khan, Logan Lewis, Annie McCrone, Ijaz Nabi, Imran Rasul, Piruz Saboury, and seminar participants at the IFABS conference (Angers), MWIEDC conference (Purdue), IGC-SBP Conference (Karachi) and AD Conference (LSE, Lahore). We would like to acknowledge the exceptional research assistance from Mukhtar ul Hasan and Khurram Ashfaq Baloch. All data collection was funded by the State Bank of Pakistan. The findings and conclusions in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System, the views of any other person associated with the Federal Reserve System, or the State Bank of Pakistan.

by randomly varying loan contractual terms across farmers. We find that enforcement technology is the primary friction for limiting bank lending. Third, using a final survey, we document that farmers tend to correctly identify the financial consequences of non-repayment. Fourth, our results suggest one possible solution to overcome this financial friction—a motivated and interlinked intermediary.

JEL Classification: C93, G21, O16

Keywords: Credit markets, banks, asymmetric information, contracts

1 Introduction

Credit access, and financial inclusion more broadly, can greatly improve consumer welfare, yet access remains severely lacking. Greater access to credit can increase income (Karlan and Zinman [2009b]), reduce inequality (Solis [2017]), increase insurance (Udry [1994]), smooth consumption (Gross and Souleles [2002]), and increase entrepreneurship (Banerjee et al. [2017]). However, credit in rural areas is limited, expensive, and unreliable in many less economically developed countries (Besley [1994]). Despite having been highlighted for decades (Braverman and Guasch [1986]), the problem of having no access to financial access still affects over half of the global population (Cull et al. [2013]). Some argue that low profitability from serving low income households limits formal financial participation but there is significant empirical evidence that this does not preclude financial access (Chaia et al. [2009]).¹ Therefore, considering both the virtues of financial access against the dearth of financial opportunities raises important academic and policy questions: What are the key financial obstacles to greater financial access and are there innovative methods to overcome these obstacles? This paper focuses on credit access obstacles and suggests one possible solution.

The existing literature outlines various theories for the cause of the lack of formal credit access but the empirical evidence is limited (Karlan and Morduch [2010]). This paper fills this gap in the literature through a randomized controlled trial in rural Pakistan that finds large credit demand but limited bank credit supply. Our paper’s main contributions are that, in our setting, we identify poor bank enforcement technology—rather than screening technology—as the key financial friction that limits greater credit access. Moreover, our experiment suggests one possible solution to overcome this friction—a motivated and interlinked financial intermediary.

To identify reasons for limited bank lending in rural areas, we use a novel three stage experimental design with three key stakeholders, a set of sugarcane farmers, a non-traditional

¹Chaia et al. [2009] find that that at least two-thirds of those who use formal financial services live on less than \$5 a day in Africa, Asia, and the Middle East and argue that, “it is possible to serve low-income communities at scale with financial services, but there are still billions left to reach”.

credit intermediary (a sugar mill), and a bank.

In the first stage, we document the extent of credit demand that the bank and the non-traditional intermediary would be willing to service. Specifically, we ask the two different possible lenders to screen a set of sugarcane farmers that have signalled an interest in procuring a loan. To ensure accurate due diligence, we inform the non-traditional intermediary that they will be the residual claimant for a strict subset of the loans that it positively screened.²

In the second stage, we test the relative importance of information frictions and enforcement technology frictions for limiting credit access. Due to the close geographic, financial, and economic links between the sugar mill and the sugarcane farmers, relational contracting and relationship banking suggests that the non-traditional intermediary (the mill) will have superior enforcement technology (Levin [2003] and Boot [2000]). Furthermore, interlinked relationships (Braverman and Stiglitz [1982]) suggests that the mill has superior information than the bank for lending to the farmers. However, given the lack of credit specialization, the non-traditional intermediary is likely to be hampered by higher funding costs.

In our experiment, we classify farmers into groups according to whether the bank and the mill are willing to lend to the farmer, and subsequently randomize the loan's contractual terms across these groups.³ Specifically, some farmers get loans direct from the bank and some bank loans are guaranteed by the mill. By comparing the loan outcomes for farmers that are selected by *different* lenders, but get the *same* loan contract, we aim to isolate differences in lenders' screening technology. By comparing the loan outcomes for farmers that are selected by the *same* lender, but get *different* loan contracts, we aim to isolate differences in enforcement technology. Our research methodology is inspired by Karlan and Zinman [2009a] who empirically identify moral hazard and adverse selection by randomly varying loan offer contractual terms *before* and *after* the consumer has accepted the

²To be precise, the intermediary was asked to guarantee the repayment on loans to this strict subset of loans.

³Specifically, using the results in the first stage of the experiment, we can classify farmers into four mutually exclusive groups: those farmers that (i) no lender is willing to lend to, (ii) only the mill is willing to lend to, (iii) only the bank is willing to lend to, and (iv) both lenders are willing to lend to.

loan terms, as well as, randomizing contractual terms before the loan's renewal.

In the final stage, to uncover the reasons for the superior performance of loans that were guaranteed by the mill, we interview the farmers to understand how the lenders tried to collect repayment and to understand why the farmers did, or did not, repay.

Turning to our experiment's results: In the first stage, we find that the bank is only willing to service a small portion of the demand for loans (27 percent), and though the intermediary is willing to lend to more (52 percent), there is still a sizeable majority that is not served by any lender (39 percent).⁴

In the second stage, we identify that the key financial friction limiting bank lending in rural areas is *poor bank enforcement*—rather than screening—technology. Specifically we find that the overdue rates were nearly *60 percentage points* lower for loans where the mill was the residual claimant than for those loans where the bank was the residual claimant. Moreover, the mill's enforcement power led to low default rates with under five percent of those loans that were guaranteed by the mill overdue more than 90 days—a default rate that is economically sustainable for all stakeholders. Somewhat surprisingly, we do not find that the mill had superior *screening technology* than the bank; specifically, repayment rates were similar for those farmers that were chosen by different lenders but got the same loan contract.

Given the large statistically and economically significant effects in stage 2 of our experiment, we consequently decided to add a final survey of farmers. From this survey we learn that farmers perceived different costs of non-repayment according to the loan contract. For farmers with loans guaranteed by the mill, the majority of farmers stated non-repayment would damage their future relationship with the mill. Whereas, for farmers with direct bank loans, a relatively larger fraction believed it would cause higher interest rates in the future. Moreover, there is a large difference between the enforcement technology utilized by the mill and the bank. The mill ensures repayment by collecting the payment at the time of milling the farmer's sugarcane. In contrast, the bank uses

⁴Note these probabilities do not need to sum to 100 percent since the bank and the intermediary were both willing to lend to some farmers.

a mixture of messages to the farmer and costly visits to the farmer after the end of the harvesting season.

Our paper aims to understand a key policy and academic question in empirical economics: Why is bank lending in rural areas in developing countries so limited? Specifically, what is the market failure? The theoretical literature—building from the classical papers on information asymmetries, moral hazard and adverse selection ([Jaffee and Russell \[1976\]](#) and [Stiglitz and Weiss \[1981\]](#))—have raised a number of possible causes for the market failure; including, costly or insufficient enforcement technology ([Hoff and Stiglitz \[1998\]](#)), limited commitment ([Matthews \[2001\]](#)), limited liability ([Innes \[1993\]](#)), differentially informed lenders ([Ghosh and Ray \[2016\]](#)), insufficient borrower risk capacity or tolerance ([Boucher and Carter \[2001\]](#) and [Boucher et al. \[2008\]](#)), as well as, large transaction costs ([Giné \[2011\]](#)).⁵ However, these frictions are notoriously difficult to identify in practice ([Karlan and Morduch \[2010\]](#)). Understanding the key cause for this problem is imperative for designing the optimal policy response given the large potential benefits from credit access.

Our experiment finds that the key limitation for bank lending in rural areas is banks' inability to enforce the loan contract and that one solution that reduces this financial friction, and importantly, to potentially economically sustainable levels is through leveraging an interlinked intermediary (in this case, the sugar mill). The success of this financial innovation is evident with the bank, following our experiment, independently expanding the trial of using mills as loan guarantors to two other mills, and earmarking an expansion of the scheme from an initial \$0.7 million to \$35 million in credit disbursements.⁶

Inadequate formal financing forces individuals to rely on less formal forms of credit such as moneylenders; trade credit; rotating credit and savings associations (ROSCAs); and interpersonal financial networks ([Collins et al. \[2009\]](#), [Rutherford \[2000\]](#)). However, these

⁵As noted by [Banerjee and Duflo \[2011\]](#), these frictions often cause a multiplier effect, whereby the higher administrative costs cause higher interest rates, which in turn, amplify the initial friction, necessitating even larger rises in the interest rates.

⁶At some level, our experiment is formalising an existing trilateral lending relationship as, in a survey of moneylenders in rural Pakistan, [Aleem \[1990\]](#) discovers that that 30 percent of informal moneylenders' funds come directly or indirectly from institutional sources, such as banks, wholesalers, and cotton mills.

informal institutions often suffer from high prices, limited supply, poor resilience to financial shocks, and may exacerbate economic inequality.⁷

To remedy the problem of inadequate credit access, policymakers have tried a plethora of methods, for example, government ownership and control of banks, including rural development banks (Besley [1994]); subsidising credit to rural areas (Hoff and Stiglitz [1990]); establishing or supporting credit guarantee schemes (almost 100 countries have a scheme for some sector (Green [2003]), and in a recent global survey, more than thirty percent of credit guarantee schemes had some form of government ownership, Beck et al. [2010]); and mandating the fraction of rural bank branches (for example, India), or mandating the fraction of loans to rural areas (for example, Philippines). Overall, the success of these methods has been limited and much more innovation is needed.

In recent years, given the dearth of effective lending methods, there is interest in new mechanisms of increasing credit access, often utilizing key aspects of the borrower's social and professional network. The leading innovation is the microfinance industry (with an estimate of \$60 to \$100 billion with 120 million clients in 2015). Additionally, there are trials examining the use of community chosen loan-officer (Maitra et al. [2020]), innovative forms of asset collateralization (Jack et al. [2016]), and greater use of non-traditional data or fintech data (Bharadwaj et al. [2019]). Our paper introduces another potential solution.

The rest of the paper proceeds as follows: Section (2) presents the research design of our experiment, including outlining the sample frame, borrower summary statistics, timing, and empirical strategy. Section (3) presents and analyzes the results of our experiment. Section (4) presents a summary of our findings and concluding remarks for future research.

⁷In a survey of six rural villages in Kerala and Tamil Nadu, Dasgupta [1989] finds that average interest charged by professional money lenders is about 52 percent and provided almost half of total available credit. Moreover, Aleem [1990] documents there are significant financial frictions for moneylenders too, with rural Pakistani moneylenders' average cost of capital being over 30 percent—significantly higher than the prevailing 10 percent for bank deposit rates at that time. There is significant evidence that interpersonal credit and risk-sharing networks have limited effectiveness, especially during correlated shocks (for example, natural disasters) Dercon [2002]. Finally, Blumenstock et al. [2016] finds that following a large earthquake, better connected (larger social networks and more centrally located) and wealthier individuals were likely to receive greater financial transfers.

