Board of Governors of the Federal Reserve System

International Finance Discussion Papers

ISSN 1073-2500 (Print)
ISSN 2767-4509 (Online)

Number 1363

December 2022

Credit access and relational contracts: An experiment testing informational and contractual frictions for Pakistani farmers

M. Ali Choudhary and Anil K. Jain

Please cite this paper as:
Choudhary, M. Ali and Anil K. Jain (2022). "Credit access and relational contracts: An experiment testing informational and contractual frictions for Pakistani farmers," International Finance Discussion Papers 1363. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/IFDP.2022.1363.

Credit access and relational contracts: An experiment testing informational and contractual frictions for Pakistani farmers

M. Ali Choudhary*

State Bank of Pakistan &

Center for Economic Performance, London School of Economics

Anil K. Jain

Federal Reserve Board

December 2022

Abstract

Credit access is limited in rural areas, especially in developing economies. Using a novel two-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 by randomly varying loan contractual terms across farmers and find that enforcement technology is the primary friction. Third, using an endline survey, we document that farmers tend to correctly identify the financial consequences of non-repayment.

*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, Andrew Bird, 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 AFBC meeting (virtual), American Economic Association (virtual), Colombian Central Bank, IFABS conference (Angers), MWIEDC conference (Purdue), 8th Relational Contract Workshop (virtual), IGC-SBP Conference (Karachi) and AD Conference (LSE, Lahore). We would like to acknowledge the exceptional research assistance from Mukhtar ul Hasan and Khurram Ashfaque 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.

Fourth, our results suggest one possible solution to overcome this financial friction—a motivated and interlinked intermediary and the use of relational contracts.

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 universal access remains severely lacking, especially in rural areas in developing countries. 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]). Therefore, considering both the virtues of financial access against the dearth of financial opportunities raises an important academic and policy question: What are the key financial obstacles to greater financial access and are there innovative methods to overcome these obstacles? This paper focuses on information asymmetries and enforcement frictions that limit credit access in rural Pakistan.

The existing literature outlines various theories for the cause of the lack of formal credit access in developing countries 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 as a key financial friction that limits greater credit access. Moreover, our experiment suggests that one possible solution to overcome this friction—a motivated and interlinked intermediary and the use of relational contracts.

To identify reasons for limited bank lending in rural areas, we use a novel two-stage experimental design with an endline survey and three key stakeholders: a set of sugarcane farmers, a non-traditional 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 (the sugar mill) would be willing to service. Specifically, we ask these institutions to screen a set of sugarcane farmers that have signalled an interest in procuring a loan. For the non-traditional intermediary, we ask whether they would be willing to guarantee a bank loan to the farmer. For the bank, we ask whether they would

be willing to offer a direct loan to a farmer (that is, with no guarantee from any other party). 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 they positively screened.

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—informal agreements sustained by the value of the future relationships (Baker et al. [2002])—and relationship banking suggests that the non-traditional intermediary (the mill) will have superior enforcement technology (Levin [2003] and Boot [2000]). Furthermore, the theory of interlinked relationships (Braverman and Stiglitz [1982]) suggests that the mill has superior information than the bank for lending to the farmers. Conversely, given the lack of credit specialization, the non-traditional intermediary is likely to be hampered by higher funding costs.

In our experiment, using the results of the first stage of the experiment, we divide farmers into groups according to whether the bank was willing to lend to the farmer and whether the mill was willing to guarantee a loan to the farmer. Subsequently, we randomize the loan's contractual terms across these groups.[1] 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 as creditworthy by the *same* institution, but get *different* loan contracts, we aim to isolate differences in enforcement technology. By comparing the loan outcomes for farmers that are screened as creditworthy by the *different* institutions, but get the *same* loan contract, we aim to isolate differences in screening 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 loan terms, as well as, randomizing contractual terms before the loan's renewal.

In the final stage, we implement an endline survey to uncover the reasons for the superior performance of loans that were guaranteed by the mill. Specifically, we interview the farmers to understand how the institutions tried to collect repayment and to understand

[1]Specifically, using the results in the first stage of the experiment, we can classify farmers into four mutually exclusive groups: those farmers that (i) the bank was *unwilling* to lend and the mill was *unwilling* to guarantee a loan, (ii) the bank was *unwilling* to lend and the mill was *willing* to guarantee a loan, (iii) the bank was *willing* to lend and the mill was *unwilling* to guarantee a loan, and (iv) both the bank was *willing* to lend and the mill was *willing* to guarantee a loan.

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 guarantee more farmers (52 percent), there is still a sizeable majority that is deemed not creditworthy by either institution (39 percent).[2]

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 (that is, the mill guaranteed the bank loan to the farmer) than for those loans where the bank was the residual claimant (that is, the bank made a direct loan to the farmer. 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 deemed creditworthy by different institutions 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 an endline 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, consistent with the predictions of relational contract theory, 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. To ensure the mill is not responsible for losses on loans that the mill guaranteed, the mill collects the loan payment at the time of milling the farmer's sugarcane. In contrast, the bank uses a mixture of messages to the farmer and costly visits to the farmer at 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? 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])—

[2]Note these probabilities do not need to sum to 100 percent since the bank and the intermediary both denoted some farmers as creditworthy.

