Financing Smallholder Agriculture: An Experiment with Agent-Intermediated Microloans in India *

Pushkar Maitra[†], Sandip Mitra[‡], Dilip Mookherjee[§], Alberto Motta[¶] and Sujata Visaria[‖]

September 28, 2016

Abstract

Recent evaluations have found that traditional microloans have insignificant impacts on incomes and output. Randomly selected villages in West Bengal, India participated in a field experiment with a novel variant of microcredit called TRAIL, where the selection of borrowers of individual liability loans was delegated to local trader-lender agents incentivized by repayment-based commissions. Other randomly selected villages participated in a group-based microcredit program called GBL. TRAIL loans increased the production of the leading cash crop and farm incomes by 27–37%, but GBL loans had insignificant effects. To understand underlying mechanisms, we develop and test a theoretical model that explains borrower selection into the two schemes as well as borrower incentives to invest the loans for productive purposes. We find that borrowers selected by the TRAIL agents were more able farmers than those who self-selected into the GBL scheme; this pattern of selection explains about a third of the observed difference in income impacts.

Key words: Agricultural Finance, Agent-based Lending, Group Lending, Selection, Repayment

JEL Codes: D82, O16

*Funding was provided by the Australian Agency for International Development, the International Growth Centre, United States Agency for International Development and the Hong Kong Research Grants Council. We are grateful to Shree Sanchari for collaborating on the project. Jingyan Gao, Clarence Lee, Daijing Lv, Foez Mojumder, Moumita Poddar and Nina Yeung provided exceptional research assistance and Elizabeth Kwok provided excellent administrative support. Boston University Masters students Torry Ah-Tye, Ou Bai, Juan Blanco, Chantel Pfeiffer and Stefan Winata conducted useful analysis and provided insights from a field study of relations between agents and borrowers in the study. We thank two anonymous referees, the co-editor of this journal, Xavier Gine, Albert Park, Russell Toth, Farshid Vahid, Bruce Wydick and a large number of seminar and conference participants for helpful comments on previous and related versions. Internal review board clearance was received from Monash University, Boston University and the Hong Kong University of Science and Technology. The authors are responsible for all errors.

[†]Pushkar Maitra, Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia. Pushkar.Maitra@monash.edu.

[‡]Sandip Mitra, Sampling and Official Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India. Sandip@isical.ac.in.

[§]Dilip Mookherjee, Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA. dilipm@bu.edu.

[¶]Alberto Motta, School of Economics, UNSW Australia, NSW 2052, Australia. motta@unsw.edu.au.

[‖]Sujata Visaria, Department of Economics, Lee Shau Kee Business Building, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. svisaria@ust.hk.

1 Introduction

Microcredit was famously heralded as a solution to global poverty; yet a large number of experimental evaluations have found no evidence of significant impacts on borrower incomes or production (Kaboski and Townsend, 2011, Banerjee, Karlan, and Zinman, 2015). This is true not only for the traditional group-based lending schemes that have been evaluated most commonly, but also for variants with individual liability loans (Giné and Karlan, 2014, Attanasio, Augsburg, De Haas, Fitzsimons, and Harmgart, 2015). Other experiments have found that when the rigid repayment schedules that restrict borrowers' project choice are relaxed, microloans increase farm activity and business incomes, but at the cost of increased default rates (Field, Pande, Papp, and Rigol, 2013, Feigenberg, Field, and Pande, 2013).[1] Thus far no microcredit variant studied in the literature has generated significant average treatment effects on borrower output or incomes and maintained high repayment rates at the same time. The reasons for this are not well understood. It is also not known whether alternative variants would be more likely to succeed.

In this paper we explore the hypothesis that part of the reason traditional group-based microfinance schemes fail to increase borrower incomes is that they do not screen out unproductive borrowers. Given their greater likelihood of default, unproductive borrowers pay higher interest rates in the informal credit market, and so have a strong incentive to apply for MFI loans if they are available. Since MFI loan officers lack fine-grained information about the risk and productivity characteristics of poor borrowers, they cannot screen them with sufficient precision. Group lending contracts also do not help to screen out unproductive borrowers in practice, because they are rarely as sophisticated as needed to generate the selection patterns that have been highlighted by theoretical formulations (Ghatak, 1999, 2000).[2]

These considerations motivated us to experiment with an alternative method of selecting borrowers called agent-intermediated lending: formal financial institutions or MFIs could collaborate with local informal lenders and tap into their knowledge of the characteristics of different borrowers within the community. This paper considers a version (called trader-agent intermediated lending or TRAIL) in which the lender delegates borrower selection to an agent randomly chosen from among the informal trader/lenders in the community. The agent is incentivized by commissions that depend on interest paid by the clients he recommended. This could motivate the agent to select borrowers who are less likely to default. Under the plausible hypothesis that default risk and productivity are negatively correlated, this would result in a borrower pool with high average productivity.

To test this idea, we devised and conducted a field experiment in two districts of West Bengal in India. The experiment implements TRAIL in randomly selected villages and compares its effects with group-based lending (GBL), implemented in another set of randomly selected villages. In each TRAIL village one agent was randomly selected from a list of established trader-lenders within the village, and asked to recommend, as potential borrowers, 30 village residents who owned no more than 1.5 acres of cultivable land. A random subset of these recommended households were

[1] Field, Pande, Papp, and Rigol (2013) find that a longer grace period for repaying individual liability microloans increased weekly business profits by 41% and household incomes by almost 20%, but also tripled the default rate.

[2] MFIs do not typically offer menus of credit contracts within the same village that would induce different borrower groups to sort according to their type. They do tend to include savings requirements and group meetings, but there is no clear evidence on how well these succeed at screening out less productive borrowers.

offered individual liability loans at below-market interest rates, repayable in a singe lumpsum at the end of four months. The agent was promised a commission equal to 75 percent of the interest payments received from borrowers he had recommended. He also incurred penalties for borrower defaults. Borrowers were incentivized to repay because repayment was tied to future growth in credit access: in any subsequent four month cycle borrowers could take a new loan worth 133 percent of the principal they had repaid in the previous cycle. Besides charging an interest rate substantially below average rates in the informal credit market, the loan contracts also provided insurance against covariate risks.

A Kolkata-based MFI called Shree Sanchari implemented the GBL scheme. In each GBL village, residents owning less than 1.5 acres of cultivable land could form 5-member groups. Consistent with standard MFI practices in rural West Bengal, groups were required to meet with loan officers each month and make regular savings deposits for the six months before the loan scheme began.[3] A random subset of the groups that completed this initiation process were offered joint liability loans. The monthly group meetings and savings continued throughout the loan cycle. As described further below, GBL loans featured the same interest rate, loan duration, growth in credit access and covariate risk insurance as the TRAIL loans. The MFI received a commission equal to 75 percent of the interest payments received from GBL borrowers. Neither the TRAIL agent nor the MFI were responsible for providing loan capital: this was funded through a research grant.

As is clear from the discussion above, the two schemes had different selection procedures as well as different liability rules. In the GBL program groups were jointly liable for repayment, but TRAIL borrowers were individually liable for their own loans.[4] It follows that the two schemes provided borrowers with different incentives. In a joint liability contract the borrower may be called upon to pay up on behalf of a defaulting group member, and this raises the effective interest rate she faces. This can limit borrower incentives to expand the scale of borrowing. Equally, to avoid incurring this "joint liability tax", group members might monitor each other and discourage risky projects, such as the adoption of high-value high-risk cash crops (Fischer, 2013). The agent might also provide help to the farmer and/or monitor the borrower differently from how GBL group members monitor and help each other. These different incentive effects could generate significant differences in the impacts of the two schemes on borrowers' crop output and value-added, even in the absence of any selection differences.

To better understand and distinguish between these underlying mechanisms, we develop a theoretical model of borrower heterogeneity and incentives. The model assumes borrowers have heterogenous ability, where ability is negatively correlated with default risk and positively correlated with productivity. It builds on the Ghatak (2000) model of selection in group borrowing, by adding a productivity dimension to the types of borrowers and by allowing for an informal credit market.[5]

[3]Many group-lending schemes in different parts of the world require that members save regularly for a pre-assigned duration or meet a savings target before they can begin to borrow. It is often argued that this builds the financial discipline required to repay regularly.

[4]Ideally we would have separately varied selection methods and loan liability rules to study the respective effects of each. We were unable to do this because TRAIL agents had no experience with and were unwilling to be involved with group loans, whereas Shree Sanchari had no experience with individual liability loans.

[5]The productivity dimension is needed to explain the positive effects of the loans on borrower's output and incomes, which is our key experimental finding. The informal credit market is needed to model the borrower selection choices of the TRAIL agent, who is a local informal lender and observes borrower types within his own clientele. This also helps to explain how the selection pattern in the TRAIL scheme differs from self-selection

The model explains why TRAIL agents are motivated to select high ability borrowers, irrespective of whether or not there is corruption in the selection process. In contrast, the GBL scheme attracts both low and high ability borrowers, because there is no alternative loan contract offer that helps to separate the high from the low ability borrowers. The effect of this difference in selection (the *selection effect*) is compounded by differential *incentive effects*: which imply that for a borrower of a given ability, the GBL loans increase incomes by less than TRAIL loans do, because the joint liability tax raises the effective interest rate. Both selection and incentive effects work in the same direction, implying that the TRAIL scheme creates larger average treatment effects on production and farm incomes than the GBL scheme does.

The model generates a number of other testable predictions. These include the following: more able borrowers devote more land to cultivation and produce higher output; more able borrowers pay lower interest rates on the informal market, and the TRAIL agent is more likely to recommend the more able borrowers from his segment. Under additional assumptions, more able borrowers experience larger loan treatment effects on borrowing, cultivation, output and farm incomes. To test these predictions, we impose a Cobb-Douglas functional form on the production function, and postulate that farmer ability is a composite effect of fixed factors owned and other household attributes. We also impose a constant elasticity relationship between ability and crop failure risk. The model then provides a method of estimating each farmer's ability as a farmer fixed effect from a regression of the logarithm of cultivation scale or of output on farmer and year dummies, in the spirit of Olley and Pakes (1996) and Levinsohn and Petrin (2003).[6] The generated ability estimates allow us to test the detailed predictions mentioned above.[7] They also allow us to decompose the difference in the average treatment effects in the TRAIL and the GBL schemes into the respective contributions of selection and incentive differences.[8]

Our first main experimental finding is that the TRAIL loans generated significant average treatment effects (ATE) on production and incomes, but GBL loans did not. The ATE differences between the two schemes are large in magnitude: for example, TRAIL loans increased average farm value-added by 25 percent over the mean, whereas GBL loans had a statistically non-significant effect of 4 percent. This difference is also statistically significant (p-value=0.077). The large treatment effects of the TRAIL loans are driven by increased cultivation of potatoes, the leading cash crop in this region, whose cultivation the loans were designed to facilitate.

The model makes no definite predictions about how the repayment rates of the TRAIL and GBL schemes should compare. On the one hand, GBL borrowers have lower ability on average and therefore a higher risk of crop failure. On the other hand, conditional on ability a GBL borrower benefits from the joint liability feature of his contract, because group members might repay on his behalf if his crop fails. We find that repayment rates were an equally high 95% over the 3 years in both the TRAIL and GBL schemes. However loan take-up rates were significantly higher in

patterns in the GBL scheme, where a borrower's propensity to form a group increases with the gap between the interest rate she pays to the informal lender and the interest rate on the group loan.