2 Research Design

This paper reports the results from a three stage field experiment in Pakistan. In the first stage, we analyzed what type of lenders are willing to lend to different farmers and the key farmer characteristics that determine the lender’s decision to offer loans. In the second stage, by varying the farmer’s loan contractual terms, we identify how the importance of both different lender information technology and lender enforcement technology determine loan repayment outcomes. Finally, we survey the farmers to understand why they did or did not repay their loans.

Our experiment was started in June 2016 in 60 villages in the district of Matiari, Sindh, Pakistan.⁸ This rural area is 22 miles from the closest city of Hyderabad, with the main occupation being sugar and wheat farming. Matiari is significantly less developed and poorer than the rest of Pakistan, with the average monthly household income only around Pakistani Rupee 21,000 (about 209 US dollars) in 2015—30 percent lower than the average income in Pakistan—and with a literacy rate of only 45 percent, significantly below the national average of 60 percent ([Pakistan Bureau of Statistics \[2015\]](#)).

Our experiment uses three key stakeholders: (i) the bank, (ii) the sugar mill, and (iii) the sugarcane farmers. To obtain a baseline survey for our sampling frame, we asked all sugarcane farmers in the Matiari district if they would like a bank loan for the next sugar growing season.⁹ A few key characteristics stand out. Of the 528 farmers that wanted a loan, most farmers sell the majority of their sugarcane crop to this sugar mill (on average 73 percent of their total crop), and sugar is one of their main crops (23 percent of their crop sales) suggesting that the sugar mill plays a substantial economic role for these farmers. Table (2) provides more details about the farmers.

Our sampling frame is the set of sugar farmers in Matiari who expressed an interest in taking a bank loan (528 farmers).

⁸We use the word village to describe what are known as “Deh” in the local language, Sindhi. Formally, deh are areas that have three key characteristics (i) a separate record-of-rights or (ii) have been separately assessed to land revenue or (iii) the Board of revenue may declare to be a Deh.

⁹The original list of sugarcane farmers in the Matiari district came from a State Bank of Pakistan 2011-2012 door-to-door survey of farmers in that area.

2.1 Experiment design: Stage 1—Testing willingness to lend by lender type

To understand how different lender types assess credit risks, we first asked the sugar mill and the bank to *independently* assess the creditworthiness of all the farmers that expressed an interest in getting a loan. The bank was asked to assess whether the bank would be willing to offer a loan (with no loan guarantee from the mill) to the farmer and the mill was asked whether they would guarantee a loan to the farmer.¹⁰ To ensure accurate assessments, the mill was informed that they would guarantee a strict subset of those farmers they said they were willing to guarantee (by randomizing the set of farmers that got a loan guarantee we can identify the importance of the lender’s enforcement and information technology—the focus of section (2.2)).

To ensure consistency across the experiment—apart from the loan size and whether the loan was guaranteed by the mill—the loan contract’s terms were standard across all farmers. The farmer was charged an interest rate of 13 percent on each loan, irrespective of whether the loan was guaranteed by the mill. If the mill guaranteed a loan, the bank would pay a 2 percent fee (of the loan amount) to the mill, and the mill would repay the bank if the loan was not repaid. Therefore, if the farmer took a \$100 loan from the bank that was guaranteed by the mill, at the end of the loan term the farmer would owe \$113 to the bank. If the loan was fully repaid, the mill would receive \$2 from the bank, and the bank would earn a net return of \$11. If the farmer did not repay the loan, the mill would pay the bank \$111, and the bank would still earn a net return of \$11.¹¹

To identify the farmer characteristics that are important for each lender’s classification of a farmer’s creditworthiness, we run regression (1) separately for each lender type.

¹⁰The results of this bank and mill investigation were not shared with the farmer; rather, depending on the bank’s and the mill’s assessment and the subsequent randomization, the farmer was either offered a specific loan contract or no loan (more details in the next section).

¹¹An alternative explanation is that if the mill was guaranteeing the loan for the farmer, the contractual net payments for the mill are identical to the scenario where the mill borrowed \$100 from the bank at an interest rate of 11%, and subsequently lent \$100 to the farmer at an interest rate of 13%.

$$\text{Creditworthy}_{lf} = \beta_l \text{Farmer characteristics}_f + \epsilon_{lf} \quad (1)$$

Where ‘Creditworthy’ is a binary variable, defined as whether the bank (mill) was willing to offer a loan (guarantee a bank loan) to farmer f . The lender type (either the bank or the mill) is denoted by l . In table (3), in section (3.1), we regress the mill’s (column 1) and the bank’s (column 2) classification as whether the farmer is creditworthy on the farmer’s characteristics. In the third column we analyse the differences in each lender’s classification by computing the difference in these coefficients.¹²

To explore similarities and differences across the lender’s choice, we define three mutually exclusive groups: group (A) “creditworthy only mill $_f$ ”, group (B) “creditworthy both $_f$ ”, and group (C) “creditworthy only bank $_f$ ”. Outcome (A) is equal to one if farmer f was chosen only by the mill, outcome (B) is equal to one if farmer f was chosen by both lenders, and outcome (C) is equal to one if farmer f was chosen by only the bank. Figure (1) illustrates the possible set of each borrower. Out of the 528 farmers, 185 farmers were in group (A), 88 farmers in group (B), and 51 farmers in group (C).

2.2 Experiment design: Stage 2—Testing asymmetric information and differing enforcement technology

The second stage of our experiment varied the loan contractual terms across farmers to identify the relative importance of (i) different lender screening technology and (ii) lender enforcement technology to determine loan repayment outcomes.

From the willingness to lend experiment, farmers were allocated to four mutually exclusive groups: (A) only the mill defined the farmer as creditworthy, (B) both the mill and the bank defined the farmer as a creditworthy, (C) only the bank defined the farmer as a creditworthy, and (D) neither lender defined the farmer as creditworthy.

To explore the effect on repayment, we randomized loan contract offers to farmers de-

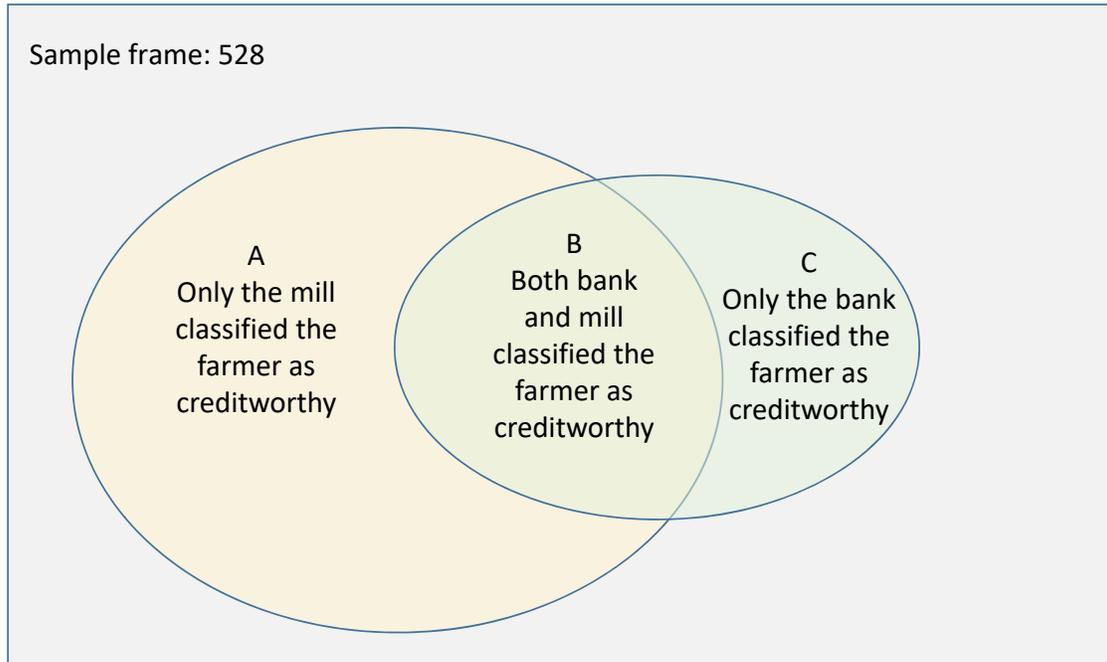
¹²Note that in the third column of table (3), we have double the number of observations as columns 1 and 2 because we have two data points for each farmer—one for each lender.

pending on their group (stratifying on the number of acres each farmer planned to grow sugarcane). The ideal experiment would randomize contracts (direct loan from the bank or guaranteed loan by the mill) across all possible groups (A) to (D). However, in practice, the mill was only willing to guarantee loans that the mill had previously denoted as creditworthy. Moreover, in the interest of greater experimentation, the bank was willing to offer some loans to farmers that they had previously denoted as not creditworthy. Specifically, the bank was willing to offer a direct bank loan to farmers that either the mill or the bank had classified as creditworthy but not to farmers that neither the mill or the bank had denoted as creditworthy. Therefore, we were able to randomize both lender contracts to farmers in groups (A) and (B) but farmers in group (C) only received direct bank loans. Since no lender was willing to lend to group (D), they were omitted from the rest of the study. For reference, tables (11) and (12) in Appendix (5.1) report the definition of each group.

Across the various treatment arms, the bank was willing to lend a total of approximately \$700,000. Given that the initial interest in receiving loans was larger than the available funds for lending, farmers were randomized on whether they were offered a loan, and if so, the loan's contractual terms. To ensure the allocation of farmers across treatments were more likely to be balanced ex-post, we stratified farmers on the number of acres that the farmer planned to plant with sugarcane. Since we were unable to randomize contracts to farmers in group C—thereby limiting the potential inference—we selected a greater fraction of farmers in groups A and B, than group C. In total, just over 90 percent of farmers that were screened as creditworthy in groups A and B were offered a loan, and just over 50 percent of farmers that were screened in group C were offered a loan. Overall, as described later, around 75 percent of farmers accepted the loan offers.

Overall, following the randomization of loan offers, there were five mutually exclusive groups that took loans. From the set of borrowers that the mill were willing to guarantee (group A), one set (group A1) were offered bank loans with a loan guarantee from the mill and a second set (group A2) were offered direct loans from the bank. Note that group A2 were not classified as creditworthy farmers by the bank—a key part of our experimental

Figure 1: Possible classification of each farmer between the mill and the bank



Note: The 528 farmers that requested a loan was screened by both the mill and the bank. As discussed in section (2.2), farmers could be allocated to four mutually exclusive groups: (A) only the mill defined the farmer as creditworthy (185 farmers), (B) both lenders defined the farmer as creditworthy (88 farmers), (C) only the bank defined the farmer as creditworthy (51 farmers), or (D) no lender defined the farmer as creditworthy (204 farmers).

design. From the set of farmers that both the mill and the bank classified as creditworthy (group B), one set (group B1) were offered bank loans with a loan guarantee from the mill, and a second set (group B2) were offered direct loans from the bank. Finally, the set of borrowers that only the bank classified as creditworthy (group C) were offered only direct bank loans. Figure (2) illustrates the possible groups, and for reference, table (12) in Appendix (5.1) repeats the definitions.