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]). Moreover, as 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. 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 a key financial friction for bank lending in rural areas is banks' inability to enforce the loan contract. Further one solution that reduces this friction—and importantly to potentially economically sustainable levels—is through leveraging an interlinked intermediary (in this case, the sugar mill) with strong pre-existing economic relationships. 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. At 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.

Our results contribute to the growing empirical work in developing countries that find that strategic default is a key binding constraint for lending (Blouin and Macchiavello [2019]) and demonstrate that social and economic relationships can overcome this constraint. The closest paper to our work is Bryan et al. [2015] that uses an innovative randomized control trial to separate peer information from peer enforcement effects in consumer lending. By randomizing (ex ante and ex post) referral bonuses for loan approvals and loan repayments, similar to our paper, they find strong peer enforcement effects but little evidence for peer information. However, the magnitude of the estimated effect is substantially smaller. Specifically, they find that a small referral bonus (less than 3 percent of the total loan size), which is contingent on the referred borrower repaying the loan, increased repayment rates by only 10 percentage points. In our setting, the bank induces

substantially larger incentives for the mill to successfully enforce the contract (the mill is the residual claimant for any loss on the loan contract therefore the mill's enforcement incentive is both the loan's principal and interest, so around 113 percent of the loan size) with much larger increases in repayment (close to 60 percentage points). Moreover, the bank's cost to induce these stronger incentives and higher repayment is only 2 percent of the loan size (the size of the loan guarantee paid by the bank to the mill).

Our paper also contributes to the growing literature on the effectiveness of relational contracts in developing countries. Developing countries have significantly weaker economic institutions, (Acemoglu et al. [2005]), for instance, weaker creditor protections cause banks to issue smaller loans, shorter maturities, and higher interest rates (Bae and Goyal [2009]). To overcome these friction there is growing attention on the role of relational contracts. Macchiavello and Morjaria [2021] studies competition and relational contracts in the Rwandan coffee industry and find that greater competition impedes the use of relational contracts, and crucially, to such an extent that all parties are made worse off. Our paper directly compares the effectiveness of relational contracts between a farmer and a mill with a formal contract between a farmer and a bank. Consistent with the results in Macchiavello and Morjaria [2021], in our setting, we find that the relational contract between the mill and the farmer is far superior to that of the formal contract between the bank and the farmer. More generally, there is a long literature describing how relationships can overcome weak contractual protections in developing countries, such as McMillan and Woodruff [1999] in the Vietnamese informal credit market and Banerjee and Duflo [2000] in the Indian software market.

The results of our experiment highlight one possible avenue for overcoming the lack of formal financing in rural areas in developing countries—a key problem that has been highlighted in the academic and policy literature. For instance, 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 informal institutions often suffer from high prices, limited supply, poor resilience to financial shocks, and may exacerbate economic inequality.[3]

[3]In 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 interper-

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). However, there is emerging evidence for some of the drawbacks for wider use of microfinance. Some of the innovations that facilitate cheaper financing, such as regular, frequent repayments and group repayments (Feigenberg et al. [2013]) likely reduces the benefits of credit access. Microfinance is less optimal for large, seasonal borrowings—for example, farmers borrowing for sowing season—because repayments are usually required to start immediately. Moreover, some microfinance requires group lending, and consequently does not necessarily facilitate equal access and possibly rules out some of the least socially connected individuals. Other forms of formal financing are also being trialled such as 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 that leverages the monopsony power of a mill.

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.

2 Research Design

This paper reports the results from a two-stage field experiment with an endline survey in Pakistan. In the first stage, we analyze differences in screening decisions between different institutions. In the second stage, by randomly varying the farmer's loan contractual terms, we identify the relative importance of different institution's information and enforcement technology for determining loan repayment. Finally, we survey the farmers to understand why they did or did not repay their loans.

sonal 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.

Our experiment started in June 2016 in 60 villages in the district of Matiari, Sindh, Pakistan.[4] 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.[5] 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).

2.1 Key details on sugarcane farming and relational contracts

The study of the Pakistani sugarcane industry is well suited to analyze the effectiveness of relational contracts to overcome informational and enforcement frictions in credit markets for economic and logistical reasons.

The key economic reason to analyze sugarcane farming is that sugarcane mills are particularly well-suited to enforcing contracts. The high fixed cost of building mills ensures there are only a small number in any particular area (in the area we conduct our experiment, Matiari district, there are only two mills). Moreover, sugarcane must be quickly processed after harvesting (even if refrigerated, sugarcane can, at most, be stored for two weeks). Therefore, the limited competition for a farmer's sugarcane and the necessity to quickly sell the harvested sugarcane, ensures that the mills have significant market power

[4]We use the word village to describe what are known as "deh" in the local language, Sindhi. Formally, deh are areas that have at least one of three key characteristics (i) a separate record-of-rights, (ii) have been separately assessed to land revenue, or (iii) Pakistan's Board of Revenue declare to be a Deh.