[6] Our model can be viewed as a special case of theirs, where farmer ability is fixed over time rather than following a first order Markov process.

[7] For each prediction that uses the generated ability measure, we test statistical significance on the basis of a distribution of 2000 cluster-bootstrapped estimates.

[8] The selection effect is a weighted average of difference in selected proportions of different ability types, with TRAIL treatment effects for given types serving as weights. The incentive effect is a weighted average of difference between TRAIL and GBL treatment effects for given types, with GBL selection proportions for different types serving as weights.

the TRAIL scheme.[9]

The experimental evidence is also consistent with the more detailed predictions of the model. The distribution of estimated ability among households recommended by TRAIL agents first order stochastically dominates the distribution of households who self-selected into GBL groups.[10] Households with higher estimated ability paid lower (annualized) interest rates on informal credit taken before the study began. TRAIL treatment effects on borrowing, output and incomes were larger for more able borrowers. Hence selection differences contributed positively to the observed difference in average treatment effects between the TRAIL and GBL schemes. Our decomposition procedure suggests that the selection effect contributes 27–36 percent of this difference.

We also address the concern that TRAIL agents may have abused their power to extract benefits from the borrowers they recommended. Our data show no evidence that they manipulated the terms of other trading relationships with treated borrowers: either to siphon off their benefits, or to create large positive effects by subsidizing inputs or enabling them to realize higher prices for output sales. Finally, the administrative costs of the TRAIL scheme were lower than those of the GBL scheme. This is because commission rates for both the TRAIL agents and the MFI that implemented the GBL scheme were the same, but the MFI's loan officers incurred substantial costs on high-frequency meetings with borrowers in the GBL scheme, which were not part of the TRAIL design. Combined with its higher take-up rates and identical repayment rates, this indicates that the the TRAIL scheme outperformed the GBL scheme on financial sustainability.

Our paper contributes to the literature by exploring whether selection problems can provide an explanation for the disappointingly low effects of traditional microcredit on output and borrower incomes. Our focus on heterogeneity and endogenous selection patterns is similar to that of Beaman, Karlan, Thuysbaert, and Udry (2015), who conduct a field experiment with group loans in Mali. However, their study compares group loans with grants, and focuses on differences between borrowers and non-borrowers in the group loan scheme. We focus instead on differences between the two alternative methods of selection in the TRAIL and GBL schemes. An additional contribution of our paper is therefore the design and implementation of a new approach to microlending.

A few qualifications are in order. The scale of our intervention was smaller than most other microcredit experiments, since only ten loans were offered in each village. The results of this experiment cannot be used to predict the consequences of a larger scale intervention.[11] Also, our analysis was restricted to impacts on production and incomes; we have not examined impacts on consumption smoothing, liquidity management, investment or social empowerment. We defer an examination of the distributive impacts of the TRAIL relative to the GBL scheme to a subsequent paper. This paper does not claim that the TRAIL scheme generated welfare-superior outcomes. Instead, our objective was to understand better why group-based lending has failed to generate large growth impacts, and to initiate the exploration of a promising alternative.

[9] Loan records show that 92% of households that were offered TRAIL loans took the loan in the first four-month cycle of the scheme. At the end of three years, the take-up rate was 62%. In the GBL scheme the take-up rate was 88% to begin with, and fell to 49% by the end of the third year.

[10] The estimated ability distributions are significantly different according to the Kolmogorov-Smirnov test (p-value = 0.00).

[11] The small scale of our interventions also imply that spill-overs on non-beneficiaries in the experimental villages were unlikely.

2 Experimental Design and Data

Our experiment was conducted in the Hugli and West Medinipur districts of the state of West Bengal, India. These are among the largest producers of potatoes in West Bengal, which itself produces about a third of all the potato output in India. Potatoes are a particularly high-value crop: as we shall show below, they generate the highest value-added per acre of all crops grown in the area. For this reason, the loan products were designed so that they could be used for potato cultivation.

In both the TRAIL and the GBL scheme, borrowers were offered repeated loans of 4-month durations at an annual interest rate of 18%, which was below the prevailing market rate of 25 percent. The first loans were capped at ₹2000 (equivalent to approximately $US40 at the prevailing exchange rate), and were disbursed in October-November 2010, to coincide with the potato-planting season. Repayment was due in a single lumpsum after 4 months. In each subsequent cycle, borrowers who repaid the entire amount that was due became eligible for a 33 percent larger loan, on the same terms as before. Those who repaid less than 50 percent of the repayment due were not allowed to borrow again. Others were eligible to borrow 133 percent of the principal repaid.[12] Both schemes had an in-built index insurance scheme, according to which the required repayment would be revised downwards if the revenue per acre for potatoes fell 25 percent below a three year average in the village, as assessed through a separate village survey.[13]

Each sample village was at least 10 kilometers away from all other sample villages, to minimize contamination of the experimental interventions through the spread of information. The MFI had not operated in any of the sample villages before our project started, and in general MFI penetration was low in these regions. A research grant held by the project team provided the funds for all loans in the two schemes.

Table 1 summarizes the differences between the TRAIL intervention and other related microcredit interventions recently studied in the literature (see the summary presented in Banerjee, Karlan, and Zinman, 2015, Table 1). Apart from the method of borrower selection in the TRAIL scheme, an important difference is in repayment frequency: loans were due in a single lumpsum at the end of 4 months in both the TRAIL and GBL schemes, whereas repayment was due on weekly, bi-monthly or monthly schedules in the other studies. Many of the other loan features are similar across TRAIL, GBL and other microcredit programs.

As stated above, we rationed loan offers to 10 borrowers in each village. In contrast, the scale of most other interventions was determined by the demand for the loan product: any eligible individual in the treatment slum or village could participate in the loan scheme. The impacts estimated in those studies combine selection and loan treatment effects. They can be interpreted as the effects of MFI entry on a representative member of the eligible sub-population within that sampling unit, where loan take-up within the sub-population is entirely demand-determined.

[12] To facilitate credit access for post-harvest storage, borrowers were allowed to repay the loan in the form of cold storage receipts (or “bonds”) instead of cash. In that case the repayment was calculated at the prevailing price of the bonds.

[13] In yet another 24 villages, an alternative version of the agent intermediated lending scheme (called GRAIL) was implemented, where a member of the village council (*Gram Panchayat*) was appointed as the agent. The GRAIL agent is likely to have been motivated by the political benefits of participating in the scheme. The treatment effects of the GRAIL program will be analysed in a separate paper.

In contrast, in our study, estimates of loan treatment effects control for "selection" into the scheme (either through recommendation by a TRAIL agent or through participation in a GBL group). This is possible because only a subset of households who were recommended (in the TRAIL villages) or joined groups (in the GBL villages) were offered the program loans. In TRAIL villages, the agent recommended 30 individuals for loans, and 10 of these were randomly chosen through a public lottery to receive them. In GBL villages, two of the groups that had survived a 6-month initiation period were randomly chosen to receive loan offers. The loan treatment effects are then estimated as differences in outcomes between those randomly chosen to receive a loan offer (we call these Treatment households in what follows), and those who were recommended or formed a group but were unlucky in the lottery and did not receive the loan offer (we call these Control 1 households). This is similar to the analysis of loan treatment effects in Karlan and Zinman (2011), where loan assignment was randomized among borrowers deemed marginally creditworthy by a credit scoring algorithm. Our design therefore allows us to separately identify selection effects (comparing Control 1 households with those not recommended in TRAIL or those not forming groups in GBL) from loan treatment effects conditional on selection (comparing Treatment with Control 1 households).

The villages where the experiment was conducted had an average of 350 households per village. More than three-quarters of villages had a primary school, a quarter had a primary health centre, 14% had a bank branch and 35% of the villages had access to a metalled road. Households had 5 members on average. The majority of the households were Hindu, and among them, there were roughly equal proportions of high and low castes. The average landholding of village households was 0.46 acres. Nearly 95 percent of households had male heads, about 42% of the household heads had completed primary schooling and about half reported that agricultural cultivation was their primary occupation (see Table 2).

2.1 The Trader-Agent-Intermediated Lending (TRAIL) Scheme

Starting in September 2010, we consulted with prominent persons in each TRAIL village to draw up a list of traders and business people who had operated a business in the village for at least three years, and had at least 50 clients. One person from this list was randomly chosen and invited to become an agent.[14] The agent was asked to (confidentially) recommend as potential borrowers 30 village residents who owned no more than 1.5 acres of agricultural land. In October 2010, our project officer selected 10 out of these 30 names in a lottery conducted in the presence of village leaders. Loan officers visited the treated households in their homes to explain the loan terms and later to disburse the loan if it was accepted.

At the beginning of the scheme, the agent was required to put down a deposit of ₹50 per borrower. The deposit was refunded to the agent at the end of two years, in proportion to the loan repayment rates of his recommended borrowers. At the end of each loan cycle he received as commission 75% of the interest received on these loans. The agent's contract was terminated at the end of any cycle in which 50% of borrowers whom he had recommended failed to repay. Agents were also promised an expenses-paid holiday at a local sea-side resort if they survived in the program

[14] The experimental protocol stated that if the person approached rejected the offer, the position would be offered to another randomly chosen person from the list. However the first person offered the position accepted it in every village.

for two years.

Interactions between loan officers and borrowers were limited to single visits to the borrowers' residences at the beginning of each cycle to disburse loans and at the end of each cycle to collect loans. They were not required to engage in any monitoring or collection effort beyond this. Borrowers were not required to report to the loan officers their intended or actual use of the loan.[15]

A potential concern with the TRAIL intervention is that agents might act in ways that undermine the purpose of the scheme. For instance, they might ask for bribes to recommend borrowers, select unsuitable borrowers (with high default risk, less productive individuals, wealthy individuals, or cronies in exchange for bribes or favors), extract borrower benefits by manipulating other transactions with them, collude with borrowers (encourage them to default and divide up the loan funds instead) or coerce them to repay. To help guard against these possibilities, all loan transactions took place directly between project loan officers and the borrower. The research team verified that the agent recommended only landless and marginal landowners (households owning $\leq$ 1.5 acres, as per the protocol). The team also communicated clearly to all borrowers that the interest rate was fixed, there were no other charges for participation, and that all payments were to be made only to the project loan officers. In any case, in what follows we shall examine the evidence on borrower recommendation patterns, and also check if transactions between borrowers and the TRAIL agent changed as a result of the intervention.

2.2 The Group-based Lending (GBL) Scheme

The MFI began operations in the GBL villages in February-March 2010 by inviting residents to form 5-member groups, and then organizing bi-monthly meetings for each group, where each member was expected to deposit ₹50 per month into the group account. Of the groups that survived until October 15, 2010, two were randomly selected into the scheme through a public lottery. Each group member received a loan of ₹2,000 in Cycle 1, repayable in a single lump sum at the end of four months. Thus the entire group received ₹10,000. All group members shared liability for the entire sum: if less than 50% of the due amount was repaid in any cycle, all members were disqualified from future loans; otherwise the group was eligible for a new loan, which was 33% larger than the previous loan. Bi-monthly group meetings continued throughout, in keeping with the MFI's standard protocol for joint liability lending. At the end of each loan cycle the MFI received as commission 75% of the interest received on these loans.[16]

2.3 Data and Descriptive Statistics

From December 2010 to December 2013, we conducted repeated surveys of 50 households in each village. The surveys collected information about household demographics, assets, landholding,

[15] However in our household surveys we did ask respondents to tell us how they used each loan.