To minimize the ex-ante differences across our treatment groups we stratified our sample on the number of acres each farmer planned to grow sugarcane. As evidence for successful randomization, tables (14) and (15) in the appendix show that there were no statistically significant differences within each group (that is, between farmers in groups A1 and A2, and between farmers in groups B1 and B2).

Another key statistical concern is selective attrition. As is common in randomized control trials, even though some farmers were approved to receive a bank loan, some farmers—following the randomization—decided to not exercise their bank loan option. To test for possible selective attrition, in table (16) in the Appendix, we examine whether the farmers that took the loan were significantly different on observables (on measures such as assets, farm size) to the farmers that did not take the loan but were offered a loan. Moreover, our experiment had high take-up rates with around 75 percent of farmers eventually taking the loan. Table (17) in the Appendix show the relative uptake of the various loans across the various groups, and as is quickly evident, there were no significant differences across randomized treatments in groups A and B and there is no evidence for selective attrition in our experiment.

Overall, a total of 204 loans were dispersed with an average loan size of \$3,400—a relatively large sum given that the average monthly household income in this area was \$209 (Pakistan Bureau of Statistics [2015]).

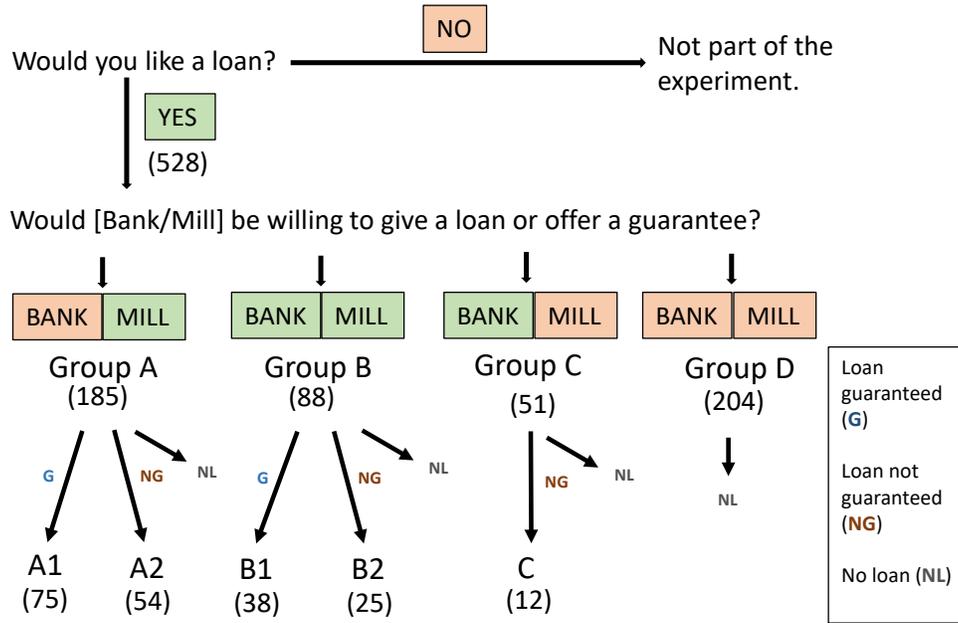
2.2.1 Empirical specifications

Enforcement technology

To test the importance of each lender’s enforcement technology, we compare the repayment rates for those farmers that were defined as creditworthy by the same lender but got different loan contracts. Our research design is similar to Karlan and Zinman [2009a]’s seminal paper but tests a different economic question. Given our research design, we can undertake two similar tests but on different sets of farmers.

First, in table (4) columns 1 and 2, we analyse differences in repayment rates for the farmers defined as creditworthy by only the mill, therefore holding the credit screening technology, but received different contracts (*groups A1 and A2*). Specifically, using the group of growers which the mill exclusively selected we compare the repayment rates between (i) farmers that received a guaranteed loan by the mill (*group A1*) and (ii) farmers that received a direct bank loan (*group A2*). Second, in columns 3 and 4, we analyse differences

Figure 2: Experimental Design



Note: The numbers in parenthesis are the number of farmers in each group. For example, there were 528 farmers who were interested in a bank loan, and of these 528 farmers, 185 farmers were classified as creditworthy by the mill but not by the bank (Group A). 88 farmers were classified as creditworthy by the mill and the bank, and 51 were classified as creditworthy by only the bank. The farmers in groups A, B, and C, were randomly allocated to three possible treatments: direct bank loan (loan not guaranteed), bank loan guaranteed by the mill, or no loan. Given that neither the bank nor the mill classified the farmers in group D as creditworthy, none of these farmers were offered loans. Some farmers even though they initially described interest in taking a loan, decided against taking a loan, as described in section 2.2.

in repayment rates for the common set of farmers passing the creditworthy criterion of both the mill and the bank (*groups B1 and B2*), specifically (i) farmers that received a guaranteed loan by the mill (*group B1*) and (ii) farmers that received a direct bank loan (*group B2*).

$$\text{Overdue}_f = \beta_E \text{Loan guarantee}_f + \beta_1 \text{Farmer characteristics}_f + \epsilon_f \quad (2)$$

where “Overdue_f” is a binary variable equal to zero if farmer *f* repaid the bank on time and one otherwise. “Loan guarantee_f” is a dummy variable equal to one if the farmer

received a loan guarantee from the mill.

The coefficient of interest, β_E , estimates whether those farmers that were selected to receive a loan guarantee had higher default rates than those farmers that received a direct bank loan. This regression tests the hypothesis that the mill's loan enforcement technology—via the loan guarantee—is superior to the bank's enforcement technology.

Finally, in table (5), we examine whether specific farmer characteristics were correlated with farmer loan defaults.¹³ To do so, we repeat the regressions in table (4) but we interact the loan contract with farmer characteristics. Specifically, we run regressions similar to the following form:¹⁴

$$\begin{aligned} \text{Overdue}_f = & \beta_{CE} \text{Loan guarantee}_f \times \text{Farmer characteristics}_f \\ & + \beta_1 \text{Loan guarantee}_f + \beta_2 \text{Farmer characteristics}_f + \epsilon_f \end{aligned} \quad (3)$$

Screening technology

To test the importance of each lender's screening technology, we compare the repayment rates between those farmers who received the same loan contract, therefore holding constant the enforcement ability, but were classified as creditworthy by different lenders. By comparing borrowers that were selected by different lenders but were offered the same loan terms, we can identify the relative effectiveness of each lender's screening technology.

First, in table (6) columns 1 and 2, for farmers that received the same contract (guaranteed bank loan) we analyse differences in repayment rates between (i) farmers who only the mill defined as creditworthy (*group A1*) and (ii) farmers who both the bank and mill defined as creditworthy (*group B1*). Second, in table (6) columns 3 to 6, we conduct a symmetric test and compare repayment rates for the farmers who got a direct bank loan (*groups A2, B2, and C*) but were defined as creditworthy by different lenders.

¹³These regressions were added following the end of the experiment to investigate why farmers were more likely to default.

¹⁴Note for column 3, we also include a control for whether only the mill classified whether the farmer was creditworthy, that is, whether the farmer was in groups A1 or A2.

Specifically, in table (6), we run regressions similar to the following form on the selected samples:

$$\text{Overdue}_f = \beta_S \text{Creditworthy only Mill}_f + \beta_1 \text{Farmer characteristics}_f + \epsilon_f \quad (4)$$

where “Overdue_{*f*}” is a binary variable equal to zero if farmer *f* repaid the bank on time and one otherwise. “Creditworthy only Mill_{*f*}” is a dummy variable equal to one if only the mill selected farmer *f* as creditworthy, and zero otherwise.

The coefficient of interest, β_S , estimates whether those farmers that were selected as creditworthy by the mill had lower default rates than those farmers that were selected by the bank. This regression tests the hypothesis that the mill’s loan screening technology is superior to the bank’s technology.

Finally, in table (7), we conduct a pooled regression that examines the effect of the screening technology but controls for the different loan contract. By controlling for the loan contract, we can increase our sample size and consequently increase the power of our test.

2.3 Experiment design: Stage 3–Final survey

Following the completion of our experiment, we were struck by the large effects that our experiment uncovered. To explore these results in more detail, we conducted an additional survey that asked the farmers that received loans for why they did or did not repay the loan, and specifically asked questions on what they expected the consequences of not repaying the loan would be. In March 2020, we tried to interview all farmers that took a loan, and we successfully managed to interview (via telephone) 128 farmers, of which 60 farmers received a guaranteed loan from the sugar mill and 68 farmers received a direct loan from the bank.¹⁵

¹⁵A successful callback rate of over 60 percent.

2.4 Timing

The experiment started in mid-2016 and ended in early-2020, and figure (3) shows the timeline of the experiment. In June 2016, we interviewed all sugar farmers that grew sugarcane in the district of Matiari to assess whether they were interested in availing a new bank loan. In total we interviewed 1,455 sugar farmers. Of these 1,455 farmers only 528 farmers were interested in taking a bank loan. Of the farmers that were interested in a loan, we conducted a baseline survey.

In July and August 2016, the bank and the mill simultaneously (and independently) screened the set of farmers that were interested in taking a loan.

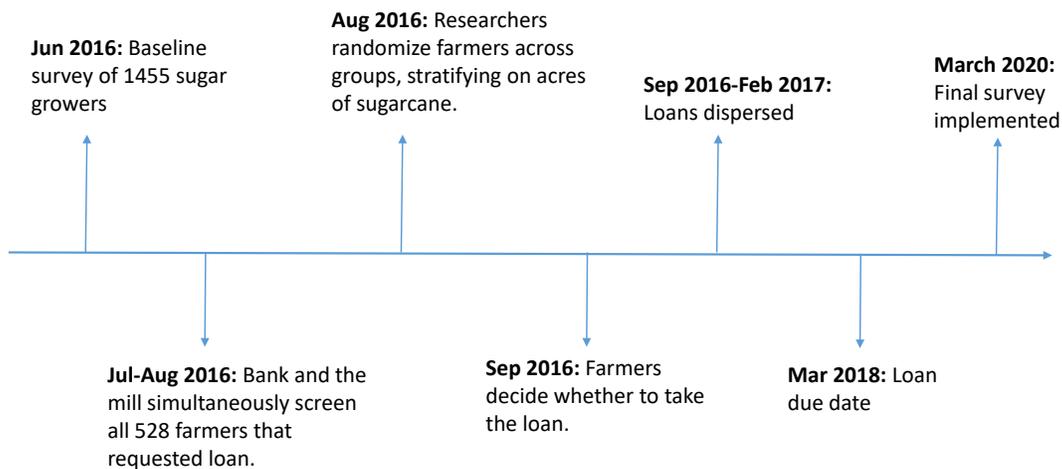
By August 26 2016, we had received the lending decisions of the bank and the mill, and subsequently randomized loan contracts across the appropriate farmers on the basis of the baseline surveys. By September 15 2016, farmers were notified of the decision of whether they would receive a loan, and received the details of the loan contract where applicable. The growers were given 15 days to decide on whether to accept the loan offer. Loans were dispersed from late September 2016 through February 2017 (that is, before sugarcane is sowed) with a due date for March 31 2018 (that is, a tenor of approximately 18 months and after the harvesting season for sugarcane). Finally, in March 2020, we surveyed farmers on their reasons for repaying or not repaying their loans.