[5]The 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.

at the time of purchase.

The key logistical reason to analyze sugarcane farming in Pakistan is the time from sowing to harvesting sugarcane is relatively short (only 12 to 18 months) thereby facilitating a faster experiment that is both cheaper and more likely to have lower attrition. Moreover, the authors have access to the official Pakistani credit registry and subsequently can combine data from the experiment with official farmer credit history.

Appendix (A.1) provides additional background detail on the (i) sugar industry in Pakistan; (ii) cultivation, harvesting, and processing of sugarcane; and (iii), sugar farming in Matiari district.

2.2 Experiment design: Stage 1–Testing differences in designations of creditworthiness

To understand how different institutions 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. The mill was asked whether they would guarantee a loan to the farmer.[6] Importantly, even though both institutions are asked to assess the "creditworthiness" of the farmers, this classification will vary across institution. For the bank, creditworthiness is defined as whether the farmer would repay the bank. For the mill, creditworthiness is defined as whether the farmer will repay the bank conditional on the mill guaranteeing the loan; that is, would the farmer be willing to potentially renege on the relational contract between the farmer and the mill.

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 institution's enforcement and information technology—the focus of section (2.3)).

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 farm-

[6]The 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).

ers. 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.[7]

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

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

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 institution (either the bank or the mill) is denoted by l.

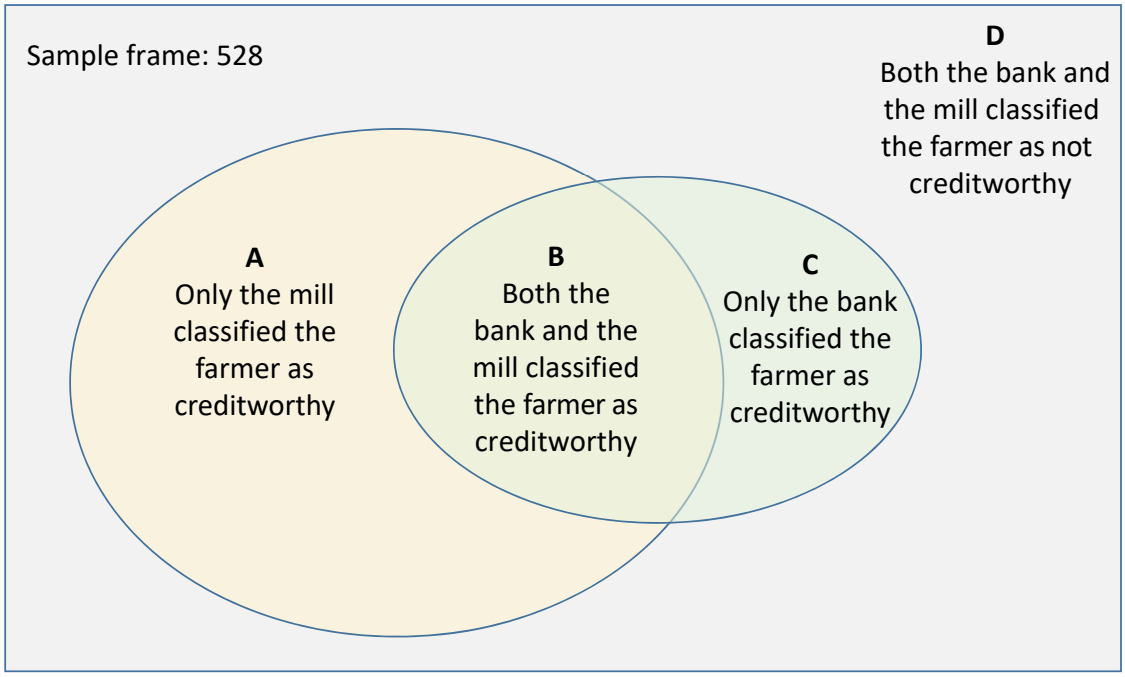
To explore similarities and differences across the institution's choice, we define four mutually exclusive groups: *group (A) "creditworthy only mill$_f$"* (farmers that were solely denoted as creditworthy by the mill), *group (B) "creditworthy both$_f$"* (farmers that were solely denoted as creditworthy by the mill and the bank), *group (C) "creditworthy only bank$_f$"* (farmers that were solely denoted as creditworthy by the bank), and *group (D) "creditworthy neither$_f$"* (farmers that were denoted by both the bank and the mill as not creditworthy). 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), 51 farmers in group (C) and 204 farmers in group (D).

2.3 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 institution screening technology and (ii)

[7] An 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%.