[16] Thus the incentives provided to TRAIL agents and to the MFI were identical. Both faced the same formula for commissions. The paid holiday for surviving in the scheme offered to TRAIL agents was akin to the internal bonus that Shree Sanchari loan officers could expect if their job performance was considered satisfactory.

cultivation, land use, agricultural input use, sale and storage of agricultural output, credit received and given, incomes, and economic relationships within the village. In each village, the household sample was composed of three sub-groups. In TRAIL villages, the agent recommended 30 borrowers for loans, 10 of whom were randomly chosen to receive the loan offer. All 10 of these Treatment borrowers were included in the sample. Of the remaining 20 recommended individuals, a random subset of 10 were also included in the sample; these constitute the Control 1 group. Finally, we included 30 households randomly chosen from those that were not recommended (Control 2). In the GBL villages, of all the groups that formed, two groups were randomly selected to receive the loan offer, and all 10 households from these two groups (Treatment households) were included in the sample. Two groups that had formed but were not offered loans were also randomly chosen into the sample (Control 1). Finally, 30 households that did not form groups were randomly chosen to be included (Control 2).

Our analysis is restricted to the 2070 sample households who owned less than or equal to 1.5 acres of land. We conducted surveys every four months over a three year period. The high frequency of the data collection helped minimize measurement error. There was no attrition in the sample over the three years. In each sample household the same respondent answered survey questions in each round.

Panel A in Table 2 provides checks of balance across the villages randomly assigned to the TRAIL versus GBL treatment arms. As can be seen, there were no significant differences in village-level characteristics across the two groups. Within each treatment category, Panel B checks whether the randomization of selected households (recommended households in TRAIL villages/participating households in GBL villages) into Treatment and Control 1 groups led to a balance of household characteristics. For most characteristics, we see only minor differences across households. The F-statistic shows that we cannot reject the joint hypothesis of no differences across the two arms in either the TRAIL or GBL villages.

Table 3 describes credit market transactions that took place during September–December 2010 in all sample households that owned less than 1.5 acres of land. Since this was the planting season for potatoes, the crop with the highest working capital requirements in this region (as shown below in Table 4), these data provide a picture of the main sources of agricultural credit, and characteristics of the loans. The sample households self-reported all borrowing, regardless of source or loan purpose. We present here data on both total borrowing and borrowing for agricultural purposes.[17] Nearly 67 percent of sample households borrowed in this 4-month period. Traders and moneylenders provided 63% of all agricultural credit and thus were the single most important lender category. Credit cooperatives provided about a quarter of the agricultural credit, but they loaned mainly to households with relatively larger landholdings.[18]

The average interest rate on loans from traders and moneylenders was 26%, substantially above the 18% interest rate charged on the TRAIL and GBL program loans. Loans from family and friends were also more expensive than the program loans.[19] The average duration of loans from traders

[17] Importantly, we use our detailed survey data documenting the purchase of inputs to ensure that all purchases of inputs on trade credit are included in borrowing.

[18] Consistent with the fact that this region had low MFI penetration at the time our intervention began, a very small share of the overall credit taken by our sample households came from MFIs.

[19] Note, we do not consider loans where the repayment amount due was reported to be equal to the principal, since these loans could include insurance features.

and moneylenders was 4 months, reflecting the 4-month agricultural cycles in this area. Loans from family and friends were given for about 6 months. It was extremely rare for any of the informal loans to be secured by collateral. Cooperatives and government banks charged substantially lower interest rates, required more collateral and had longer average durations. However the share of informal lenders in agricultural credit became progressively larger as household landholding decreased from 1.5 acres to zero. Landless households received 87% of their agricultural credit from them, and only 6% from cooperatives (statistics available upon request). Presumably this is because cooperatives lend against collateral: more than three quarters of cooperative loans were collateralized.

Table 4 describes the mean characteristics of the major categories of crops grown by sample farmers during the three years of our study. Paddy was grown two or three times a year, on an average of 0.47 acres of land. Potatoes and sesame are both winter crops planted only once a year, and the average farmer planted each on similar quantities of land: potatoes on 0.31 acres and sesame on 0.21 acres. The table makes it clear that potatoes were the highest-value crop for the villages in this study: they accounted for a significant proportion of acreage, had the highest working capital needs, and generated nearly three times the value-added per acre of other major crops.

3 Theoretical Model of Selection

Our model is based on two key features: borrower heterogeneity, and a segmented informal credit market. Borrowers vary in (exogenously-determined) ability; more able borrowers have lower default risk and higher productivity. Ability variations could reflect either differences in total factor productivity, such as experience or farming skill or in the ownership of complementary fixed factors, such as land or household labor stock. Any selection-based exploration of output or income effects of microcredit must incorporate such heterogeneity in borrower ability.[20] The model abstracts from moral hazard, although similar results can be obtained in extensions that incorporate moral hazard (presented in previous versions of this paper). Defaults arise from incidents of crop failure (such as a pest attack) combined with limited liability: when their crop fails, farmers do not have the means to repay their loans. More able farmers are less likely to experience crop failure because they are better at preventing the pest attack. The risk of crop failure is not correlated across farmers. Besides productivity, the model incorporates associated variations in default risk in order to explain the TRAIL agent's induced selection choices.

Each farmer endogenously chooses the scale of cultivation, measured by area cultivated or expenditure on variable inputs. Conditional on their crop succeeding, more able farmers are more productive insofar as they produce more output from a given scale of cultivation. Specifically, a farmer of ability i experiences crop failure with probability $(1 - p_i) \in (0, 1)$ and produces nothing; otherwise he produces $\theta_i f(l)$ where l denotes the level of input ($\equiv$ loan size) chosen by the farmer. The production function f is smooth, strictly increasing and strictly concave with $f'(0)$

[20]Thus "ability" in our model represents more than just intrinsic characteristics of a farmer, but also includes human capital that could have been acquired over time (before the study began), and physical capital (which we assume remains fixed during the study), all of which may contribute to higher productivity and higher likelihood of crop success.

large enough to ensure interior production for all parameter values and ability levels. Both p_i and θ_i are non-decreasing in i, while their product (or expected productivity) $\bar{\theta}_i \equiv p_i\theta_i$ is strictly increasing. It will turn out that the limited liability constraint will never bind in the absence of a crop failure: farmers will always cultivate on a scale that generates sufficient output to repay their loans. Informal lenders are able to monitor whether their borrower's crop succeeds, and can impose sufficient penalties to deter voluntary default. Hence the default risk of a farmer of ability i is $1 - p_i$.

In the simplest version of the model, there are only two possible ability levels: high ($i = H$) and low ($i = L$), with $H > L$. A given proportion μ_H of borrowers are highly able. Extension to the case of more types is straightforward. To keep the exposition simple we restrict attention to the two-type case for the time being. In Section 3.5 we allow for specific functional forms and for ability to vary continuously.

3.1 Pre-Intervention Informal Credit Market

Each village is partitioned into S different segments on the basis of physical or social proximity. These can be thought of as hamlets, neighborhoods or networks. There are N borrowers in the village divided equally across these S segments, and each segment has the same proportion of H type borrowers. Each segment also has at least two informal lenders who can distinguish borrower types in their own segment, but not in any other segment. All lenders have the same cost of capital ρ per unit loaned, and face no capacity constraints. They compete with one another in Bertrand fashion to make credit offers consisting either of an interest rate (with the borrower deciding how much to borrow), or of a loan size and interest rate pair. The location of each agent in the village is determined exogenously.

Standard arguments imply that the lenders in any given segment will specialize in lending to highly able borrowers in their own segment, and will compete with each other so that in equilibrium they will offer them any amount at interest rate $\frac{\rho}{p_H}$. Low ability borrowers will be able to borrow from any lender in the village at the interest rate $\frac{\rho}{p_L}$, because all lenders will be willing to lend to any borrower in the village at this rate.[21]

Thus, before the MFI intervention, borrower of type i will borrow $\bar{l}_i$ where

$$\bar{\theta}_i f'(\bar{l}_i) = \rho \tag{1}$$

which is a Walrasian allocation. The segmentation of the market has no consequence for the allocation. However, segmentation affects the outcomes of the TRAIL intervention, to which we now turn.

[21] An informal lender will not be willing to lower the interest rate below $\frac{\rho}{p_L}$ for any low ability borrower in his own segment. He will not offer borrowers from other segments an interest rate below $\frac{\rho}{p_L}$ because the only borrowers who would accept that offer would be the low ability ones, resulting in losses.

3.2 TRAIL Intervention

Suppose now that the MFI enters and offers loans at interest rate r_T which is below ρ, the cost of capital for informal lenders. The MFI's comparative advantage over the informal lenders is its lower capital cost. However, it suffers from an informational disadvantage: it is unable to identify the ability of any given borrower. To overcome this, it randomly selects an informal lender, and appoints him as its agent. The agent is asked to recommend to the MFI n borrowers from the village as potential borrowers for TRAIL individual liability loans at interest rate r_T. The MFI then offers loans to a a randomly selected fraction of those recommended. The agent is paid a commission at the rate of $m \in (0, 1)$ per unit of interest repaid by the borrowers he recommended. This incentivizes the agent to recommend borrowers who have a lower risk of crop failure. As with informal loans, we assume that the borrower always has the incentive to repay the loan, so that there is no voluntary default.[22]

The TRAIL agent's selection incentives are as follows. Assuming that he does not collude with borrowers, he tries to maximize the likelihood that the TRAIL loans are repaid.[23] To achieve this, his most-preferred borrowers are the H-type borrowers from his own segment. His second preference is for randomly chosen borrowers from other segments, and this is followed finally by L-type borrowers in his own segment. If $n \leq \frac{N}{S}\mu_H$, then all the borrowers he recommends are H-type from his own segment. Otherwise, he recommends all the H-type borrowers from his own segment and then fills the remaining slots with randomly chosen borrowers from other segments.[24]

We assume that the TRAIL loans do not crowd out the informal loans that the borrowers already have from informal lenders.[25] We shall empirically verify the validity of this assumption. We also simplify by assuming that the TRAIL credit limit is not binding: each farmer's desired TRAIL loan size is smaller than the amount the MFI offers. The main conclusions continue to apply when the limit is binding for some borrowers.[26]

We can now predict the impact of the TRAIL intervention. A selected farmer of ability i will select a TRAIL loan l_i^T satisfying

$$\bar{\theta}_i f'(\bar{l}_i + l_i^T) = p_i r_T \tag{2}$$

Conditions (1) and (2) can easily be used to compare *levels* of borrowing, output and farmer income across types, both before and after the intervention, as stated in the lemma below.

[22]This can be because defaulting borrowers are cut off from future access to TRAIL loans, or because the informal lender pressurizes the borrower to repay.

[23]We shall discuss below how the analysis changes if the agent colludes with borrowers.