2.5 Data

Our paper utilizes three key sources of data: hand-collected survey data, credit registry data from the State Bank of Pakistan's e-CIB database (this data has also been used in [Khwaja and Mian \[2005\]](#) and [Choudhary and Jain \[2020\]](#)), and administrative data provided by the sugar mill.

To procure more data on the farmers in our sample, we conducted surveys in three waves; first, to determine interest for credit, we conducted a telephonic survey of all sugar farmers in Matiari; second, to understand the characteristics of those farmers that wanted a loan, we solicited information on farmers' assets, crops, and demographics. Third, we

Figure 3: Timeline of the experiment



conducted a final survey to learn more about why farmers did or did not repay.¹⁶

To ascertain better knowledge on farmers' access to credit, we matched each farmer to their corresponding credit registry entry in Pakistan's national credit registry, e-CIB, run by the State Bank of Pakistan. If farmers had previously held any formal credit in the last five years, their details would be in the registry. Of the 528 farmers in our sample, only 34 percent (180 farmers) had entries in the Pakistan's credit registry, reinforcing the evidence for a lack of formal credit markets in rural Pakistan. If a farmer's details are in the registry, we match details on total amount of credit the farmer previously had, and whether the farmer repaid the loan on time. Summary statistics on the farmer's credit history (by treatment arm) are provided in table (2).

Finally, most farmers have a pre-existing relationship with the mill, so we supplement the farmer's survey data on the farmer's sales to the mill in previous years. We use information on the length of the farmer's relationship with the mill, the fraction of produce that is sold to the mill, the value of produce sold to the mill, and the distance of the farmer to the mill to complete the evidence for the farmer's relationship with the mill.

¹⁶All costs related to data collection and surveying were funded by the State Bank of Pakistan.

3 Results

In section (3.1), we start by examining the differences in the mill and the bank's determination of a farmer's creditworthiness. We explore what characteristics of the farmer that make the bank and the mill more likely to lend. In section (3.2), we analyze the difference in repayment rates following our randomization of contract types across farmers.

3.1 Stage 1: Willingness to lend by lender type

To examine how the set of farmers who are judged to be creditworthy across different lenders with different information and enforcement technology, we asked each lender to independently assess the creditworthiness of each farmer that had expressed interest in procuring a loan. We proceed in three steps. First, we outline significant aggregate differences in the set of farmers each lender chooses. Second, we document differences in farmer characteristics across the different selected groups. Finally, we more formally test the key farmer characteristics that are correlated with whether the lender defined the farmer as creditworthy.

Starting with the aggregate differences in the lenders' choices, in table (1), we immediately identify that the set of farmers that were defined as creditworthy between the two lenders significantly differed in size and composition.

The mill was significantly more willing to lend to farmers than the bank. Out of the set of farmers that expressed an interest in a loan, the mill defined almost double the number of farmers as creditworthy (group A and group B) than the bank (group B and group C).¹⁷

Additionally, the bank and the mill evaluate farmers on significantly different characteristics. There was a large number of farmers (235 farmers) that only one lender defined as creditworthy (group A and group C). Whereas, only 88 farmers were selected as creditworthy by both the bank and the mill (group B).

¹⁷ Consistent with the definitions used in the earlier part of the paper, group A are farmers that only the mill defined as creditworthy, group B are farmers that both the mill and the bank defined as creditworthy, group C are farmers that only the bank defined as creditworthy, and group D are farmers that no lender defined as creditworthy.

Table 1: Size of each group

	A	B	C	D
Number of farmers	185	88	51	204
Fraction of farmers	35%	17%	10%	39%
Observations	528			

Group A is the set of farmers that only the mill defined as creditworthy, group B is the set of farmers that both the mill and the bank defined as creditworthy, group C is the set of farmers that only the bank defined as creditworthy, and group D is the set of farmers no one defined as creditworthy.

Turning to the differences in farmer characteristics across each group, in table (2), we report summary statistics for farmers according to which lenders defined the farmer as creditworthy. In the final column (“total”), we report aggregate statistics for the sample of farmers who expressed an interest in procuring a loan.

The starkest differences in the characteristics of farmers that each lender defined as creditworthy are apparent when comparing groups A and C—the farmers that were only chosen by one lender. The bank was more significantly likely to lend to farmers that had larger farms and owned expensive farming equipment (so more likely to be richer) and have *good* credit history. Whereas, the mill was more likely to lend to farmers that sold relatively and absolutely more of their output to the mill (so more likely to have a deeper economic relationship). Surprisingly, the length of the farmer’s relationship with the mill (the number of years selling to the mill) was lower for the farmers selected by the mill than for the bank.

Table 2: Baseline characteristics and summary statistics (mean) by group

	A	B	C	D	Total
Sugarcane planted (acres)	15.14	24.69	23.48	16.03	17.46
Income from agriculture (percent)	85.88	82.58	89.81	86.58	85.93
Income from sugar (percent)	21.97	22.53	21.15	22.83	22.47
Relative value of crop sales (decile)	4.381	6.580	6.333	5.718	5.456
Educ. below high school (Y=1;N=0)	0.470	0.506	0.481	0.560	0.519
Farm size (acres)	21.15	42.75	49.53	33.10	31.62
Years selling to the mill	11.95	22.22	21.22	11.86	13.96
Sales to the mill (pct of total sales)	76.77	85.69	35.72	42.71	59.78
Within 5km of the mill (Y=1;N=0)	0.571	0.840	0.481	0.393	0.523
Relative sales to the mill (decile)	6.399	7.531	4.593	3.881	5.278
Formal credit history (Y=1;N=0)	0.298	0.432	0.444	0.345	0.348
Prev. bank loan overdue (Y=1;N=0)	0.0357	0.0370	0	0.0159	0.0246
Own-tractor (Y=1;N=0)	0.190	0.309	0.407	0.345	0.294
Own-thresher (Y=1;N=0)	0.185	0.321	0.296	0.302	0.267
Own-blade (Y=1;N=0)	0.208	0.346	0.407	0.357	0.311
Own-cultivator (Y=1;N=0)	0.214	0.321	0.407	0.357	0.309
Own-raja (Y=1;N=0)	0.214	0.346	0.444	0.353	0.313
Own-gobal (Y=1;N=0)	0.208	0.309	0.370	0.357	0.303
Own-bundmaker (Y=1;N=0)	0.143	0.247	0.259	0.246	0.214
Own-harvesting machine (Y=1;N=0)	0.0298	0.0494	0.0741	0.0159	0.0284
Observations	528				

Note: Group A is the set of farmers that only the mill defined as creditworthy, group B is the set of farmers that both the mill and the bank defined as creditworthy, group C is the set of farmers that only the bank defined as creditworthy, and group D is the set of farmers no one defined as creditworthy. The final column shows the aggregate value across all farmers. Table (13) in Appendix (5.1) describes the variable definitions.

To formally test the differences in the mill and the bank's choice of creditworthy farmers, we regress the determination of the farmer's creditworthiness by each lender on the farmer's observable characteristics in table (3). In column 1 (column 2) we regress whether the farmer was designated as creditworthy by the mill (bank) on a set of farmer characteristics and in the final column we compute the average difference between these two measures.¹⁸

The main differences between the lenders across the classification of a farmers' creditworthiness are similar to the results seen in table (2). The bank was relatively more willing to lend to farmers that had observable characteristics that are highly likely to be correlated with greater farmer wealth; for instance, the bank was relatively more willing to lend to farmer's with higher crop sales, larger farm sizes, and were more likely to own expensive farm machinery. In contrast, relatively more important characteristics for the mill were those that suggested a deeper or more valuable partnership with the mill. Therefore, those farmers who were closer to the mill, or sold more of their crop to the mill were relatively more likely to be classified as creditworthy by the mill. Surprisingly, those farmers that had had been selling to the mill for a greater time were relatively more likely to be defined as creditworthy by the *bank*. One possible explanation for this counter-intuitive result is that the farmers that have been growing sugar the longest are also likely the richest farmers, who are relatively preferred by the bank.

¹⁸To be precise, we do the following regression $\text{Creditworthy}_{lf} = \beta_1 \text{Farmer characteristics}_f + \beta_2 \text{Farmer characteristics}_f * \text{mill}_l + \epsilon_{lf}$, and report the coefficients β_2 in table (3). As the model is saturated, the coefficient β_2 is the exact difference between the regression coefficients reported in columns 1 and 2. To calculate the t-statistics in column 3, we cluster the standard errors at the farmer-level.