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. Farmers could be allocated to four mutually exclusive groups: (A) only the mill defined the farmer as creditworthy *(185 farmers)*, (B) both institutions defined the farmer as creditworthy *(88 farmers)*, (C) only the bank defined the farmer as creditworthy *(51 farmers)*, or (D) no institution defined the farmer as creditworthy *(204 farmers)*.

institution enforcement technology to determine loan repayment outcomes.

To explore the effect on repayment, using the results from the willingness to lend experiment (stage 1), we randomized loan contract offers to farmers depending 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, we were only able to partially randomize contracts across these groups.

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, in addition to the farmers that the bank deemed creditworthy, the bank was also willing to offer a direct bank loan to farmers that the mill classified as creditworthy. However, the bank was unwilling to lend to farmers that neither the mill or the bank had denoted as creditworthy. Unfortunately, the mill was only willing to guarantee loans that the mill had previously

denoted as creditworthy.

Overall, farmers in groups (A) and (B) received both loan contracts and farmers in group (C) received only direct bank loans. Since both the bank was unwilling to directly lend and the mill unwilling to guarantee loans to farmers in group (D), they were omitted from the rest of the study. For reference, tables (A.11) and (A.12) in Appendix (A.2) report the definition of each group.

Across the various treatment arms, the bank was willing to lend a total of approximately $700,000 in direct and guaranteed loans. Given that the initial interest in receiving loans was larger than the available funds for lending, we randomized on whether farmers were actually 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. In total, as described later in more detail, 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 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 (A.12) in Appendix (A.2) 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 (A.14) and (A.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).

An additional 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. We find no evidence for selective attrition. First, table (A.16) in the Appendix finds that farmers that took the loan were mostly similar on observables (on measures such as assets, farm size) to the farmers that did not take the loan but were offered a loan. Second, our experiment had high take-up rates with close to 75 percent of farmers taking the offered loan. Finally, table (A.17) in the Appendix shows that the loan take-up rate across the various treatment groups is similar. Moreover, the Chi-squared test at the 10 percent significance level is unable to reject the null hypothesis that the take-up rates are the same across treatment groups.

Overall, a total of 204 loans were disbursed 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]).

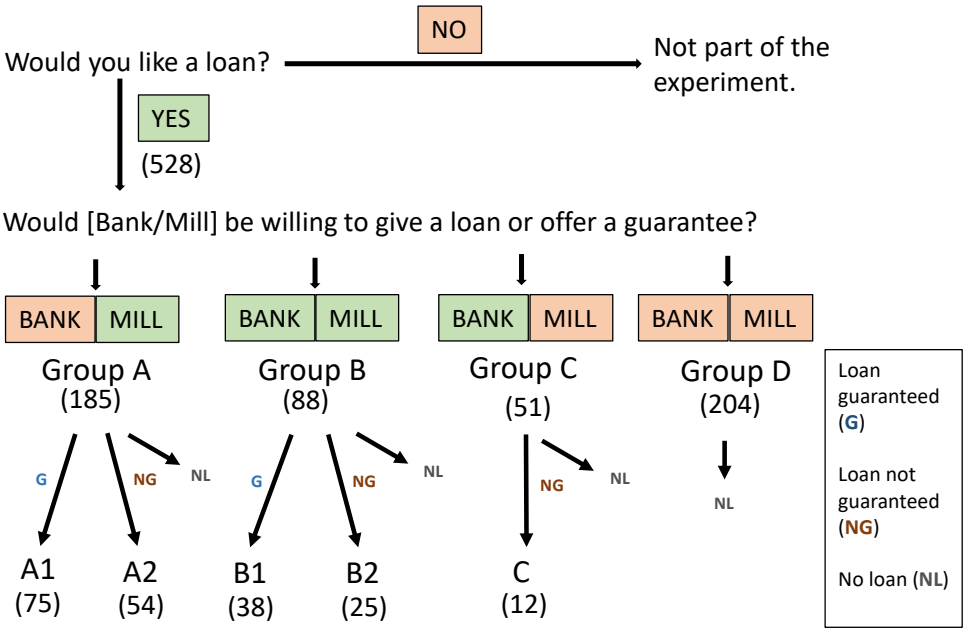
2.3.1 Empirical specifications

Enforcement technology

To test the importance of each institution's enforcement technology, we compare the repayment rates for those farmers that were defined as creditworthy by the *same* institution 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, we analyse differences in repayment rates for the farmers defined as creditworthy by only the mill, therefore holding the credit screening technology constant, 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, we analyse differences 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*). Specifically, in table (4), we run

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 2.3.

regressions similar to the following form on the selected samples:

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

where "Overdue$_f$" is a binary variable equal to zero 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.

Screening technology

To test the importance of each institution'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 institutions. Importantly, even though both institutions are asked to assess the creditworthiness of the farmers, this classification will vary across institution. For the bank, creditworthiness will mean will the farmer repay a direct loan. For the mill, creditworthiness will be whether the farmer will repay the bank conditional on the mill guaranteeing the loan; that is, would the farmer be willing to potentially renege on the relational contract between the farmer and the mill. By comparing borrowers that were selected by different institutions but were offered the same loan terms, we can identify the relative effectiveness of each institution's screening technology.