[24]This is under the reasonable assumption that the total population of other segments exceeds n.

[25]This could be because borrowers are uncertain about how long the TRAIL intervention will be available and so are reluctant to disrupt their pre-existing credit channels. Alternatively, TRAIL loans may not be close substitutes for informal loans, which have more flexible durations or repayment terms.

[26]A binding credit ceiling will not affect the default risk, so leaves the TRAIL agent's selection incentives unaffected. If the ceiling were binding for both high and low ability borrowers, the TRAIL loan size would be the same for both, while the higher ability type would borrow more before the TRAIL scheme was introduced. This would imply that the loan treatment effect was decreasing in ability. Instead we see that the loan treatment effect is increasing in ability. It follows that even if the ceiling is binding at all, it cannot bind for the low ability type. In this case it can be readily be verified that parts (a) and (b) in Lemma 2 will continue to apply. In the empirical analysis these are the two parts that turn out to be relevant.

Lemma 1 Selection (Comparing Levels): *Higher ability types borrow, produce and earn more than lower ability types, both before and after being offered the TRAIL loan.*

The less trivial question is how treatment effects on borrowing, output or income vary by borrower type. This is ambiguous in general. Starting with the loan treatment effect, the question is will more able farmers take larger TRAIL loans? There are three relevant forces here:

(a) Productivity Difference: More able farmers have higher productivity, so they derive larger benefits from expanding the scale of cultivation;

(b) Diminishing Returns: More able farmers produced more before the intervention, and so they have a lower marginal rate of return to expanding cultivation, controlling for productivity differences;

(c) Subsidy Difference: More able farmers paid a lower interest rate on the informal market before the intervention, so the intervention lowers their interest rate by less.

The productivity difference induces more able farmers to take larger TRAIL loans, but the diminishing returns and smaller interest rate subsidy work in the opposite direction. As a result it is unclear whether the overall treatment effect would be larger for more able types.

Consider the case where high and low ability farmers are equally productive, so that they only vary in default risk. Then it follows from the above that the loan treatment effect will be decreasing in ability.[27] Now introduce productivity differences, so that θ_i increases in i. Then higher ability borrowers who are offered TRAIL loans borrow a larger total volume $(\bar{l}_i + l_i^T)$. The pre-intervention scale of borrowing depends entirely on expected productivity $\bar{\theta}_i$. Therefore if expected productivity $(\bar{\theta}_i)$ is constant and productivity (θ_i) accounts for more of it, so that the crop success rate (p_i) accounts for less of it, then total borrowing after the intervention $(\bar{l}_i + l_i^T)$ increases more steeply in ability i than pre-intervention borrowing $(\bar{l}_i)$ does. This means that loan treatment effects increase in ability. In the limiting case where crop risk does not vary at all with ability, we show below that the loan treatment effect must increase in i. Hence the relative importance of productivity variations relative to crop risk variations in ability determines how loan treatment effects vary with ability.

In the following result, we restrict attention to production functions satisfying a *Regularity Condition* (RC): $\frac{-f''}{f'}$ is decreasing. This condition is satisfied by the constant elasticity function ($f(l) = \frac{1}{\alpha}l^{\alpha}$ with $\alpha < 1$, $\alpha \neq 0$, which corresponds to the logarithmic function, as well as the exponential function ($f(l) = \Gamma[1 - \exp(-al)]$ with $a > 0$).

Lemma 2 Selection Effects (Comparing TRAIL Treatment Effects Across Types): *Suppose that the production function satisfies RC, and that expected productivity $\bar{\theta}_i$ is strictly increasing in ability i.*

[27]To see why, note that any given borrower of type i selects the TRAIL loan size $l = l_i^T$ to maximize net income conditional on crop success $\theta_i f(\bar{l}_i + l) - r_T l$. If there are no productivity differences, θ_i does not vary with i: then all ability types would have the same aggregate borrowing, cultivation, output and income (conditional on crop success). Since higher ability types borrow more before the credit intervention, the loan treatment effect would decrease in i.

(a) If the loan treatment effect is rising in ability, then output treatment effect will also be rising in ability.

(b) If variation in productivity accounts for all (or most) of the variation in expected productivity (so that the crop success probability p_i is entirely or nearly independent of ability), then loan, output and income treatment effects will be rising in ability,

(c) If all (or most) of the variation in expected productivity is accounted for by variation in the probability of crop success (so that productivity is entirely or nearly independent of ability), then loan and output treatment effects will be falling in ability.

The proof of Lemma 2 is in the Appendix. Parts (b) and (c) show that how the treatment effects vary with ability depends on whether productivity or crop risk is more sensitive to variations in ability.[28]

The empirical analysis in subsequent sections will examine how loan, cultivation and income treatment effects vary with ability. The results above help to see why the model must incorporate variations in both default risk and productivity. If we had assumed farmers vary only in default risk, part (c) of Lemma 2 shows that TRAIL treatment effects would be falling in ability, which would have unduly restricted the predictions of the model and rendered it unable to accommodate the opposite pattern. If instead farmers vary only in productivity, then we would be unable to explain the TRAIL agents' selection patterns, because the agent is incentivized on repayment rates and not on the borrower's output.

Importantly the model enables us to empirically disentangle the two sources of variation: differences in informal interest rates reflect variations in default risk, and, given Lemma 2, the pattern of variation of TRAIL treatment effects then reveals the importance of productivity differences. For example, if we find that treatment effects are rising while interest rates are falling in ability, then we can infer that higher ability farmers have lower default risk and are also significantly more productive.

3.2.1 Collusion between the TRAIL agent and borrowers

Now consider the consequences of corruption, where the TRAIL agent can charge bribes in return for recommendations. Loan sizes could also be collusively chosen, so that recommended TRAIL borrowers internalize the larger commissions that the agent would earn if the loan were to become larger.

In this case, the effective interest rate on the loan for the coalition would be $(1-m)r_T$ (where m is the agent's commission rate) instead of the r_T from the non-collusive equilibrium. Lemma 2 would continue to hold, with the effective TRAIL interest rate adjusted from r_T to $(1-m)r_T$, as above. If productivity variations are larger than default risk variations, case (b) applies and the borrower income treatment effects increase in ability. Then high ability borrowers benefit more

[28] In case (c) we are not able to provide a definite result about how treatment effects on farm income vary across types. It can be shown that they decrease in ability if the scale of the TRAIL loans is small enough, i.e., when $[\frac{\rho}{p_L} - r_T]$ is not too large.

from the loan than low ability borrowers, and are willing to pay larger bribes. Thus collusion reinforces the agent's incentive to recommend high ability borrowers.[29]

3.3 GBL Intervention

As is standard in the literature (see for example Besley and Coate, 1995, Ghatak, 1999, 2000), we simplify the analysis by assuming that each GBL group consists of two members. The MFI requires individuals to self-select into groups. Group members then apply for a joint liability loan, which is offered at the same interest rate r_T as the TRAIL loan. Each member is potentially liable for the loans of both members. In addition, the GBL program requires members to periodically attend group meetings and meet savings targets. The cost of meeting these requirements varies idiosyncratically in the population and is uncorrelated with their type: we assume the cost for any borrower c is drawn from a distribution with positive density g over the nonnegative reals. As in the analysis of the TRAIL scheme, we abstract from repayment incentives, and assume that borrowers honor their obligations whenever their own project does not fail.

In contrast to Ghatak (1999, 2000), the scale of cultivation and hence the loan size is variable. Consistent with Ghatak's formulation we assume that members of a group cooperate, i.e. can make side payments without any friction in order to internalize externalities they exert on each other. Then the loan size choices l^G_{ij}, l^G_{ji} for any group (whose members have types i, j) will maximize the sum of their respective *ex ante* payoffs: $\bar{\theta}_i f(\bar{l}_i + l^G_i) + \bar{\theta}_j f(\bar{l}_j + l^G_j) - r_T[p_i\{l^G_i + (1 - p_j)l^G_j\} + p_j\{l^G_j + (1 - p_i)l^G_i\}]$, implying

$$\bar{\theta}_i f'(\bar{l}_i + l^G_{ij}) = [p_i + (1 - p_i)p_j]r_T \tag{3}$$

The expected value of the extra liability that group member j bears in the event that i's crop fails is $(1 - p_i)p_j r_T l_i$. This "joint liability tax" raises the effective cost of the GBL loan relative to the TRAIL loan. So the GBL borrower chooses a lower scale of borrowing than the TRAIL borrower of the same ability. Hence we obtain

Lemma 3 Incentive Effect (Comparing TRAIL and GBL Impacts for a Given Borrower Type)*: For any given ability type, the TRAIL treatment impact on loan size, cultivation scale, output and income is larger than the GBL treatment impact.*

Treatment effects on borrowing and income will therefore be smaller for GBL loans than for TRAIL loans, controlling for ability. A similar effect would arise if the model were extended to incorporate help or monitoring by the TRAIL agent that enhances productivity by more than similar services by other group members, or MFI officers.

As they have similar costs of attending group meetings and meeting savings requirements, *both* high and low ability borrowers have an incentive to participate in the GBL scheme. To see this, consider first a homogenous group, i.e. one in which both members are of type i. Each group

[29] This would obtain regardless of whether the collusion game were modeled cooperatively with stable matching followed by Nash Bargaining, or non-cooperatively, where either side makes a take-it-or-leave-it offer to the other. We omit the details here.

member faces an expected interest rate of $p_i(2-p_i)r_T$, which is lower than what she pays in the informal market, since $r_T < \frac{\rho}{p_i(2-p_i)}$.[30] Hence, homogenous groups of either type would prefer a GBL loan to the status quo. If positive assortative matching does not obtain, heterogenous (H, L) groups could also form.[31] Either way, both low and high risk types would join groups and apply for GBL loans.

The composition of the GBL applicant pool would depend on how the benefits to different groups were rank-ordered. However, the key point is that the proportion of low ability GBL applicants will be bounded away from zero: even with positive assortative matching and the resulting homogenous groups, both high ability groups and low ability groups have an incentive to borrow, and with negative assortative matching even mixed groups would form. Thus, unlike the TRAIL scheme where the agent acts a gate-keeper, there is no mechanism in the GBL scheme that keeps low ability borrowers out. We therefore expect the TRAIL agent to recommend a larger proportion of high ability borrowers than those who self-select into the GBL scheme.[32]

Lemma 4 Differences in Selection Patterns between TRAIL and GBL*:*

(a) If $n \leq \frac{N}{S}\mu_H$, all TRAIL borrowers are H-type, but only a fraction of GBL groups are of H-type.

(b) If $n > \frac{N}{S}\mu_H$, the proportion of borrowers who are H-type in the TRAIL scheme is weakly larger than μ_H. The TRAIL scheme also has more H-type borrowers than the GBL scheme, unless GBL treatment effects are rising in ability and n is sufficiently large relative to $\frac{N}{S}\mu_H$.