Table 3: Analyzing the observable characteristics that determine the farmer's creditworthiness

	(1)	(2)	(3)
	Mill	Bank	Difference
Sugarcane planted (acres)	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)
Value of crop sales (decile)	-0.024*** (0.008)	0.006 (0.007)	-0.030*** (0.011)
Educ. below high school (Y=1;N=0)	-0.011 (0.038)	-0.018 (0.032)	0.007 (0.048)
Farm size (acres)	-0.001** (0.000)	0.001 (0.001)	-0.002** (0.001)
Years selling to the mill	0.002 (0.003)	0.029*** (0.002)	-0.028*** (0.004)
Sales to the mill (pct of total sales)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)
Within 5km of the mill (Y=1;N=0)	0.137*** (0.041)	0.052 (0.032)	0.085 (0.053)
Relative sales to the mill (decile)	0.070*** (0.010)	0.003 (0.009)	0.067*** (0.014)
Formal credit history (Y=1;N=0)	0.020 (0.039)	0.028 (0.035)	-0.008 (0.051)
Prev. bank loan overdue (Y=1;N=0)	0.005 (0.081)	-0.201*** (0.063)	0.205** (0.096)
Own-tractor (Y=1;N=0)	-0.028 (0.124)	0.300*** (0.094)	-0.328** (0.156)
Own-thresher (Y=1;N=0)	0.131 (0.086)	-0.079 (0.074)	0.210* (0.119)
Own-blade (Y=1;N=0)	-0.096 (0.121)	0.049 (0.138)	-0.145 (0.184)
Own-cultivator (Y=1;N=0)	-0.008 (0.158)	-0.248 (0.170)	0.239 (0.231)
Own-raja (Y=1;N=0)	-0.029 (0.154)	0.291** (0.124)	-0.320 (0.218)
Own-gobal (Y=1;N=0)	-0.004 (0.147)	-0.316*** (0.107)	0.313* (0.178)
Own-bundmaker (Y=1;N=0)	-0.047 (0.075)	-0.080 (0.070)	0.033 (0.104)
Own-harvesting machine (Y=1;N=0)	0.062 (0.108)	0.043 (0.132)	0.019 (0.184)
Constant	0.159*** (0.056)	-0.250*** (0.041)	0.409*** (0.063)
Observations	528	528	1056

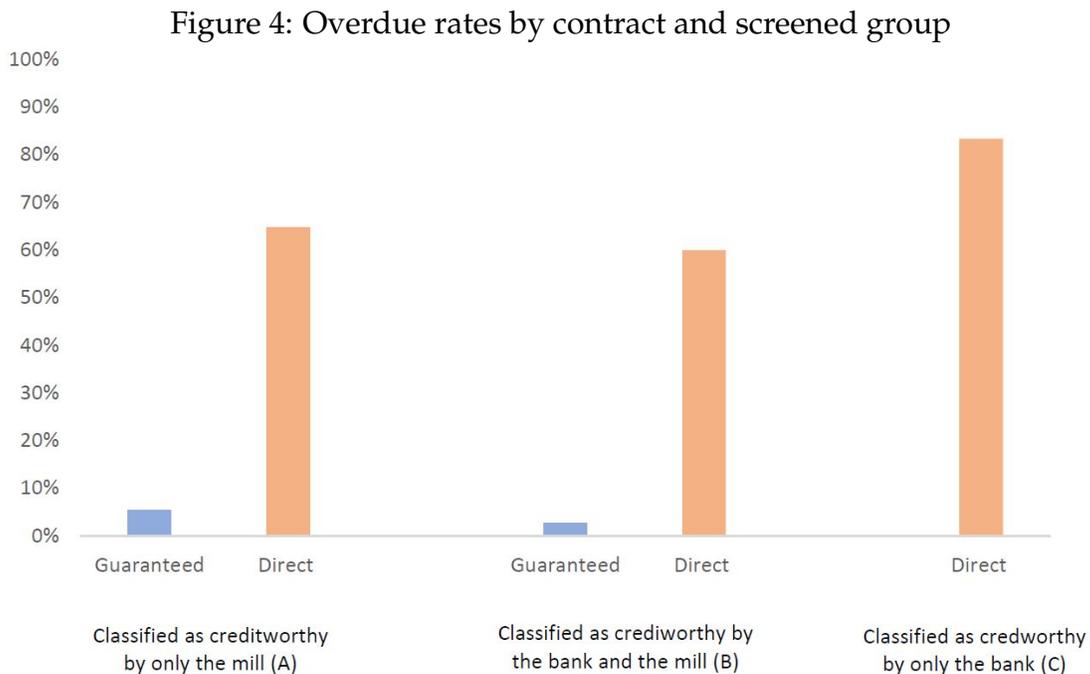
Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The result in column 1 (column 2) is the regression of the mill's (bank's) classification of the farmer's creditworthiness on farmer characteristics. The standard errors in columns 1 are 2 robust standard errors. In column 3, we report the difference between these two coefficients. Standard errors are calculated using the following regression $Creditworthy_{lf} = \beta_1 \text{Farmer characteristics}_f + \beta_2 \text{Farmer characteristics}_f * \text{mill}_l + \epsilon_{lf}$ and are clustered at the farmer level. Table (13) in Appendix (5.1) describes the variable definitions.

3.2 Stage 2: Contractual frictions

To preview our main result, figure (4) shows overdue rates by treatment arm. Across all arms, the striking result is direct bank loans (the orange bars) had substantially greater overdue rates than the loans that were guaranteed by the mill (the blue bars). The average overdue rate for direct bank loans was over 65 percent, yet, for loans that were guaranteed by the mill, the average overdue rate was just over 4 percent. This result demonstrates that the bank faces severe enforcement frictions. The rest of this section formally analyzes this result in greater detail. First, in section (3.2.1), we analyze contractual frictions by testing how farmers that were selected by the *same lender* but were randomly assigned *different contracts* affected farmer repayment. Second, in section (3.2.2), we analyze information frictions by testing how farmers that received the *same contract* but were selected by *different lenders* affected farmer repayment.



Note: This graph shows overdue rates (percent) for each treatment arm in our study. The orange bars correspond to overdue rates for loans received directly from the bank, and the blue bars for loans that were guaranteed by the mill. The key inference from this figure is that direct bank loans had substantially higher overdue rates than the loans that were guaranteed by the mill.

3.2.1 Enforcement frictions

To understand whether the mill has superior enforcement technology than the bank, we test for whether the repayment rates are relatively higher for those farmers that were guaranteed by the mill. To ensure we are identifying the effect of differences in lenders' enforcement technology, we compare repayment rates between farmers that received different contracts (different residual claimant on loan proceeds) but were initially selected as creditworthy by the same lender.

In table (4), we report the results from regressing repayment on whether the farmer's loan was guaranteed by the mill, and additional controls.¹⁹ Columns 1 and 2 restrict attention to those farmers that were identified as creditworthy by only the mill (that is, farmers in group A). Columns 3 and 4 restrict attention to those farmers that were identified as creditworthy by both the mill and the bank (that is, farmers in group B). Finally, columns 5 and 6 include all farmers that were identified as creditworthy by the mill (that is, farmers in groups A and B). Columns 2, 4, and 6 include controls for farmer characteristics.

As previewed in figure (4), there is strong evidence that the mill has superior enforcement technology than the bank. Across all four specifications, those farmers that had a loan guarantee from the mill were around *60 percentage points* less likely to be overdue on their loan than those farmers that received a direct bank loan. The high rate of overdue loans for banks (and significantly higher than the interest rate on the loans) illustrates the severity of the banks' problem to offer rural loans. Moreover, the results are robust to including additional controls for farmer characteristics (columns 2, 4, and 6), separately estimating the results on the various classifications of farmer creditworthiness—farmers that were identified as creditworthy only by the mill (columns 1 and 2), by both the bank and the mill (columns 3 and 4), and estimating the results on all farmers that were identified as creditworthy by the mill (columns 5 and 6).

Interestingly, there are no significant differences in the repayment behavior between those farmers that were solely selected by the mill, and those farmers that were selected by both the mill and the bank, after controlling for contract type. This lack of difference

¹⁹More detail on the empirical strategy is provided earlier in section (2.2.1).

in repayment behaviour is evident by the small coefficient estimate for “creditworthy-only mill”, an indicator variable for whether the farmers were in group A—that is, those farmers that were only selected by the mill. Therefore, this surprising result suggests that from the set of borrowers chosen by the mill, the bank does not choose borrowers that are, on average, more likely to repay the bank. Specifically, you may expect that the bank’s screening technology would facilitate the bank to choose higher quality borrowers leading to higher repayments.

Table 4: Does the mill have superior enforcement technology than the bank?

	(1)	(2)	(3)	(4)
	Overdue	Overdue	Overdue	Overdue
Loan Guarantee	-0.59*** (0.071)	-0.62*** (0.070)	-0.57*** (0.10)	-0.55*** (0.10)
Sugarcane planted (acres)		-0.0034 (0.0024)		-0.00091 (0.0014)
Relative value of crop sales (decile)		0.018 (0.017)		-0.027 (0.021)
Educ. below high school (Y=1;N=0)		-0.0075 (0.071)		-0.015 (0.084)
Farm size (acres)		0.00071 (0.0011)		0.0012 (0.0011)
Years selling to the mill		-0.010** (0.0047)		-0.020* (0.011)
Sales to the mill (pct of total sales)		0.00032 (0.0011)		-0.0042** (0.0019)
Within 5km of the mill (Y=1;N=0)		0.018 (0.068)		0.19 (0.12)
Relative sales to the mill (decile)		0.0032 (0.020)		0.0041 (0.027)
Formal credit history (Y=1;N=0)		0.016 (0.075)		0.071 (0.080)
Prev. bank loan overdue (Y=1;N=0)		0.019 (0.058)		-0.063 (0.096)
Observations	129	129	63	63
Groups	A1 & A2	A1 & A2	B1 & B2	B1 & B2

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Table (13) in Appendix (5.1) describes the variable definitions.

Having found that the mill has superior enforcement technology, we investigate for *which farmers* the mill's enforcement technology is superior. To do so, in table (5), we supplement the regressions in table (4) with additional independent variables. Specifically, we interact the dummy variable 'whether a farmer had a loan guarantee', with the farmer's characteristics. Overall, the regressions suffer from a lack of power, which is not surprising given the small sample size. Nonetheless, there are two notable results. The mill was relatively better at enforcing repayment by those farmers that planted less sugarcane (coefficient on the regressor "Guarantee X Sugarcane planted" is consistently positive and statistically significant in some specifications) and those farmers that owned larger farms (coefficient on the regressor "Guarantee X Farm size" is consistently negative and statistically significant in some specifications). These results suggest that the mill has superior technology for those farmers with larger farms, and for whom sugar is a small part of the farmer's crop portfolio. Somewhat surprisingly, we do not see significant coefficients on some of the variables that you may expect to be correlated with the mill's superior technology, namely, those variables that suggest a stronger relationship between the mill and the farmer (such as the fraction of sales to the mill, short distance from the mill, and the tenure of the relationship with the mill).

Table 5: For which farmers does the mill have superior enforcement technology?

	(1)	(2)	(3)
	Overdue	Overdue	Overdue
Guarantee x Sugarcane planted	0.0057 (0.0071)	0.0054 (0.0057)	0.0087** (0.0036)
Guarantee x Relative value of crop sales (decile)	-0.023 (0.038)	0.16*** (0.047)	0.013 (0.030)
Guarantee x Educ. below high school	0.075 (0.16)	-0.00045 (0.21)	-0.057 (0.13)
Guarantee x Farm size (acres)	-0.0019 (0.0027)	-0.0074*** (0.0016)	-0.0042*** (0.0015)
Guarantee x Years selling to the mill	0.0074 (0.011)	0.036 (0.030)	0.0068 (0.0086)
Guarantee x Sales to the mill (pct of total sales)	-0.0013 (0.0026)	0.0095*** (0.0033)	0.0016 (0.0020)
Guarantee x Within 5km of the mill	-0.13 (0.16)	0.060 (0.30)	-0.12 (0.13)
Guarantee x Relative sales to the mill (decile)	0.014 (0.047)	-0.030 (0.052)	0.0041 (0.036)
Guarantee x Formal credit history	-0.048 (0.17)	-0.014 (0.16)	-0.0047 (0.13)
Creditworthy only Mill			-0.042 (0.066)
Constant	0.66** (0.26)	2.91*** (0.60)	0.96*** (0.25)
Observations	129	63	192
Groups	A	B	A & B
Additional Controls	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: We have interacted the farmer characteristics with an indicator variable for whether the farmer received a loan guarantee from the mill; these variables are prefixed with the word 'Guarantee', for instance, "Guarantee x Acres Sugarcane" refers to the variable "Loan Guarantee" interacted with the variable "Acres Sugarcane". In all the regressions we include the non-interacted variables (that is, Loan guarantee, acres sugarcane, value of crop sales, etc) but in the interest of space we have omitted the coefficient results. Table (13) in Appendix (5.1) describes the variable definitions.