First, for farmers that received a 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, 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 institutions.

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

$$\text{Overdue}_f = \beta_S \text{Creditworthy only Mill}_f + \beta_1 \text{Farmer characteristics}_f + \epsilon_f \qquad (3)$$

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 screen-

ing 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.

2.4 Experiment design: Stage 3–Endline 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.[8]

2.5 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.

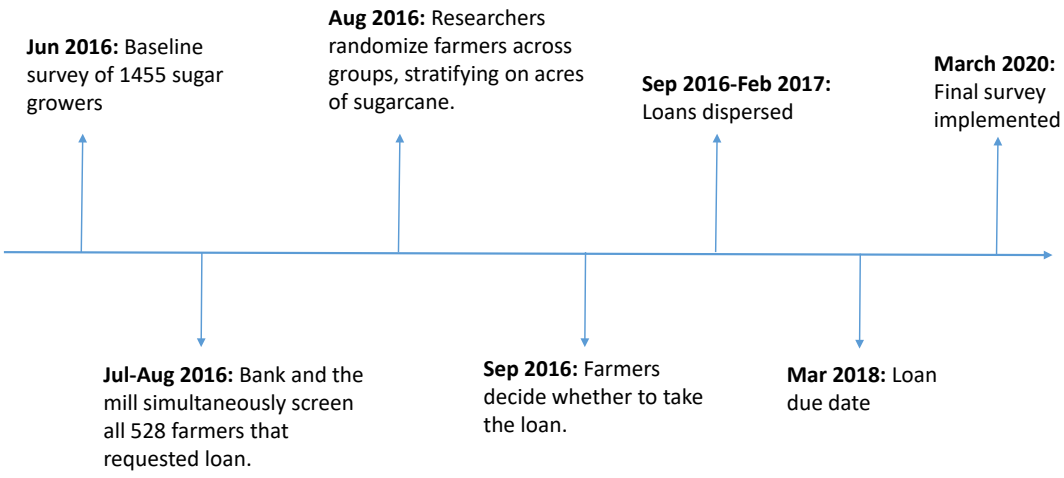
In July and August 2016, the bank and the mill simultaneously (and independently) screened the set of farmers that were interested in taking a loan.

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 disbursed 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

[8] A successful callback rate of over 60 percent.

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.

Figure 3: Timeline of the experiment



2.6 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 conducted an endline survey to learn more about why farmers did or did not repay.[9]

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

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

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.

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.

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: Screening technology

To examine how the set of farmers who are judged to be creditworthy across different institutions with different information and enforcement technology, we asked each institution 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 institution 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 institution defined the farmer as creditworthy.

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

The mill was significantly more willing to be the residual claimant on loans to farmers than the bank. Out of the set 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).[10]

Additionally, the bank and the mill evaluate farmers on significantly different characteristics. There was a large number of farmers (235 farmers) that only one institution 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).

Table 1: Size of each group

Group	A	B	C	D
Defined as creditworthy by:	*Only the mill*	*Both*	*Only the bank*	*Neither*
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 neither insitution 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 institutions 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 institution defined as creditworthy are apparent when comparing groups A and C—the farmers that were only chosen by one institution. The bank was significantly more 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 guarantee loans to farmers that sold relatively and absolutely more of their output to the mill (so more likely to have a deeper economic relationship and more likely to be able to sustain a relational contract). 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.

[10]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 institution defined as creditworthy.

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	0.0000	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 (A.13) in Appendix (A.2) 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 institution 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.[11]

The results in table (3) are broadly similar with the results 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 positively screened farmer's with high crop sales, large farm sizes, and owned expensive farm machinery. In contrast, consistent with theory of relational contracts, 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.