3.4 Predicting and Decomposing TRAIL-GBL Differences in Average Treatment Effects

We can express the average treatment effect on any given outcome of intervention v (where $v = \text{T}$ if the scheme is TRAIL, and $v = \text{G}$ if the scheme is GBL) as an average of the treatment effects for different borrower types, using as weights the proportion of selected borrowers that belong to the type, as follows:

$$T^v \equiv \omega^v T_H^v + (1-\omega^v)T_L^v \tag{4}$$

where for intervention v, T_i^v denotes the treatment effect on a type i borrower, T^v denotes the average treatment effect and ω^v denotes the fraction of H types selected. The difference between TRAIL and GBL average treatment effects can then be expressed as

$$T^T - T^G \equiv \quad [\omega^G(T_H^T - T_H^G) + (1-\omega^G)(T_L^T - T_L^G)] + (\omega^T - \omega^G)(T_H^T - T_L^T) \tag{5}$$

The difference in average treatment effects is the sum of two expressions. The first we call the Incentive Effect, a weighted average of the 'incentive effect' for various types, using as weights the

[30]This follows from the fact that $p_i(2-p_i) < 1$, for any i.

[31]In this setting where loan sizes are endogenously chosen, it is difficult to pin down the exact conditions under which positive assortative matching would result.

[32]This is provided the number of recommendations required does not greatly exceed the number of high ability borrowers in the agent's own segment of the informal credit market.

selection likelihoods in the GBL scheme. The second which we refer to as the Selection Effect, is the product of the difference in TRAIL treatment effects between the two types, and the difference in the proportion of H-types between the two interventions. From this and the preceding lemmas it follows that

Lemma 5 Sufficient Condition for Comparing Average Treatment Effects: *The average treatment effect of TRAIL loans is larger than the average treatment effect of GBL loans if TRAIL treatment effects increase in ability, and the TRAIL agent's recommendations contain a larger proportion of H types than the borrowers who self-select into the GBL scheme (e.g. if n is smaller than* $\frac{N}{S}\mu_H$*, or is not much larger).*

Note that this is a sufficient condition, but is not necessary. The purpose of this lemma is to show that the model provides a possible explanation for the larger average treatment effects of the TRAIL scheme than the GBL scheme.

3.5 Specific Functional Forms

The results in the lemmas above depend on assumptions on unknown parameters, and on covariations between observable variables and farmer ability, which is unobserved by the researcher. As a result they are not directly testable. For the empirical analysis we therefore impose a specific functional form that allows us to estimate ability from data we do observe, so that we obtain testable predictions. This also allows us to evaluate the respective roles of selection and incentive effects in driving the difference between the treatment effects of the TRAIL scheme and the treatment effects of the GBL scheme.

We assume the production function is Cobb-Douglas:

$$Y = \theta^{1-\gamma}[\frac{1}{1-\alpha}l^{1-\alpha}]\epsilon \tag{6}$$

where θ denotes ability, l is the scale of cultivation chosen by the farmer, $\gamma, \alpha \in (0, 1)$ and ϵ is an error term which we assume is uncorrelated with ability.

The probability of crop success is given by

$$p(\theta) = P\theta^{1-\nu} \tag{7}$$

where $\nu \in (0, 1)$ and P is the average crop success rate or yield within the village. To keep probabilities between zero and one, we impose an upper bound $\Theta < \infty$ on ability and then restrict $P \leq [\Theta]^{\nu-1}$. A particular example of this is $P = \chi[\bar{a}]^{\nu-1}$ for some $\chi \in (0, 1)$, so that

$$p(a) = \chi[\frac{\theta}{\Theta}]^{1-\nu} \tag{8}$$

Note the following features of this specification:

(a) If ν is close to 1 while γ is not close to 1, then most of the variation in expected productivity is driven by variation in productivity rather than default risk, corresponding to case (b) of Lemma 2. Conversely, if γ is close to one while ν is not, then most of the variation is accounted by default risk, and case (c) holds.

(b) Previously we considered only two borrower types: high and low. In this version ability varies continuously. So we keep track of how pre-, post- and treatment effects vary with ability, and can construct a continuous ability index.

(c) To simplify exposition, we abstract from the error term ϵ. This is justified if the farmer is risk neutral and does not observe the realization of ϵ before he decides on the scale of borrowing and cultivation. Alternatively, the farmer may know the realization of ϵ in advance, but ϵ is uncorrelated with ability.

A control group farmer of ability θ borrows from informal lenders, and so maximizes $\theta p(\theta)\frac{1}{1-\alpha}l^{1-\alpha} - \rho l$. This gives us an expression for the scale of cultivation l^C.

$$\log l^C = \frac{1}{\alpha}\log A + \frac{1}{\alpha}[\log P - \log \rho] \tag{9}$$

where

$$A \equiv \theta^{2-\gamma-\nu} \tag{10}$$

varies monotonically with ability, which varies across households. In what follows below we will therefore use A or θ interchangeably to measure ability. The second term on the right-hand-side of (9) includes covariate shocks to yields and the cost of capital, which varies at the village-year level, but not across households within a given village-year.

A TRAIL treated farmer of ability θ (or equivalently measured by A as in equation (10) above) selects the TRAIL loan l^* to maximize $p(\theta)\theta\frac{1}{1-\alpha}[l^C + l^*]^{1-\alpha} - p(\theta)r_T l^*$, implying that

$$\log l^* = \delta\frac{1}{\alpha}\log A - \frac{1}{\alpha}\log r_T \tag{11}$$

where $l^T \equiv l^C + l^*$ denotes aggregate scale of cultivation for treated farmers, and

$$\delta \equiv \frac{1-\gamma}{2-\gamma-\nu} \tag{12}$$

which lies between 0 and 1. We see here that the expected cost of borrowing increases in ability for treated borrowers but not for control borrowers. As a result the scale of cultivation varies less sharply with ability for TRAIL treated borrowers than for control borrowers. The intuitive reason is that informal lenders are able to offer more able borrowers lower interest rates, unlike the MFI.

Returning to condition (a) from the above, if ν is close to 1 while γ is not close to 1, then most of the variation in expected productivity is driven by variation in productivity rather than default risk and case (b) of Lemma 2 applies. We see from equation (12) above that this also implies that δ is close to 1. Therefore in the empirical analysis we will check the value of δ, and if we find that it is close to 1 we will expect TRAIL treatment effects to be larger for households with greater ability.

Averaging across groups, the effective borrowing cost for a member with ability θ_i is $p(\theta_i)[1-\bar{p}]+\bar{p}$ which is increasing linearly in $p(\theta_i)$. Here $\bar{p}$ denotes the average success probability.

Similar expressions also arise for the expected output of treated and control households. For control households

$$\log E[Y^C] = \frac{1}{\alpha}\log A + \frac{1}{\alpha}[\log P - (1-\alpha)\log\rho] - \log(1-\alpha) \tag{13}$$

while for TRAIL treated households:

$$\log E[Y^T] = [\delta\frac{1}{\alpha} + (1-\delta)]\log A + \log P - \frac{(1-\alpha)}{\alpha}\log r_T - \log(1-\alpha) \tag{14}$$

and again we see that log output varies less with ability for treated households than for control households.

In GBL villages, expressions (9) and (13) continue to apply for control households. For treated households, however, the expressions for effective cost of borrowing depend on the pattern of matching and do not have closed-form solutions. Therefore we cannot estimate the ability of GBL treated households without making additional assumptions. Fortunately for our subsequent analysis we do not need ability estimates for these households.

3.5.1 Estimating Ability

From this point onwards, we denote households by h. We assume that the ability of household h depends on observable farmer characteristics $X_{kh}, k = 1, \ldots$ such as land owned, number of household members engaged in cultivation, gender, caste and religion of head:

$$A_h = T_h X_{1h}^{\psi_1} X_{2h}^{\psi_2} \ldots \tag{15}$$

where $\psi_k > 0$ are unknown parameters to be estimated, and T_h is a household specific component which is unobservable to us and MFI officials, although it may be observed by borrowers and agents. Household characteristics are assumed to be time-invariant.

From equations (15) and (7), the scale of cultivation or output of control group household h located in village v in year t satisfies:

$$\log l_{ht}^C = \frac{1}{\alpha}[\log T_h - \log\rho_{vt} + \log P_{vt}] + \frac{1}{\alpha}\sum_k \psi_k X_{kh} \tag{16}$$

thereby generating the regression specification

$$\log l_{ht}^C = \sum_k \beta_k X_{kh} + \mu_{vt} + u_h + \epsilon_{ht} \tag{17}$$

which can be estimated by ordinary least squares or random effects regressions. Under the strong assumption that observable household characteristics are uncorrelated with unobservable characteristics or the error term, the coefficients $\beta_k \equiv \frac{1}{\alpha}\psi_k$ provide consistent estimates of ability. They can be used to construct a continuous ability index equal to the predicted value

$$\frac{1}{\alpha}\log A_h = \sum_k \hat{\beta}_k X_{kh} \tag{18}$$

for *both* control and treated households. An alternative procedure that allows for both observable and unobservable components of ability and requires weaker assumptions, estimates ability as the household fixed effect in regressions of cultivation scale or output, as follows:

$$\log l^C_{ht} = \zeta_h + \mu_{vt} + \epsilon_{ht} \tag{19}$$

$$\log l^T_{ht} = \delta\zeta_h + K + \mu_{vt} + \epsilon_{ht} \tag{20}$$

where K is a constant representing the mean difference $\log \bar{\rho} - \log r_T$ in the cost of borrowing between control and treated households, and subscript v denotes the village in which h resides. We then obtain estimates of ability

$$\frac{1}{\alpha}\log A_h \equiv \zeta_h \tag{21}$$

For control households, equation (19) delivers estimates of ζ_h, but for treated households equation (20) delivers estimates of $\pi_h \equiv \delta\zeta_h + K$. To isolate ζ_h for treated households we utilize the fact that households recommended by the TRAIL agent were randomly assigned to treatment, so that Treatment and Control 1 households are drawn from the same distribution of ζ_h. It follows that both the Treatment and Control 1 groups must have the same mean and variance of ζ_h. Hence

$$E[\pi_h|h \in T] = K + \delta E[\zeta_h|h \in T] = K + \delta E[\zeta_h|h \in C1] \tag{22}$$

and

$$\text{Var}[\pi_h|h \in T] = \delta^2 \text{Var}[\zeta_h|h \in T] = \delta^2 \text{Var}[\zeta_h|h \in C1] \tag{23}$$

These two moment conditions allow us to estimate δ and K (where hats denote sample estimates) as follows:

$$\hat{\delta} = \left[\frac{\hat{\text{Var}}[\pi_h|h \in T]}{\hat{\text{Var}}[\zeta_h|h \in C1]}\right]^{\frac{1}{2}} \tag{24}$$

$$\hat{K} = \hat{E}[\pi_h|h \in T] - \hat{\delta}\hat{E}[\zeta_h|h \in C1] \tag{25}$$

$$\hat{\zeta}_h = \frac{\hat{\pi}_h - \hat{K}}{\hat{\delta}} \tag{26}$$

We can then examine how the estimated TRAIL treatment effect on farm value-added varies with $\hat{\zeta}_h$, by regressing the farm value-added in TRAIL villages on the treatment dummy, interacted with ability. This reveals the heterogeneity of the TRAIL treatment effect with respect to ability, denoted by $T^v(\zeta)$.