3.2.2 Screening technology: Information frictions

To understand whether the mill has a superior information about farmer's creditworthiness than the bank, we test for whether the repayment rates are relatively higher for those farmers selected by the mill than for the bank. To ensure we are identifying the effect of differences in lenders' screening technology, we first compare repayment rates between farmers that received the same contract but were selected as creditworthy by different lenders.

In table (6), we report the results from regressing repayment on whether the farmer was chosen by the mill, and additional controls.²⁰ Columns 1 and 2 restrict attention to those farmers that received a guaranteed loan from the mill and were selected by the mill as creditworthy (that is, farmers in groups A1 and B1). Columns 3 and 4 restrict attention to those farmers that received a direct bank loan and were selected by the mill as creditworthy (that is, farmers in groups A2 and B2). Finally, columns 5 and 6 restrict attention to all the farmers that received a direct bank loan (that is, farmers in groups A2, B2, and C). Columns 2, 4, and 6, include controls for farmer characteristics.

Overall, we find there is little economic or statistical significance for the mill possessing superior information on farmer's creditworthiness than the bank. Across the set of regressions in table (6), the estimated casual effect from the mill's greater information on repayment rates varies around zero, with the highest estimated effect being increasing repayment by nearly eight percentage points (column 6) or reducing repayment by nearly five percentage points (column 3).

Analyzing the results for the control variables, it is evident that across the regressions there was little systemic differences by farmer characteristics. There are only two variables that are both statistically significant in some specifications and consistent in the coefficient's sign across the various specifications. The farmers that plant more sugarcane ("acres sugarcane") and have been selling for more years to the mill ("years selling to the mill"), on some specifications, had statistically significant higher repayments rate, suggesting these farmers are—observationally—more creditworthy.

²⁰More detail on the empirical strategy is provided earlier in section (2.2.1).

Table 6: Does the mill have superior information technology than the bank?

	(1)	(2)	(3)	(4)	(5)	(6)
	Overdue	Overdue	Overdue	Overdue	Overdue	Overdue
Creditworthy only Mill	0.027 (0.037)	-0.044 (0.030)	0.048 (0.12)	-0.038 (0.16)	-0.028 (0.10)	-0.078 (0.15)
Sugarcane planted (acres)		-0.00030 (0.00028)		-0.0090** (0.0037)		-0.0042 (0.0041)
Relative value of crop sales (decile)		0.0049 (0.012)		-0.0076 (0.028)		0.0067 (0.023)
Educ. below high school (Y=1;N=0)		-0.0058 (0.041)		0.051 (0.13)		-0.012 (0.12)
Farm size (acres)		-0.00020 (0.00017)		0.0040** (0.0015)		0.0023** (0.0011)
Years selling to the mill		-0.0067* (0.0037)		-0.013 (0.0090)		-0.014 (0.0086)
Sales to the mill (pct of total sales)		0.00026 (0.00074)		-0.0014 (0.0019)		-0.0011 (0.0018)
Within 5km of the mill (Y=1;N=0)		-0.031 (0.049)		0.084 (0.13)		0.12 (0.11)
Relative sales to the mill (decile)		0.0063 (0.015)		0.0021 (0.033)		-0.014 (0.026)
Formal credit history (Y=1;N=0)		0.0065 (0.049)		0.012 (0.12)		0.036 (0.11)
Prev. bank loan overdue (Y=1;N=0)		-0.018 (0.026)				
Observations	113	113	79	79	91	91
Groups	A1&B1	A1&B1	A2&B2	A2&B2	A2,B2&C	A2,B2&C
Contract: Loan Guarantee	Yes	Yes	No	No	No	No
Contract: Direct bank loan	No	No	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Table (13) in Appendix (5.1) describes the variable definitions.

Table (7) builds on the regressions in table (6). We increase the power of our tests by pooling the farmers across their assigned contract. Specifically, we include an indicator variable ("Loan Guarantee"), that is "one" if the loan was guaranteed by the mill. Columns 1 and 2 only include those farmers that the mill identified as creditworthy,

whereas, columns 3 and 4 include all farmers. The results in table (7) are consistent with the results in table (6), that is, holding the contract fixed, there is no statistical evidence that those farmers selected as creditworthy by the mill were more likely to repay their loans. In all the regressions, the coefficient for the indicator variable "creditworthy only mill" is not statistically significant and the causal estimate is close to zero.

Table 7: Pooled regressions: Does the mill have superior information technology than the bank?

	(1)	(2)	(3)	(4)
	Overdue	Overdue	Overdue	Overdue
Creditworthy only Mill	0.036 (0.053)	-0.038 (0.064)	0.0016 (0.051)	-0.058 (0.063)
Loan Guarantee	-0.59*** (0.058)	-0.59*** (0.057)	-0.62*** (0.054)	-0.61*** (0.054)
Sugarcane planted (acres)		-0.0023 (0.0016)		-0.0018 (0.0013)
Relative value of crop sales (decile)		0.0042 (0.013)		0.0089 (0.012)
Educ. below high school (Y=1;N=0)		0.012 (0.053)		-0.0047 (0.052)
Farm size (acres)		0.00097 (0.00065)		0.00083 (0.00053)
Years selling to the mill		-0.0090** (0.0041)		-0.0092** (0.0041)
Sales to the mill (pct of total sales)		-0.00061 (0.00089)		-0.00069 (0.00087)
Within 5km of the mill (Y=1;N=0)		0.053 (0.058)		0.059 (0.058)
Relative sales to the mill (decile)		-0.0052 (0.016)		-0.012 (0.014)
Formal credit history (Y=1;N=0)		0.026 (0.056)		0.027 (0.054)
Prev. bank loan overdue (Y=1;N=0)		0.011 (0.040)		0.0093 (0.039)
Observations	192	192	204	204
Groups	A & B	A & B	A, B & C	A, B & C
Contract	Both	Both	Both	Both

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Table (13) in Appendix (5.1) describes the variable definitions.

3.3 Stage 3: Final survey exploring the large differences in repayment rates

In this section, we present the findings from a final survey that investigates why farmers did or did not repay the loan. For farmers to repay their loans they must be both *able* and *willing* to repay the loan. If farmers' inability to pay (for example, due to a poor sugar harvest) was the primary reason for high default rates, we would expect to see similar default rates across the randomly allocated contractual treatments because there's no reason to think that the farmers that were randomly allocated to the direct bank loan would have different abilities to pay. However, in section (3.2.1), we found strikingly large differences across treatment groups, specifically farmers repaid the bank markedly less often than the mill. Taking these results together suggests that the net benefits of repaying the mill must be higher than repaying the bank. This section explores possible reasons for this result. ²¹

3.3.1 Differences in collection technology

Table (8) presents the results of the question, "How did the lender encourage repayment?" As is quickly evident, the sugar mill and the bank used substantially different methods to encourage repayment. The sugar mill leveraged its position in the farmer's production process to procure repayment at the time of milling the sugarcane. Whereas, the bank relied on predominantly calling the farmer (57 percent of all loans), and in some cases also visiting the farmer (an additional 38 percent of loans).

Visiting individual farmers is expensive because (i) the bank branch staff is located in a nearby city of Hyderabad some 22 miles (35 kilometers) away from the sugar mill and (ii) the sugar farmers are on average dispersed over 17 miles radius (30 kilometers) and the terrain is poorly connected with narrow dirt roads. For example, the remotest farm from the sugar mill is located at 28 miles (45 kilometers). Even though the bank visited many farmers, in section (3.2.1), we found that farmers were 60 percentage points more likely to

²¹As described earlier, this part of the experiment was added after the experiment after observing a sufficient period of non repayment, as we wanted to further explore the reasons farmers defaulted.

be overdue on direct bank loans than loans guaranteed by the mill. Therefore, even with the careful attention of the bank loan officer, the loan officer was unable to get high—or even close to profitable—repayment rates. In contrast, the mill’s collection process was relatively inexpensive as the mill collected repayments directly from the farmer at the time of milling the sugarcane.

Table 8: Final survey question: How did the lender collect repayment?

	Bank (mean)	Mill (mean)	Diff.	Standard Error
Collected at Mill	0.00	0.92	-0.92***	0.04
In-person Visit and Call	0.38	0.05	0.33***	0.08
Only Call	0.57	0.03	0.54***	0.08
Other	0.05	0.00	0.05	0.03
Observations	100			

T-test significance level denoted by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents the results from a final survey that was administered to farmers after the due date of their loans. The first two columns present the fractions of farmers that reported how the bank (column 1) or the mill (column 2) collected the loan repayments. The column labelled ‘Bank’ reports results from those farmers that received a direct bank loan (that is, those farmers in groups A2, B2, and C) and the column labelled ‘Mill’ reports results from those farmers that received a bank loan that was guaranteed by the mill (that is, those farmers in groups A1 and B1). Standard errors for the t-test are computed assuming that the variances of the variable of interest may differ between the bank and the mill.

3.3.2 Stated reasons for repaying on time

Table (9) presents the results of the question, “Why did you repay on time?” Overall, farmers stated relatively similar reasons across both treatments with most farmers stating the importance of ‘responsible behavior’, that is, the farmers’ felt that repaying a loan is an obligation because it is the ‘right’ thing to do. In terms of some of the differences, relatively more of the farmers with loans guaranteed by the mill suggested that they were financially capable of repaying. Interesting, even though the mill enforced the farmer contract by collecting the loan repayment at the time of milling the sugar (see table (8)), very few (less than ten percent of farmers that received a loan guaranteed by the mill) farmers noted that was the reason they repaid on time.

Table 9: Final survey question: Why did you repay on time?

	Bank (mean)	Mill (mean)	Diff.	Standard Error
Responsible behaviour	0.62	0.40	0.22*	0.10
Good crop and financial liquidity	0.05	0.22	-0.17*	0.07
Future relationship	0.12	0.20	-0.08	0.07
Financial cost	0.14	0.09	0.05	0.07
Deducted at milling	0.00	0.09	-0.09*	0.04
Other	0.07	0.00	0.07	0.04
Observations	97			

T-test significance level denoted by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents the results from a final survey that was administered to farmers after the due date of their loans. The column labelled 'Bank' reports results from those farmers that received a direct bank loan (that is, those farmers in groups A2, B2, and C) and the column labelled 'Mill' reports results from those farmers that received a bank loan that was guaranteed by the mill. Standard errors for the t-test are computed assuming that the variances of the variable of interest may differ between the bank and the mill.