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

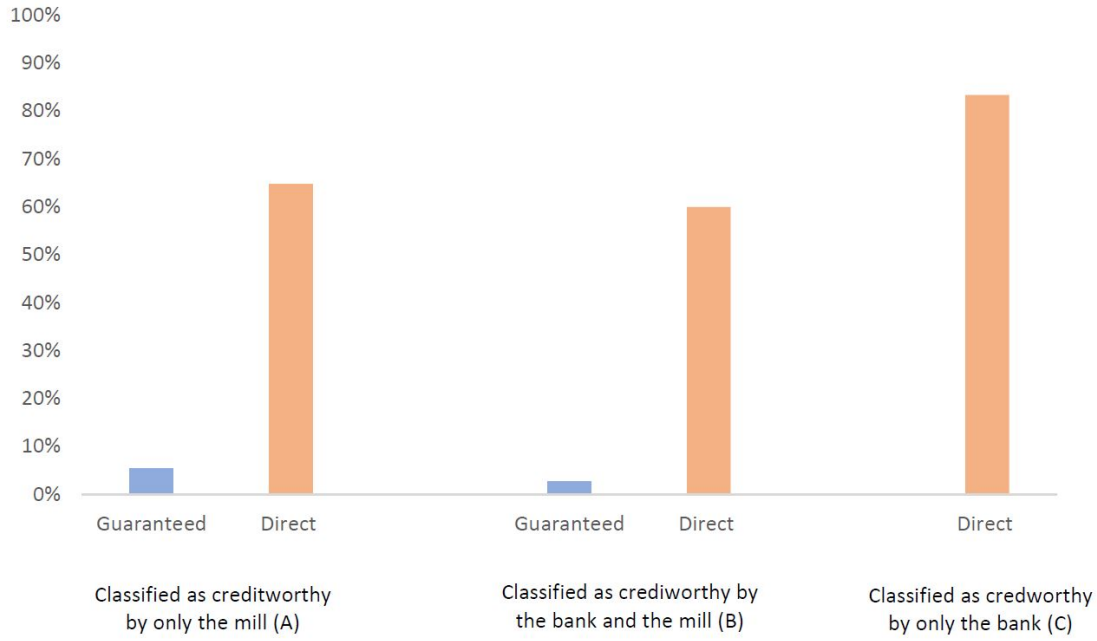
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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 $\text{Creditworthy}_{lf} = \beta_1 \text{Farmer characteristics}_f + \beta_2 \text{Farmer characteristics}_f * \text{mill}_l + \epsilon_{lf}$ and are clustered at the farmer level. Table (A.13) in Appendix (A.2) 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 institution* 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 institutions* affected farmer repayment.

Figure 4: Overdue rates by contract and screened group



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 institutions' 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 institution.

In table (4), we report the results from regressing repayment on whether the farmer's loan was guaranteed by the mill, and additional controls.[12] 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). Columns 2 and 4 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 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

[12] More detail on the empirical strategy is provided earlier in section (2.3.1).

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) Overdue	(2) Overdue	(3) Overdue	(4) Overdue
Loan Guarantee	-0.59***	-0.62***	-0.57***	-0.55***
	(0.071)	(0.070)	(0.10)	(0.10)
Sugarcane planted (acres)		-0.0034		-0.00091
		(0.0024)		(0.0014)
Relative value of crop sales (decile)		0.018		-0.027
		(0.017)		(0.021)
Educ. below high school (Y=1;N=0)		-0.0075		-0.015
		(0.071)		(0.084)
Farm size (acres)		0.00071		0.0012
		(0.0011)		(0.0011)
Years selling to the mill		-0.010**		-0.020*
		(0.0047)		(0.011)
Sales to the mill (pct of total sales)		0.00032		-0.0042**
		(0.0011)		(0.0019)
Within 5km of the mill (Y=1;N=0)		0.018		0.19
		(0.068)		(0.12)
Relative sales to the mill (decile)		0.0032		0.0041
		(0.020)		(0.027)
Formal credit history (Y=1;N=0)		0.016		0.071
		(0.075)		(0.080)
Prev. bank loan overdue (Y=1;N=0)		0.019		-0.063
		(0.058)		(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 (A.13) in Appendix (A.2) 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 surpris-

ing 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) Overdue	(2) Overdue	(3) Overdue
Guarantee x Sugarcane planted	0.0057	0.0054	0.0087**
	(0.0071)	(0.0057)	(0.0036)
Guarantee x Relative value of crop sales (decile)	-0.023	0.16***	0.013
	(0.038)	(0.047)	(0.030)
Guarantee x Educ. below high school	0.075	-0.00045	-0.057
	(0.16)	(0.21)	(0.13)
Guarantee x Farm size (acres)	-0.0019	-0.0074***	-0.0042***
	(0.0027)	(0.0016)	(0.0015)
Guarantee x Years selling to the mill	0.0074	0.036	0.0068
	(0.011)	(0.030)	(0.0086)
Guarantee x Sales to the mill (pct of total sales)	-0.0013	0.0095***	0.0016
	(0.0026)	(0.0033)	(0.0020)
Guarantee x Within 5km of the mill	-0.13	0.060	-0.12
	(0.16)	(0.30)	(0.13)
Guarantee x Relative sales to the mill (decile)	0.014	-0.030	0.0041
	(0.047)	(0.052)	(0.036)
Guarantee x Formal credit history	-0.048	-0.014	-0.0047
	(0.17)	(0.16)	(0.13)
Creditworthy only Mill			-0.042
			(0.066)
Constant	0.66**	2.91***	0.96***
	(0.26)	(0.60)	(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 additional control variables, specifically, we include the non-interacted variables (that is, Loan guarantee, acres sugarcane, value of crop sales, etc.) but in the interest of space and simplicity we do not report the coefficient results. Table (A.13) in Appendix (A.2) 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 institutions' screening technology, we first compare repayment rates between farmers that received the same contract but were selected as creditworthy by different institutions.