The exact analytical expression for $T^v(\zeta)$ is somewhat cumbersome; it is neither linear or log-linear in ζ. We can estimate a "non parametric" version by discretizing the ability index. We divide the range of estimated ability values into quartiles and then replace the ability index $\hat{\zeta}$ with dummy variables indicating the quartile it belongs to ($q_i = 1$ if and only if $\hat{\zeta}_i \in (\hat{Z}_i, \hat{Z}_{i+1}), i = 1, \ldots, 4$). From a regression of farm value-added on interactions of the treatment dummy with the ability quartile q_i, we can estimate TRAIL treatment effects $Tr^T(q_i)$ within each quartile q_i.

Finally, the difference between the TRAIL and the GBL treatment effects can be decomposed as follows. If we denote the loan scheme with v, the average treatment effect is

$$\text{Tr}^v \equiv \int \sigma^v(\zeta) T^v(\zeta) d\zeta \tag{27}$$

where $\sigma^v(.)$ denotes the density of the ability distribution of households selected to participate in scheme v. Hence the difference between the two average treatment effects can be decomposed:

$$\mathrm{Tr}^T - \mathrm{Tr}^G = \int [\sigma^T(\zeta) - \sigma^G(\zeta)] T^T(\zeta) d\zeta + \int \sigma^G(\zeta)[T^T(\zeta) - T^G(\zeta)] d\zeta \quad (28)$$

where v takes value T for the TRAIL scheme and G for the GBL scheme. The first term on the right-hand-side, the Selection Effect, is the one we compute. The second term is the Incentive effect. A discrete approximation of the Selection effect is

$$S = \sum_i [\sigma^T(q_i) - \sigma^G(q_i)] Tr^T(q_i) \quad (29)$$

Note that this requires only an estimate of difference in selection proportions between the TRAIL and GBL schemes and the heterogenous TRAIL treatment effects. Specifically, we do not need to estimate heterogeneous GBL treatment effects.

3.6 Summary of Testable Predictions

Before proceeding to the empirical analysis, it is helpful to summarize the theoretical predictions that can be tested.

Prediction 1 ***TRAIL Selection Patterns:*** *The borrowers whom the TRAIL agent recommends from his own clientele are more able than those whom he recommends from outside his clientele, and are also more able than those within his clientele whom he does not recommend.*

Prediction 2 ***Ability-Informal Interest Rate Relationship:*** *Higher ability borrowers pay lower interest rates in the informal market.*

Prediction 3 ***Compression:*** $\delta < 1$*; or, the scale of cultivation varies less with ability for treated borrowers than for control borrowers.*[33]

Prediction 4 ***Treatment Effect Heterogeneity:*** *If the TRAIL treatment effect on borrowing is rising in ability, so is the TRAIL treatment effect on output.*[34]

Prediction 5 ***Selection Effect:***

(a) The Selection Effect is smaller than the average treatment effect difference.[35]

(b) If the ability distribution among TRAIL selected borrowers first order stochastically dominates the ability distribution among GBL selected borrowers, and TRAIL treatment effects are rising in ability, then the Selection Effect is positive, and the average treatment effect in the TRAIL scheme is larger than in the GBL scheme.

[33]This follows from a comparison of equations (9) and (11).
[34]This follows from part (a) of Lemma 2.
[35]This holds because Lemma 3 implies that the Incentive Effect is positive.

4 Empirical Results

We start in Section 4.1 by presenting the average treatment effects in the two loan schemes on borrowers' cultivation, output and farm value-added. This is followed in Section 4.2 by a discussion of the repayment and take-up rates of the loans and the administrative costs and overall financial performance of the two schemes. Next, in Section 4.3 we test the model's predictions, and whether and to what extent the difference in selection patterns can explain the difference in their average treatment effects. Finally, in Section 4.4 we address some ancillary issues, such as the changes in treatment impacts over time, and concerns that TRAIL agents and borrowers might have entered into side-transactions that changed the benefits to borrowers.

4.1 Empirical Results About Loan Treatment Impacts on Borrower Production and Income

To examine the average treatment effects of the two lending mechanisms, we rely on the fact that only a randomly chosen subset of the selected borrowers (recommended by the agent in the TRAIL villages or members of self-formed groups in the GBL villages) were offered the loans. Any difference between households that were recommended but not offered the loan (Control 1 households) and those that were both recommended and offered the loans (Treatment households) must be caused by the loans. Clearly, this estimate is conditional on the selection of these borrowers into the loan scheme and thus is the combined result of the selection pattern and the impact of the loans.

Our regression specification takes the following form:

$$\begin{aligned} y_{hvt} &= \beta_0 + \beta_1 \text{TRAIL}_v + \beta_2(\text{TRAIL}_v \times \text{Control } 1_{hv}) + \beta_3(\text{TRAIL}_v \times \text{Treatment}_{hv}) \\ &+ \beta_4(\text{GBL}_v \times \text{Control } 1_{hv}) + \beta_5(\text{GBL}_v \times \text{Treatment}_{hv}) + \gamma \mathbf{X}_{hvt} + \varepsilon_{hvt} \end{aligned} \tag{30}$$

Here y_{hvt} denotes the outcome variable of interest for household h in village v in year t. The omitted category is the Control 2 group in GBL villages, so that $\hat{\beta}_0$ estimates the mean y_{hvt} for Control 2 households in GBL villages. The other coefficients each estimate the level of y_{hvt} for a different group, relative to these GBL Control 2 households. Here $\hat{\beta}_3 - \hat{\beta}_2$ estimates the treatment effect in the TRAIL scheme and $\hat{\beta}_5 - \hat{\beta}_4$ estimates the treatment effect in the GBL scheme. All treatment effects are intent-to-treat estimates because they compare the outcomes of households *assigned* to Treatment and Control 1 groups, regardless of actual take-up.[36] The coefficients $\hat{\beta}_2$ and $\hat{\beta}_4$ measure differences between Control 1 and 2 households within TRAIL and GBL villages, respectively. $\mathbf{X}_{hvt}$ is a set of additional controls, including land owned by the households, caste, gender and educational attainment of the household head, two year dummies to control for secular changes over time and a dummy variable indicating whether the village received a separate intervention informing residents about the prevailing market price for potatoes.[37]

[36] Results are qualitatively unchanged if we instead estimate the treatment effects only on households that took up the loans, using assignment to treatment as an instrument for actual participation in the scheme. These results are presented in Table A-4 in the Appendix.

[37] The information intervention was undertaken for a separate project aimed at examining the effect of providing

To reiterate, our sample consists of 2070 households across 24 TRAIL and 24 GBL villages. In each village our sample consists of 10 Treatment, 10 Control 1 and 30 Control 2 households. We exclude households with more than 1.5 acres of land from the estimating sample. Since agricultural activity involves a long delay from planting to harvest, and the harvest could be sold over several months, we aggregate our data to the annual level in order to correctly compute the costs and revenues of each crop. Our unit of observation is then household-year. Standard errors are clustered at the hamlet level.[38]

4.1.1 Treatment effects on Borrowing, Cultivation and Farm Incomes

The regression results corresponding to equation (30) are presented in Tables 5–7. In particular, treatment effects on borrowing are in Table 5, treatment effects on cultivation of and incomes from potatoes are in Panel A of Table 6, effects on cultivation of and incomes from other crops are in Panel B of Table 6, and effects on total farm income are in Table 7.

Since we analyze a large number of outcome variables, the null hypothesis of no treatment effect could be rejected by mere chance for some, even if it were actually true. To correct for this, in each table we follow Hochberg (1988) and report a conservative p-value for an index of variables in a family of outcomes taken together (see Kling, Liebman, and Katz, 2007).[39]

Effects on Borrowing

In column 1 of Table 5 we see that participation in the TRAIL scheme increased the overall borrowing of Treatment households by ₹5050, which is a 90% increase over the ₹5590 mean borrowing by TRAIL Control 1 households. The overall borrowing of Treatment households in the GBL scheme also increased by a statistically significant ₹3732, which is a 92% increase over the mean for GBL Control 1 households.

In column 2 of Table 5 we examine if program loans crowded out agricultural loans from other sources. There is no evidence that this happened in either scheme: the treatment effects on non-program loans are small and statistically insignificant.

information about potato prices to farmers and is similar to the public information treatment described in Mitra, Mookherjee, Torero, and Visaria (2016). Villages were assigned to the information treatment randomly and orthogonally to the credit intervention that is the focus of this paper. The results are unchanged if we do not include this information village dummy in the regression specification.

[38] In Indian villages households are clustered by *paras* (or hamlets/neighbourhoods). Administrative definitions of a village correspond to a collection of 10-15 *paras* geographically isolated from each other. Households within the same *para* tend to be more homogenous, are more likely to interact with each other, and experience geographic shocks to cultivation and market prices that are highly correlated. The results are robust to clustering at the village-level instead (see columns 1 and 2 of Table A-1 and Table A-2). The treatment effects are also unchanged qualitatively if we restrict the sample to the Treatment and Control 1 households only (see columns 3 and 4 of Table A-1 and Table A-3).

[39] The variables are normalized by subtracting the mean in the control group and dividing by the standard deviation in the control group; the index is the simple average of the normalized variables. To adjust the p-value of the treatment effect for an index, the p-values for all indices are ranked in increasing order, and then each original p-value is multiplied by $(m - 1 + k)$, where m is the number of indices and k is the rank of the original p-value. If the resulting value is greater than 1, an adjusted p-value of > 0.999 is assigned.

When we consider both borrowing outcomes together in column 3, we find that TRAIL loans caused a 0.24 standard deviation increase in household borrowing, which is significant according to the more conservative Hochberg test (p-value = 0.025). However, the effect of the GBL treatment is not statistically significant (effect = 0.18 sd, Hochberg p-value = 0.368).

Effects on Cultivation and Farm Incomes

We now check if the increase in agricultural borrowing further led to increases in agricultural activity, output and incomes. Since the loan cycles matched the potato production cycle, we first present the estimated effects on potato cultivation. Note first from column 1 in Panel A of Table 6 that 72 percent of TRAIL Control 1 households cultivated potatoes per year. Although the TRAIL loans did not increase this likelihood of cultivation significantly, column 2 shows they did increase the amount of land placed under potatoes by a statistically significant 28 percent. TRAIL loans also caused borrowers to increase their expenditure on inputs (column 4) and to produce 27 percent greater output (column 3). As a result TRAIL treatment borrowers earned 28% higher revenue (column 5) and 37% higher value-added (column 6) than they otherwise would have. Their imputed net profit from potato cultivation increased by ₹1939, or 41% of the mean (column 7).[40]

Although the GBL loans did not significantly affect households' decisions of whether and how much land to place under potatoes, they did increase expenditure by 27%. However the positive effect this had on revenue, value-added and profits could not be estimated precisely. The point estimate of percent growth in value-added and imputed profit was 14%, in contrast to 37-40% in TRAIL.

In Panel B of Table 6 we consider the acreage and value-added of the other main crops: sesame, paddy and vegetables. TRAIL loans significantly increased the acreage that Treatment households allocated to paddy and sesame. The treatment effect on value-added is also positive for all three crops, but it is only statistically significantly different from zero for sesame. GBL loans did not have significant effects on the acreage or value-added for any of the crops.[41]

Finally, column 1 of Table 7 presents the treatment effects on total farm value-added, computed by aggregating across the four crop categories. The TRAIL loans led to a 25% increase in overall farm value-added over the Control 1 mean. The GBL treatment effect was comparatively negligible (4%) and statistically insignificant. As the penultimate row in column 1 shows, the TRAIL treatment effect on farm value-added was significantly larger than the GBL treatment effect (p-value = 0.077). In column 2 we see that neither the TRAIL nor the GBL loans significantly affected borrowers' non-agricultural incomes. However, when we take both farm and non-farm income into account in column 3, we see that TRAIL loans increased borrower incomes by 11.6 standard deviations (Hochberg p-value = 0.034). GBL loans had no effect (Hochberg p-value > 0.999).