3.3.3 The perceived consequences of not repaying

Table (10) presents the results of the question: "What do you think the consequences of not repaying the loan would be?" Interestingly, if the farmer did not repay the loan guaranteed by the mill, they perceived the main cost as jeopardizing their future relationship with the mill, whereas, those farmers with a direct bank were more concerned about higher financing costs in the future. Not surprisingly, given the limited effectiveness of the legal system in rural Pakistan, in both treatments, the threat of legal actions was not rated highly.²²

The results to this question reiterate the importance of relational lending. Farmers with loans that were guaranteed by the mill were clearly cognizant of the potential hit to their relationship with the mill if they were to default on the loan. By the mill guaranteeing the loan, the mill was bundling its production process and intermediating in the credit process. The farmers' answers to this question show that the farmers were aware of this

²²Djankov et al. [2003] documents the difficulties of evicting a tenant in various legal systems. In Pakistan, they found that it took 365 days to evict a tenant, versus a median of 180 days in legal systems with similar origins, or 49 days in the United States.

bundling and the potential adverse impact on their relationship with the mill. In contrast, those farmers with direct bank loans were relatively more concerned with potential adverse financial impacts from not repaying the loan (higher future interest rates, and/or penalties).

Table 10: Final survey question: What do you think the consequences of not repaying the loan would be?

	Bank (mean)	Mill (mean)	Diff.	Standard Error
Jeopardize future relationship	0.44	0.70	-0.25*	0.10
Higher future interest rates/ penalties	0.44	0.15	0.29**	0.09
Legal actions	0.12	0.15	-0.04	0.07
Observations	98			

T-test significance level denoted by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents the results from a final survey that was administered to farmers after the due date of their loans. The column labelled 'Bank' reports results from those farmers that received a direct bank loan (that is, those farmers in groups A2, B2, and C) and the column labelled 'Mill' reports results from those farmers that received a bank loan that was guaranteed by the mill. Standard errors for the t-test are computed assuming that the variances of the variable of interest may differ between the bank and the mill.

3.3.4 Summary of the final survey

Taking the results together from the final survey we learn that the mill has two comparative advantages in ensuring high repayments. First, the mill is able to leverage its role in milling the sugarcane to physically collect repayments in an efficient and cheap manner. Second, finance emphasises the importance of relational lending to ensure repayment; in this context, it's clear that the farmers highly value the importance of their relationship with the mill, and do not want to jeopardize their interlinked relationship.

4 Conclusion

On a promising note for increasing credit provision, leveraging less traditional credit intermediaries may increase credit efficiency and improve credit access. The non-traditional

intermediary (the mill) was willing to guarantee loans to an *additional* 184 farmers that the bank was unwilling to lend to—more than doubling the number of farmers that were eligible for loans. We find that the mill’s key advantage over traditional forms of lending for improving credit access is the mill’s superior enforcement technology. Moreover, the mill’s enforcement technology was sufficiently effective that the trilateral credit relationship between the farmer, the mill, and the bank was both incentive-compatible and economically sustainable. In the experiment, both the mill and the bank were able to make positive profits and the farmer received low interest rates.²³ The success of this financial innovation is evident with the bank, following our experiment, expanding the trial of using mills as loan guarantors to other two other mills, and earmarking an expansion of the scheme from the initial \$0.7 million to \$35 million in credit disbursements.

In terms of screening, somewhat surprisingly, the mill and the bank had similar effectiveness at screening farmer’s creditworthiness, that is, the mill did not seem to have better screening technology even though the farmer has had significantly more interaction with the mill than the bank.

On a less promising note for credit provision, a substantial number of farmers were classified as not ‘creditworthy’. Out of 528 farmers that requested loans more than 200 were not selected as creditworthy by either lender. Moreover, this statistic is likely an underestimate for the number of farmers that would be excluded because the high level of loan defaults for the bank suggests that the bank’s selection criteria was insufficiently rigorous given the bank’s enforcement capabilities.

The key additional finding is that banks struggle with high farmer defaults even from farmers with large farms or have more education. The key constraint seems the banks’ limited capacity or high cost to enforce the credit contract. Additional policies and methods that might alleviate this constraint are worth exploring; prior literature have examined whether additional pledgeable collateral, such as increasing property rights, or using

²³The rate of overdue loans on the guaranteed loans was slightly above the fee earned by the mill—therefore there was no net profit for guaranteeing the loans for the mill (and likely some additional administrative costs). However, guaranteeing the loans may indirectly boost the mill’s profit because it may have led farmers to grow more sugarcane, in turn, leading to additional income for the mill from the mill’s margin from processing the farmer’s sugarcane into sugar

social capital may increase the farmers' incentive to repay, but more creative solutions are worth considering. One avenue that is gaining greater attention is the greater use of transactional data—and a digital footprint more broadly (see [Berg et al. \[2018\]](#), [Bharadwaj et al. \[2019\]](#), [Frost et al. \[2019\]](#)).

We should also touch upon some of the drawbacks of our study. Our experiment mostly only examined a one-shot game, that is, a big Pakistani bank offered loans, but did not explicitly offer future bank loans. A key finding of the credit literature is that relationship lending (see [Boot \[2000\]](#) for an overview), specifically the promise of a future loans, increases consumer repayment. Therefore, our experiment may have led to a higher default rate than we may have in equilibrium because we did not explicitly link the repayment of this bank loan to future bank loans. In mitigation, the farmers that successfully repaid (either the direct bank loan or the guaranteed bank loan) would have a positive mark in their credit report that should increase their ability to procure future loans and as the results of the experiment's stage 3 demonstrates borrowers were cognizant of the potential consequences from non-repayment.

Our study may also over-estimate the number of borrowers the bank was willing to lend to. A key design choice of our experiment was that the bank was required to do due diligence on all the borrowers that requested a bank loan. It is highly possible that the screening cost for all farmers that request a loan is too high given the potential expected returns from those borrowers that actually procure loans, thereby causing a missing financial market.

References

- I. Aleem. Imperfect information, screening, and the costs of informal lending: a study of a rural credit market in pakistan. *The World Bank Economic Review*, 4(3):329–349, 1990.
- A. V. Banerjee and E. Duflo. *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs, 2011.
- A. V. Banerjee, E. Breza, E. Duflo, and C. Kinnan. Do credit constraints limit entrepreneurship? heterogeneity in the returns to microfinance. *Heterogeneity in the Returns to Microfinance (September 1, 2017)*. *Global Poverty Research Lab Working Paper*, (17-104), 2017.
- T. Beck, L. Klapper, and J. C. Mendoza. The typology of partial credit guarantee funds around the world. *Journal of Financial Stability*, 6(1):10–25, 2010.
- T. Berg, V. Burg, A. Gombović, and M. Puri. On the rise of fintechs—credit scoring using digital footprints. Technical report, National Bureau of Economic Research, 2018.
- T. Besley. How do market failures justify interventions in rural credit markets? *The World Bank Research Observer*, 9(1):27–47, 1994.
- P. Bharadwaj, W. Jack, and T. Suri. Fintech and household resilience to shocks: Evidence from digital loans in kenya. Technical report, National Bureau of Economic Research, 2019.
- J. E. Blumenstock, N. Eagle, and M. Fafchamps. Airtime transfers and mobile communications: Evidence in the aftermath of natural disasters. *Journal of Development Economics*, 120:157–181, 2016.
- A. W. Boot. Relationship banking: What do we know? *Journal of Financial Intermediation*, 2000.
- S. R. Boucher and M. R. Carter. Risk rationing and activity choice in moral hazard constrained credit markets. Technical report, 2001.

- S. R. Boucher, M. R. Carter, and C. Guirkingner. Risk rationing and wealth effects in credit markets: Theory and implications for agricultural development. *American Journal of Agricultural Economics*, 90(2):409–423, 2008.
- A. Braverman and J. L. Guasch. Rural credit markets and institutions in developing countries: Lessons for policy analysis from practice and modern theory. *World development*, 14(10-11):1253–1267, 1986.
- A. Braverman and J. E. Stiglitz. Sharecropping and the interlinking of agrarian markets. *The American Economic Review*, 72(4):695–715, 1982.
- A. Chaia, A. Dalal, T. Goland, M. J. Gonzalez, J. Morduch, and R. Schiff. Half the world is unbanked, 2009.
- M. A. Choudhary and A. K. Jain. How public information affects asymmetrically informed lenders: Evidence from a credit registry reform. *Journal of Development Economics*, 2020.
- D. Collins, J. Morduch, S. Rutherford, and O. Ruthven. *Portfolios of the poor: how the world's poor live on \$2 a day*. Princeton University Press, 2009.
- R. Cull, R. J. Cull, A. Demirgüç-Kunt, and J. Morduch. *Banking the World: Empirical Foundations of Financial Inclusion*. MIT Press, 2013.
- A. Dasgupta. Reports on informal credit markets in india: Summary. *New Delhi: National Institute of Public Finance and Policy*, 1989.
- S. Dercon. Income risk, coping strategies, and safety nets. *The World Bank Research Observer*, 17(2):141–166, 2002.
- S. Djankov, R. La Porta, F. Lopez-de Silanes, and A. Shleifer. Courts. *The Quarterly Journal of Economics*, 118(2):453–517, 2003.
- J. Frost, L. Gambacorta, Y. Huang, H. S. Shin, and P. Zbinden. Bigtech and the changing structure of financial intermediation. *Economic Policy*, 2019.

- P. Ghosh and D. Ray. Information and enforcement in informal credit markets. *Economica*, 83(329):59–90, 2016.
- X. Giné. Access to capital in rural thailand: An estimated model of formal vs. informal credit. *Journal of Development Economics*, 96(1):16–29, 2011.
- A. Green. Credit guarantee schemes for small enterprises: An effective instrument to promote private sector-led growth? sme technical working paper no. 10, unido, vienna, austria. 2003.
- D. B. Gross and N. S. Souleles. An empirical analysis of personal bankruptcy and delinquency. *Review of financial studies*, 15(1):319–347, 2002.
- K. Hoff and J. E. Stiglitz. Introduction: Imperfect information and rural credit markets: Puzzles and policy perspectives. *The world bank economic review*, 4(3):235–250, 1990.
- K. Hoff and J. E. Stiglitz. Moneylenders and bankers: price-increasing subsidies in a monopolistically competitive market. *Journal of Development Economics*, 55(2):485–518, 1998.
- R. Innes. Financial contracting under risk neutrality, limited liability and ex ante asymmetric information. *Economica*, pages 27–40, 1993.
- W. Jack, M. Kremer, J. De Laat, and T. Suri. Borrowing requirements, credit access, and adverse selection: Evidence from kenya. Technical report, National Bureau of Economic Research, 2016.
- D. M. Jaffee and T. Russell. Imperfect information, uncertainty, and credit rationing. *The Quarterly Journal of Economics*, 90(4):651–666, 1976. URL <http://qje.oxfordjournals.org/content/90/4/651.short>.
- D. Karlan and J. Morduch. Access to finance. In *Handbook of development economics*, volume 5, pages 4703–4784. Elsevier, 2010.