In table (6), we report the results from regressing repayment on whether the farmer was chosen by the mill, and additional controls.[13] 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.

[13] More detail on the empirical strategy is provided earlier in section (2.3.1).

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

	(1) Overdue	(2) Overdue	(3) Overdue	(4) Overdue	(5) Overdue	(6) Overdue
Creditworthy only Mill	0.027	-0.044	0.048	-0.038	-0.028	-0.078
	(0.037)	(0.030)	(0.12)	(0.16)	(0.10)	(0.15)
Sugarcane planted (acres)		-0.00030		-0.0090**		-0.0042
		(0.00028)		(0.0037)		(0.0041)
Relative value of crop sales (decile)		0.0049		-0.0076		0.0067
		(0.012)		(0.028)		(0.023)
Educ. below high school (Y=1;N=0)		-0.0058		0.051		-0.012
		(0.041)		(0.13)		(0.12)
Farm size (acres)		-0.00020		0.0040**		0.0023**
		(0.00017)		(0.0015)		(0.0011)
Years selling to the mill		-0.0067*		-0.013		-0.014
		(0.0037)		(0.0090)		(0.0086)
Sales to the mill (pct of total sales)		0.00026		-0.0014		-0.0011
		(0.00074)		(0.0019)		(0.0018)
Within 5km of the mill (Y=1;N=0)		-0.031		0.084		0.12
		(0.049)		(0.13)		(0.11)
Relative sales to the mill (decile)		0.0063		0.0021		-0.014
		(0.015)		(0.033)		(0.026)
Formal credit history (Y=1;N=0)		0.0065		0.012		0.036
		(0.049)		(0.12)		(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
Additional controls	No	Yes	No	Yes	No	Yes
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 (A.13) in Appendix (A.2) 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) Overdue	(2) Overdue	(3) Overdue	(4) Overdue
Creditworthy only Mill	0.036	-0.038	0.0016	-0.058
	(0.053)	(0.064)	(0.051)	(0.063)
Loan Guarantee	-0.59***	-0.59***	-0.62***	-0.61***
	(0.058)	(0.057)	(0.054)	(0.054)
Sugarcane planted (acres)		-0.0023		-0.0018
		(0.0016)		(0.0013)
Relative value of crop sales (decile)		0.0042		0.0089
		(0.013)		(0.012)
Educ. below high school (Y=1;N=0)		0.012		-0.0047
		(0.053)		(0.052)
Farm size (acres)		0.00097		0.00083
		(0.00065)		(0.00053)
Years selling to the mill		-0.0090**		-0.0092**
		(0.0041)		(0.0041)
Sales to the mill (pct of total sales)		-0.00061		-0.00069
		(0.00089)		(0.00087)
Within 5km of the mill (Y=1;N=0)		0.053		0.059
		(0.058)		(0.058)
Relative sales to the mill (decile)		-0.0052		-0.012
		(0.016)		(0.014)
Formal credit history (Y=1;N=0)		0.026		0.027
		(0.056)		(0.054)
Prev. bank loan overdue (Y=1;N=0)		0.011		0.0093
		(0.040)		(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 (A.13) in Appendix (A.2) describes the variable definitions.

3.3 Endline survey: Survey exploring the large differences in repayment rates

In this section, we present the findings from an endline 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. [14]

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 disbursed 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 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

[14] As 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.

relatively inexpensive as the mill collected repayments directly from the farmer at the time of milling the sugarcane.

Table 8: Endline survey question: How was payment collected?

	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 an endline 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: Endline 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 an endline 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.[15]

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 bundling and the potential adverse impact on their relationship with the mill. In contrast,

[15] Djankov 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.

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: Endline 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/ penalities	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 an endline 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, specifically for large productive investments. 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.[16] 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 institution. 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 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. [2020], Bharadwaj et al. [2019], Frost et al. [2019]).

We should also touch upon some of the drawbacks of our study. Our experiment mostly

[16]The 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 indirectly boosts the mill's profit because it likely 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

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. Moreover, the results of the experiment's stage 3 demonstrated that borrowers are 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

D. Acemoglu, S. Johnson, and J. A. Robinson. Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 1:385–472, 2005.

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.

K.-H. Bae and V. K. Goyal. Creditor rights, enforcement, and bank loans. *The Journal of Finance*, 64(2):823–860, 2009.

G. Baker, R. Gibbons, and K. J. Murphy. Relational contracts and the theory of the firm. *The Quarterly Journal of Economics*, 117(1):39–84, 2002.

A. V. Banerjee and E. Duflo. Reputation effects and the limits of contracting: A study of the indian software industry. *The Quarterly Journal of Economics*, 115(3):989–1017, 2000.

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. Berg, V. Burg, A. Gombović, and M. Puri. On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies*, 33(7):2845–2897, 2020.

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.

A. Blouin and R. Macchiavello. Strategic default in the international coffee market. *The Quarterly Journal of Economics*, 134(2):895–951, 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. Guirkinger. 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.