[40] Value-added in column 6 is computed by subtracting from revenues only the costs of purchased or rented inputs. In column 7 we impute the cost of family labor using the average market wage rate for hired labor in the village (an upper bound for the shadow cost of family labor).

[41] Although the magnitude of the effect on paddy is similar to that for TRAIL treatment households, it is imprecisely estimated.

4.1.2 Comparing Productivity of Selected TRAIL and GBL borrowers

Next we compute the rate of return on program loans as the ratio of the treatment effect on value-added to the treatment effect on cultivation cost. Since this is the ratio of two treatment effect estimates, we estimated cluster-bootstrapped standard errors with 2000 replications. As we see in column 4 of Table 7, in the TRAIL scheme, the rate of return on potato cultivation expenses was a statistically significant 110%. The corresponding rate of return in the GBL scheme is estimated as a substantially lower 45%, but is not statistically significant. Across all major crops, TRAIL borrowers earned a statistically significant rate of return on investment in cultivation expenditure of 122%. The estimate for GBL borrowers was 25%, but again, was not statistically different from zero. The estimated rate of return to GBL loans is too imprecisely estimated for us to infer the differences in rate of return in the two schemes with any certainty.

4.2 Loan Performance

4.2.1 Comparing Repayment and Take-up Rates

The financial sustainability of a lending program critically depends on the repayment rates on its loans. Our model does not make clear predictions about how the repayment rates on TRAIL loans and GBL loans would compare. TRAIL borrowers are likely to be more able and therefore have a lower risk of default, but GBL borrowers have the benefit of joint liability so that even if their own projects fail, their group members might repay on their behalf.[42] As seen in Panel A of Table 8, over the three years of the intervention, both TRAIL and GBL borrowers repaid their loans within one month of the due date more than 95% of the time, well within the range of the industry standard. The difference in repayment rates in the two schemes is small and a t-test indicates it is not statistically significant.

Loan take-up rates can be viewed as an indicator of the attraction of the loan product to potential borrowers. We measure these as the proportion of households that were offered a program loan in a given cycle who accepted it. Our finding in Panel A that a higher 86% of the loans offered in the TRAIL scheme than the 75% of the GBL loans offered were accepted, suggests that selected borrowers in the TRAIL scheme expected to gain more from the loans than those in the GBL scheme. Note that since loans in our study were only offered to households that had been pre-selected to participate (either through recommendations by the TRAIL agent or because they had formed GBL groups), these take-up rates cannot be compared with take-up rate measures from other studies where the entire village population is included in the set of eligible borrowers. Comparisons across TRAIL and GBL are more revealing.[43]

We also measure continuation rates, which are the number of households who took a loan in a

[42] A TRAIL loan given to a borrower of ability i would be recovered with probability p_i, whereas a GBL loan given to her would be recovered with probability $p_i + (1 - p_i)p_j$ if this borrower's group member had ability j. Hence controlling for ability, the GBL loan has a higher repayment rate. However if TRAIL borrowers are more able on average, this tilts the comparison the other way, so that the net effect is ambiguous.

[43] Recall that TRAIL and GBL loans were offered to borrowers selected according to different criteria. Therefore differences in the take-up rates reflect not whether the contractual terms of the loans were more or less attractive in general, but whether they were more or less attractive to the borrowers who were selected to participate.

given cycle as a proportion of those who were eligible in cycle 1. Households may have failed to continue in the scheme either because they had repaid less than 50% of the amount due in a previous cycle and become ineligible, or because they chose not to take a loan in this cycle. Column 3 shows that 81% of TRAIL borrowers continued, compared to 69% in GBL scheme on these indicators. Hence the differences in take-up between the two programs did reflect differences in willingness to utilize the loans, rather than differences in eligibility.

A more rigorous test of the difference in these indicators would control for seasonal variations that might affect loan take-up or repayment. Accordingly, we estimate the equation

$$y_{hvt} = \alpha_0 + \alpha_1 \text{TRAIL}_v + \gamma \mathbf{X}_t + \varepsilon_{hvt} \tag{31}$$

where $y_{hvt} = 1$ if household h assigned to treatment in village v accepted the program loan in cycle t (continuation), $y_{hvt} = 1$ if treatment household h in village v who was eligible to receive a loan in cycle t accepted it (take-up), or $y_{hvt} = 1$ if treatment household h in village v repaid a loan taken in cycle t within 30 days of the due date (repayment). Cycle dummies control for season-specific differences. Column 1 in Panel B of Table 8 confirms that the difference in repayment rates is negligible and not statistically significant. Columns 2 and 3 show that loan take-up and continuation rates were about 12 percentage points higher in the TRAIL scheme than the GBL scheme, and that this difference is statistically significant at the 10% level. This result holds whether we use as the denominator the households households that were offered the loan in that cycle (take-up, Column 2), or instead all households that were offered the loan at the beginning of the intervention (continuation, Column 3).[44]

4.2.2 Administrative Costs and Overall Financial Performance

Administrative costs were also lower for the TRAIL scheme. The per-month cost to the MFI of operating the GBL scheme in a village was ₹1463. The cost of running the TRAIL scheme was substantially lower, at ₹68 per village. This difference is largely explained by the fact that the TRAIL scheme did not require group meetings and so had lower personnel and transport costs. Recall that in both schemes the intermediary (the agent in TRAIL villages and the MFI in GBL villages) received 75 percent of interest payments as commission, and the repayment rates were similar. The capital costs of the loans were also the same. It follows that TRAIL loans generated a higher financial return for the lender, as well as higher take-up.

4.3 Testing Theoretical Predictions

We turn now to tests of our theoretical predictions about the patterns of borrower selection in the two schemes, and about how treatment effects varied with borrower ability in the TRAIL scheme.

[44]The denominator for take-up differs from the denominator for the continuation measure because past default would disqualify a treatment household from borrowing subsequently.

4.3.1 Ability Estimates

The first step in this exercise is to estimate farmer ability. In Table 9 we first investigate the observable correlates of household ability by running a household random effects regression that follows equation (17) on the logarithm of potato output (column 1) and on the logarithm of potato acreage (column 4). OLS regressions run in columns 2 and 5 provide qualitatively similar results. We find that larger households, households that owned more land, whose heads were Hindu, did not belong to the lower castes/tribes, and whose primary occupation was cultivation all devoted more land to potato cultivation and produced greater potato output.

Columns 3 and 6 show the results of the household fixed effects regressions corresponding to equation (19). These regressions arerun on all sample households in the TRAIL villages and only Control 1 and Control 2 households in the GBL villages.[45] For each household we have three annual observations. Next, we use the procedure described in equations (24)–(26) to estimate the parameters δ and K needed to adjust the fixed effect ability estimates for treated farmers in TRAIL villages. The fixed effects regressions give us our preferred ability estimates because they are based on both observed and unobserved household characteristics.

In Figure 1, we examine whether the distribution of the estimated ability index conforms with our predictions about the pattern of selection. The left panel uses fixed effects estimates based on the log of potato output; the right panel uses estimates based on the log of potato acreage. Consider first the TRAIL villages. In both panels, the distribution is bimodal and spans a wide range of abilities. As would be expected, the generated distributions for Treatment and Control 1 households in the TRAIL villages are close, since they differ only on the basis of a random draw from the set of borrowers recommended by the TRAIL agent. We can then refer to Treatment and Control 1 households together as Selected households and compare them with Control 2 households. It can be seen that a larger proportion of TRAIL Control 2 households is concentrated around the lower mode while a smaller proportion is concentrated around the higher mode. This indicates the TRAIL agent screened out some low ability farmers.

The graphs for GBL villages in Figure 1 do not include Treatment households because as discussed in footnote 45, we are unable to estimate their ability. Instead we compare only Control 1 with Control 2 households. Although the distribution here is bimodal as well, it has a greater mass in an intermediate range of ability and smaller mass at the two extremes, suggesting that GBL groups excluded both very low and very high ability borrowers.

In what follows, we will use these estimates of ability to test Predictions 1–5.

4.3.2 Test of Prediction 1

Prediction 1 states that the borrowers that the TRAIL agent recommended from within his clientele should be more able on average than borrowers he recommended from outside his clientele, and should also be more able than borrowers he did not recommend from within his clientele. We

[45] As explained on page 19, this is because the differential incentives of GBL loans make it difficult to estimate ability for GBL Treatment households. However GBL Treatment and Control 1 households are drawn from the same ability distribution, since GBL groups were randomly assigned to be either Treatment or Control 1.

test this prediction in Table 10, by running the following regression on the sample of households in TRAIL villages that owned at most 1.5 acres of land:

$$\hat{\zeta}_{hv} = \delta_0 + \delta_1 \text{Recommended}_{hv} + \delta_2 \text{Own Clientele}_{hv} + \delta_3 (\text{Recommended}_{hv} \times \text{Own Clientele}_{hv}) + u_{hv} \tag{32}$$

Here $\hat{\zeta}_{hv}$ denotes the estimated ability of household h in village v. The variable Recommended$_{hv}$ indicates whether the household was recommended by the agent for a TRAIL loan. The variable Own Clientele$_{hv}$ is measured by an indicator for whether the household had borrowed from the agent in the three years prior to the study. Finally u_{hv} denotes an idiosyncratic error term.[46] The regressions control for village fixed effects, so the regression coefficients reflect within-village comparisons. Within-village comparisons are appropriate because each TRAIL agent is restricted to selecting borrowers from a single village. In column 1 we use as the dependent variable the estimate of ability derived from the household fixed effects regression on potato output. In column 2 we use the household fixed estimates from the potato acreage regression.

By interacting the dummy for whether the household was recommended with the dummy variable indicating whether it belonged to the agent's own clientele, we are able to examine the ability of households that were recommended by the agent from within and from outside the agent's clientele. The panel titled "Total Effects" presents the predicted ability of households of these different categories. It is clear that among all the households in the sample, those recommended from within the agent's own clientele had the highest ability. That they were significantly more able than households recommended from outside his clientele (where he would not have observed their ability and so would have selected randomly) is clear from the panel titled "Difference estimates". The last row in the bottom panel also suggests that they were more able than those within his clientele whom he did not recommend, although this difference is not statistically significant.

4.3.3 Test of Prediction 2

In Table 11 we test the prediction that more able borrowers paid lower interest rates on informal loans. We were unable to conduct a baseline survey before the study began, so we restrict this analysis to Cycle 1 loans because they had been negotiated before our intervention began. If the information that particular households had been selected into the loan scheme or were receiving the treatment loans affected the terms at which informal lenders were willing to lend to these households, then households' informal credit market outcomes could potentially be affected in subsequent cycles.