- D. Karlan and J. Zinman. Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica*, 77(6):1993–2008, 2009a. URL <http://onlinelibrary.wiley.com/doi/10.3982/ECTA5781/abstract>.
- D. Karlan and J. Zinman. Expanding credit access: Using randomized supply decisions to estimate the impacts. *Review of Financial studies*, page hhp092, 2009b.
- A. I. Khwaja and A. Mian. Do lenders favor politically connected firms? rent provision in an emerging financial market. *The Quarterly Journal of Economics*, 120(4):1371–1411, 2005.
- J. Levin. Relational incentive contracts. *American Economic Review*, 93(3):835–857, 2003.
- P. Maitra, S. Mitra, D. Mookherjee, and S. Visaria. Decentralized targeting of agricultural credit programs: Private versus political intermediaries. Technical report, National Bureau of Economic Research, 2020.
- S. A. Matthews. Renegotiating moral hazard contracts under limited liability and monotonicity. *Journal of Economic Theory*, 97(1):1–29, 2001.
- Pakistan Bureau of Statistics. Pakistan social and living standards measurement. 2015.
- S. Rutherford. *The poor and their money*. Oxford University Press New Delhi, 2000.
- A. Solis. Credit access and college enrollment. *Journal of Political Economy*, 125(2):000–000, 2017.
- J. E. Stiglitz and A. Weiss. Credit rationing in markets with imperfect information. *The American economic review*, 71(3):393–410, 1981.
- C. Udry. Risk and insurance in a rural credit market: An empirical investigation in northern nigeria. *The Review of Economic Studies*, 61(3):495–526, 1994.

5 Appendix

5.1 Variable and group definitions

Table 11: Group definitions for stage 1 of the experiment

Group	Definition
A	Screened as creditworthy by only the mill (Creditworthy only mill)
B	Screened as creditworthy by both the mill and the bank (Creditworthy both)
C	Screened as creditworthy by only the bank (Creditworthy only bank)
D	Screened as creditworthy by no lender (Not creditworthy)

Table 12: Group allocations for stage 2 of the experiment

		Loan guaranteed by mill (1)	Direct bank loan (2)
Group			
Creditworthy only mill	(A)	A1	A2
Creditworthy both	(B)	B1	B2
Creditworthy only bank	(C)	omitted	C
Not creditworthy	(D)	omitted	omitted

As part of the experiment design, the mill was only willing to guarantee those farmers that the mill had screened as creditworthy (hence guaranteed bank loans were only offered to some farmers in groups A and B). Whereas, the bank was willing to offer direct bank loans to any farmer that had been denoted as creditworthy by the bank or the mill (hence direct bank loans were offered to some farmers in groups A, B, and C).

Table 13: Variable definitions

Variable	Value type	Description
Loan guarantee	<i>Binary</i>	Labels the type of contract the farmer receives: "0" equals direct bank loan and "1" equals loan guaranteed by the mill
Sugarcane planted (acres)	<i>Continuous</i>	Total number of acres of sugarcane planted by the farmer in 2016-17
Income from agriculture (percent)	<i>Continuous</i> <i>between 0 and 100</i>	Farmer's income from agriculture as a percent of total income
Income from sugar (percent)	<i>Continuous</i> <i>between 0 and 100</i>	Farmer's income from sugar as a percent of total income
Relative value of crop sales (deciles)	<i>Discrete, values from 1 to 10</i>	A (decile) ranking of each farmers' total crop sales in rupees, with 10 being the most sales
Educ. below high school	<i>Binary</i>	Education level of the farmer: "1" if the level of the farmers' highest education is below high school, "0" otherwise
Farm size (acres)	<i>Continuous</i>	Total number of acres owned for plantation by the farmer in 2016-17
Years selling to the mill	<i>Discrete</i>	Number of years the farmer has been selling to the mill
Sales to the mill (percent of total sales)	<i>Continuous</i> <i>between 0 and 100</i>	Farmer's sales (by value) to the mill as a percent of total farmer sales.
Within 5km of the mill	<i>Binary</i>	Value equal to "1" if the farmer's home is within 5 Km miles of the mill, "0" otherwise
Relative sales to the mill (decile)	<i>Discrete, values from 1 to 10</i>	A (decile) ranking of each farmers' fraction of sales of sugarcane to the mill, with 10 being the most sales
Own-"farm equipment"	<i>Binary</i>	Labels whether the household owns "farm equipment": "1" yes, "0" otherwise.
Formal credit history	<i>Binary</i>	Labels whether the farmer appears in Pakistan's formal credit registry: "1" equals formal credit history exists, "0" otherwise.
Prev. bank loan overdue	<i>Binary</i>	Labels whether the farmer had a previous overdue loan in Pakistan's formal credit registry: "1" equals yes, "0" otherwise.

5.2 Baseline characteristics and balance between treatment arms

Table 14: Baseline characteristics and balance between farmers in groups A1 and group A2

	Diff.	
Sugarcane planted (acres)	0.991	(0.28)
Income from agriculture (percent)	-2.284	(-0.72)
Income from sugar (percent)	-2.542	(-1.04)
Relative value of crop sales (decile)	-0.733	(-1.60)
Educ. below high school (Y=1;N=0)	-0.0374	(-0.47)
Farm size (acres)	-9.831	(-1.87)
Years selling to the mill	0.576	(0.46)
Sales to the mill (pct of total sales)	-3.404	(-0.62)
Within 5km of the mill (Y=1;N=0)	-0.0567	(-0.72)
Relative sales to the mill (decile)	-0.465	(-1.20)
Formal credit history (Y=1;N=0)	-0.0482	(-0.67)
Prev. bank loan overdue (Y=1;N=0)	0.0151	(0.30)
Own-tractor (Y=1;N=0)	0.0307	(0.49)
Own-thresher (Y=1;N=0)	-0.0338	(-0.56)
Own-blade (Y=1;N=0)	0.00103	(0.02)
Own-cultivator (Y=1;N=0)	-0.00887	(-0.14)
Own-raja (Y=1;N=0)	-0.0337	(-0.52)
Own-gobal (Y=1;N=0)	0.00103	(0.02)
Own-bundmaker (Y=1;N=0)	-0.0142	(-0.26)
Own-harvesting machine (Y=1;N=0)	0.0250	(0.86)
Observations	168	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table reports differences in summary statistics (mean) of various farmer characteristics between farmers in group A1 and group A2 (where a positive number indicates a larger number for farmers in group A1).

Table 15: Baseline characteristics and balance between farmers in groups B1 and B2

	Diff.	
Sugarcane planted (acres)	2.473	(0.32)
Income from agriculture (percent)	1.807	(0.41)
Income from sugar (percent)	-4.868	(-1.68)
Relative value of crop sales (decile)	0.694	(1.12)
Educ. below high school (Y=1;N=0)	0.196	(1.75)
Farm size (acres)	1.297	(0.11)
Years selling to the mill	0.0459	(0.04)
Sales to the mill (pct of total sales)	-10.85	(-1.49)
Within 5km of the mill (Y=1;N=0)	-0.0446	(-0.52)
Relative sales to the mill (decile)	0.207	(0.40)
Formal credit history (Y=1;N=0)	-0.0944	(-0.84)
Prev. bank loan overdue (Y=1;N=0)	-0.0612	(-1.77)
Own-tractor (Y=1;N=0)	-0.0453	(-0.43)
Own-thresher (Y=1;N=0)	0.0376	(0.35)
Own-blade (Y=1;N=0)	-0.00319	(-0.03)
Own-cultivator (Y=1;N=0)	-0.0140	(-0.13)
Own-raja (Y=1;N=0)	0.0485	(0.44)
Own-gobal (Y=1;N=0)	0.0580	(0.54)
Own-bundmaker (Y=1;N=0)	0.0568	(0.56)
Own-harvesting machine (Y=1;N=0)	-0.0300	(-0.64)
Observations	81	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 16: Baseline characteristics and balance between the farmers approved for loans who took loans and those farmers that did not take loans

	Diff.	
Sugarcane planted (acres)	6.778	(1.41)
Income from agriculture (percent)	3.469	(1.31)
Income from sugar (percent)	-0.0952	(-0.05)
Relative value of crop sales (decile)	0.683	(1.69)
Educ. below high school (Y=1;N=0)	0.00572	(0.08)
Farm size (acres)	9.170	(1.47)
Years selling to the mill	0.472	(0.40)
Sales to the mill (pct of total sales)	-3.422	(-0.66)
Within 5km of the mill (Y=1;N=0)	0.0531	(0.82)
Relative sales to the mill (decile)	-0.0547	(-0.15)
Formal credit history (Y=1;N=0)	-0.0621	(-0.97)
Prev. bank loan overdue (Y=1;N=0)	0.0310	(0.68)
Own-tractor (Y=1;N=0)	0.0613	(1.00)
Own-thresher (Y=1;N=0)	0.000817	(0.01)
Own-blade (Y=1;N=0)	0.0507	(0.81)
Own-cultivator (Y=1;N=0)	0.0743	(1.18)
Own-raja (Y=1;N=0)	0.0784	(1.23)
Own-gobal (Y=1;N=0)	0.0891	(1.42)
Own-bundmaker (Y=1;N=0)	0.126*	(2.14)
Own-harvesting machine (Y=1;N=0)	0.0212	(0.71)
Observations	276	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 17: Requested loans, offered loans, and final take-up

	Classified as creditworthy by only the mill (A)		Classified as creditworthy by the mill and the bank (B)		Classified as creditworthy by only the bank (C)
	Guaranteed (A1)	Direct (A2)	Guaranteed (B1)	Direct (B2)	Direct (C)
Requested	185		88		51
Offered	101	67	49	32	27
Take-up	75	54	38	25	12
Take-up (%)	74%	81%	78%	78%	44%

Note: This table outlines the number of individuals that requested a loan and were classified as creditworthy by either the mill, the bank, or both (first row in the table, labelled “Requested”). Of these farmers that requested a loan, due to constraints on the total study size, only a subset of these farmers were offered a loan (the second row in the table, labelled “Offered”). There was some attrition with only some farmers subsequently taking a loan (the third and fourth rows in the table, labelled “Take-up” and “Take-up (%)”, where “take-up (%)” is the fraction of farmers that were offered a loan that eventually took the loan). For farmers in the various treatment arms of groups A and B, we find very similar attrition rates. However, we find a larger attrition rate for farmers in group C.