G. Bryan, D. Karlan, and J. Zinman. Referrals: Peer screening and enforcement in a consumer credit field experiment. *American Economic Journal: Microeconomics*, 7(3):174–204, 2015.

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.

B. Feigenberg, E. Field, and R. Pande. The economic returns to social interaction: Experimental evidence from microfinance. *Review of Economic Studies*, 80(4):1459–1483, 2013.

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.

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. 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.

R. Macchiavello and A. Morjaria. Competition and relational contracts in the rwanda coffee chain. *The Quarterly Journal of Economics*, 136(2):1089–1143, 2021.

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.

J. McMillan and C. Woodruff. Interfirm relationships and informal credit in vietnam. *The Quarterly Journal of Economics*, 114(4):1285–1320, 1999.

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.

A Appendices

A.1 Background on the sugar industry

Sugar in Pakistan

Sugar is a major global commodity that is produced in over 100 countries and is a material component of Pakistan's economy. The sugar industry in Pakistan consists largely of growing sugarcane and the production, manufacture, and marketing of white sugar. In Pakistan, the sugar manufacturing industry has expanded substantially. At the time of independence in 1947, there were only *two* sugar mills in Pakistan. Following the setup of a government commission in 1957, the industry started to expand substantially, reaching around 50 mills by 1990 and then further expanding to 89 mills. All told, the daily crushing capacity has increased in 1990 from 90,000 metric tons to its current capacity of 505,000 tons. Sugarcane is the second most important cash crop in Pakistan. Its production contributes just over 3 percent in agriculture's value addition and just under 1 percent of Pakistani gross domestic product.

Sugarcane farming is protected by the government with a minimum price for the farmers. This policy plays an important role in persuading growers to cultivate sugarcane over other crops.

Province level data shows that Punjab is the largest sugarcane producing province in Pakistan but production in Sindh is, on average, more efficient. Sugarcane production in Sindh province—where the experiment takes place—has a cultivation area less than half of Punjab. However, production in Sindh is generally more efficient with sugarcane yield per hectare and the recovery ratio (the ratio of the amount of nutrient in the harvested crop to the amount of nutrient applied) higher than Punjab.

Cultivation, harvesting and processing of sugarcane

There are five distinct steps in the production of sugar from sugarcane.

The first stage of the cultivation is tilling or weeding the land. If sugarcane was planted in the previous year, the land solely needs to be weeded. Otherwise, the land must be tilled, which is usually done with an owned or rented tractor.

The second stage is planting and caring for the plants. During the growth process the

sugarcane plant receives care in the form of pesticide protection, fertilizer, and water. Sugarcane straws can grow up to two metres in about twelve to eighteen months

The third stage is harvesting. The harvesting of sugarcane in Pakistan is generally done by hiring local manual labor. Sugarcane is harvested when the plant is around two meters high, at which point the plant is cut near the root allowing the stem to regrow for a future crop.

The fourth stage is the transportation and sale of the sugarcane to the sugar mill. In Pakistan, the sugarcane is manually loaded onto trailers or trucks. Once at the mill, the sugar cane is weighed and sold.

The fifth stage is the milling of the sugarcane. To begin, the sugarcane is crushed into a liquid. Subsequently, this liquid is heated and filtered to form sugar crystals, which is then followed by centrifugation to remove excess water. Finally, the refined sugar is sorted, packaged, and distributed.

Sugar farming in Matiari district, Sindh Province

In the district of Matiari—the district where the experiment is conducted—sugarcane is grown under well-irrigated conditions where water is available either through surface canals or underground pumping facility using tube wells. Moreover, the district is well situated in the favorable subtropical weather conditions and good annual rainfall.

At the time of our survey, there are two sugar mills, Matiari Sugar Mill—the intermediary in the experiment—and Mehran Sugar Mill, that are located close to the farmers in our experiment. The two mills were only 30 kilometers apart. In our survey area, typically, a mill would buy sugarcane from farmers in the radius of 25-30 kilometers. Similar to other areas in Pakistan, in Matiari, sugarcane farming competes with other cash crops such as, wheat, cotton, onion and chilies. In discussions with farmers, farmers report they prefer growing sugarcane for several reasons. First, the guaranteed government minimum support price. Second, easy access to multiple buyers (two sugar mills in our case. Third, lower manual effort to grow sugarcane relative to other cash crops, such as cotton. Fourth, sugarcane plant is relatively resilient to weather shocks. Finally, there is a well-developed local market for required inputs, such as a manual labor market.

A.2 Variable and group definitions

Table A.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 institution (Not creditworthy)

Table A.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 A.13: Variable definitions

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

A.3 Baseline characteristics and balance between treatment arms

Table A.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 A.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 A.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 A.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). Overall, we do not find the take-up rates are statistically different across treatments (the Chi-squared test is unable to reject the null hypothesis that take-up results are the same across treatments at the 10 % significance level).