Since the relation between ability and interest rates is likely to be non-linear, we create dummy variables to indicate the quartile of estimated ability that this household belonged to, where the quartiles are computed on the pooled ability distribution across all TRAIL households and the

[46] Measurement errors in ability are included in the error term of the regression. Most of these are likely to be accounted for by measurement error in acreage or output. These are plausibly independent of the regressors because our data is based on farmer surveys, conducted independently of the agent.

GBL Control 1 and Control 2 households.[47] We then run the following regression:

$$r_{hv} = \sum_{i=1}^{4} \mu_i \ \hat{Q}_{hi} + \epsilon_{hv} \tag{33}$$

Here r_{hv} denotes the average interest rate that household h paid on informal loans it had contracted in Cycle 1 and $\hat{Q}_{hi}; i = 1, 2, 3, 4$ is a binary indicator variable that takes the value 1 if household h is in estimated ability quartile i.[48]

Columns 1 and 2 use ability estimates based on the quantity of potato output while columns 3 and 4 use ability estimated off potato acreage. We focus on columns 2 and 4 since these include village fixed effects, which control for village-specific unobserved variation in the informal lenders' cost of capital.[49] In both columns, the difference estimates make it clear that households in lower ability quartiles generally paid higher interest rates than households in higher quartiles. These differences are statistically significant when we compare quartile 1 households with quartile 3 or quartile 4 households.

4.3.4 Test of Prediction 3

When estimating farmer ability in Table 9 we also estimated the parameter δ, which measures the extent to which the correlation between ability and total borrowing varies between TRAIL treated borrowers and TRAIL control borrowers. Consistent with Prediction 3, the point estimates of δ are 0.951 and 0.965 in columns 1 and 4 respectively. As can be seen from a comparison of equations (9) and (11), this indicates that an increase in ability increases total borrowing by more for control borrowers than for treated borrowers. However the estimates are not precise enough to infer whether they are significantly below 1.

4.3.5 Test of Prediction 4

Notice also that while the point estimate of δ is below 1, it is close to 1. This suggests that default risk varies relatively little with ability, so that case (b) rather than (c) of Lemma 2 applies. Accordingly, we expect to find that TRAIL treatment effects increase in ability.

[47] Recall that the different incentives in GBL loans mean that we do not have a clear way of estimating the ability of Treatment households in the GBL scheme. Recall also that the expression for the informal interest rate is $P\frac{\rho}{\theta^{1-\nu}}$ and as discussed above, our estimate of δ indicates ν is close to 1.

[48] These regressions are run on a sample of 661 households. We start with 2400 households in TRAIL and GBL villages. Of these 1303 households borrow in Cycle 1 from money lenders, traders, friends and family. So interest rate data for Cycle 1 exists for 1303 households. We drop households that own more than 1.5 acres of land. We also drop loans where the household reported that the principal borrowed was equal to the amount to be repaid (i.e. where the annualized interest rate was zero), since these are likely to correspond to loans from family and friends, rather than professional lenders. This leaves us with a sample of 872 households in the estimation sample. There are 448 TRAIL households and 424 GBL households in this sample of 872 households. The estimating sample in regression (33) is restricted to Control 1 and Control 2 households only, giving us a sample of 661 households.

[49] Since ability is a generated regressor, we construct bootstrap confidence intervals for our estimates in these regressions using 2000 replications.

In Table 12 we examine the heterogeneity of TRAIL treatment effects with respect to borrower ability. We estimate treatment effects on two different dependent variables: borrowing and farm revenue. In columns 1 and 2 we regress these dependent variables on the treatment dummy interacted with quartile dummies for ability estimates based on the volume of potato output. In columns 4 and 5 the ability estimate is based on the acreage devoted to potatoes. Thus the regression specification takes the form:

$$
\begin{aligned}
y_{hvt} = & \sum_{i=1}^{4} \alpha_{1i} \ \hat{\mathrm{Q}}_{hi} + \sum_{i=1}^{4} \alpha_{2i} \ \mathrm{TRAIL}_v \times \hat{Q}_{hi} \\
& + \sum_{i=1}^{4} \alpha_{3i} \ (\mathrm{TRAIL}_v \times \mathrm{Control\ 1}_{hv} \times \hat{\mathrm{Q}}_{hi}) \\
& + \sum_{i=1}^{4} \alpha_{4i} \ (\mathrm{TRAIL}_v \times \mathrm{Treatment}_{hv} \times \hat{\mathrm{Q}}_{hi}) \\
& + \sum_{i=1}^{4} \alpha_{5i} \ (\mathrm{GBL}_v \times \mathrm{Control\ 1}_{hv} \times \hat{\mathrm{Q}}_{hi}) \\
& + \sum_{i=1}^{4} \alpha_{6i} \ (\mathrm{GBL}_v \times \mathrm{Treatment}_{hv} \times \hat{\mathrm{Q}}_{hi}) \\
& + \gamma \mathbf{X'}_{hvt} + \varepsilon_{hvt}
\end{aligned}
\tag{34}
$$

where $\hat{\mathrm{Q}}_{hi}$ indicates where the household's estimated ability is in quartile i. Equation (34) is an extension of equation (30), but does not include landholding as a regressor because the effect of landholding is included in ability. Year dummies and the dummy variable indicating the information intervention continue to be included. This extension allows us to estimate the respective treatment effects for each ability quartile separately.

Table 12 presents the implied predicted treatment effects for each quartile of estimated ability. Recall from part (*b*) of Lemma 2 that if variation in productivity drives most of the variation in ability, then TRAIL treatment effects on borrowing increase in ability. Columns 1 and 4 confirm that this is true: while a TRAIL Treatment household in the first quartile of estimated ability borrowed ₹656 more than a TRAIL Control 1 household in the same quartile, a household in the second quartile borrowed ₹2832 more, in the third quartile borrowed ₹6327 more, and in the fourth quartile borrowed ₹9474 more. This steady increase in the treatment effect of TRAIL loans on aggregate borrowing by ability is statistically significant, as can be seen in the difference estimates in the bottom panel.

Prediction 4 says that if treatment effects on borrowing increase in ability then treatment effects on output should increase in ability as well. We check this in columns 2 and 5 using aggregate farm revenue as the measure of output. Once again, the difference estimates indicate that the TRAIL loans increased the value of output of more able farmers by more.

4.3.6 Test of Prediction 5

Figure 2 plots the cumulative distribution function of estimated ability for households recommended by the TRAIL agent, and compares this with the corresponding distribution of GBL Control 1 borrowers. Again, GBL Treatment borrowers are not included in this figure: the ability index is not estimated for GBL treated households, as per the discussion on page 19.

In Figure 2 we see that at each ability level in the support of the distribution, more of the GBL selected households lay below this level than TRAIL selected households did. A Kolmogorov-Smirnov test of equality of distributions shows that the distribution of TRAIL selected households first-order stochastically dominates the distribution of GBL selected borrowers (p-value = 0.00).[50]

Given this finding, and the finding in Table 12 that TRAIL treatment effects increase in ability, Prediction 5 states that the Selection Effect is positive and the average treatment effect of the TRAIL loans is larger than that of the GBL loans. We have already shown evidence that both of the conditions for this prediction are met. Table 7 also shows that the average TRAIL treatment effect on farm value-added is significantly larger than the average GBL treatment effect.

We can also use equation (29) to estimate the Selection Effect, as shown in Table 13. Our results indicate that the Selection Effect is indeed positive, and it accounts for 27–36% of the overall difference in average TRAIL treatment effects on farm value-added, depending on how ability is measured.[51]

4.4 Ancillary Issues

We turn now to examine other possible concerns about the TRAIL scheme.

4.4.1 Impact of TRAIL on Transactions with Agents

A frequent concern about intermediary-based schemes is that they may promote corruption and distort the allocation of benefits between the intermediary and intended beneficiaries. For instance, the TRAIL agent may have extracted undue benefits from the borrowers, either by requiring bribes before he recommended them, demanding side-payments, or by manipulating other transactions with them, such as by charging higher input prices or paying lower output prices. Although it is naturally difficult to collect data on bribes or side-payments, we do have detailed data about sample households' input purchases from, output sales to, and borrowing from the

[50] In unreported results, we do not find much evidence for positive assortative matching in GBL groups. The Spearman rank correlation between the ability of any particular member and the median ability of the other members of their group is 0.48 (using either the estimates of ability from the logarithm of potato output or from the logarithm of potato acreage). About 62% of the groups had both a high and a low ability member (where high and low are defined relative to the median ability).

[51] Columns 1 and 4 of Table 9 present an index of ability based on observables, estimated from a random effects regression. When we use this ability index in the decomposition exercise, we find that the Selection Effect explains less than 15% of the overall difference in average TRAIL treatment effects on farm value-added. This suggests that unobserved characteristics constitute an important component of ability.

TRAIL agent.

In Table 14 we analyse input, output and credit transactions between sample households in TRAIL villages and the TRAIL agent. In each panel, the third row shows the mean incidence of such transactions for the Control 1 households. Note first that there is no evidence that recommended households interacted exclusively with the TRAIL agent in these markets. As can be seen in Panel A, over the 3 years, Control 1 households conducted only about 8% of their input transactions with the agent, accounting for less than 6% of the value of inputs purchased. Panel B shows that they conducted 19% of output transactions with the agent, representing less than 16% of the value of transactions, and Panel C shows that 16% of Control 1 households borrowed from the agent, accounting for 5% of their total borrowing.

It is difficult to infer whether the TRAIL agent engaged in corrupt behavior from differences in the agent's transactions between Control 1 and Control 2 households, since these households are likely to be different even in the absence of any corruption. Differences between Treatment and Control 1 households are more revealing because the only difference between them in expectation is that the Treatment households were randomly assigned to receive the TRAIL loans. In column 1 we run a regression of whether the household interacted with the TRAIL agent at all. In column 2 we regress the TRAIL agent's share in the total volume of (input or output) transactions the household carried out in that year. In columns 3–8 we run a regression of the form

$$\begin{aligned} y_{phvt} &= \lambda_0 + \lambda_1 \text{Interacted with agent}_{phv} + \lambda_2(\text{Interacted with agent}_{hv} \times \text{Treatment}_{hv}) \\ &+ \lambda_3(\text{Interacted with agent} \times \text{Control1}) + \gamma X_{hvt} + \epsilon_{hvt} \end{aligned} \tag{35}$$

where Interacted with agent indicates whether the household h in village v purchased the input p from or sold the crop p to the TRAIL agent of that village in the year t. The difference $\lambda_2 - \lambda_3$ captures whether Treatment households interacted with the TRAIL agent on different terms than Control 1 households did. In Panel A, the two significant treatment effects are a slight increase in the price at which farmers purchased seeds, and a reduction in the rate at which they rented power tillers from TRAIL agents. The two effects go in opposite directions. Given the large number of input prices we examine, we worry that the null hypothesis of no treatment effect could be rejected by chance even if it were actually true. To correct for this, we follow Kling, Liebman, and Katz (2007) and regress an index variable of all input prices in column 3. The Hochberg (1988) p-value of 0.773 indicates that overall, borrowers assigned to receive a TRAIL loan continued to pay the same input prices to TRAIL agents, as households who were recommended but not offered a loan.

In Panel B we find no significant effects on the quantities of output that borrowers sold to TRAIL agents, or the prices at which they sold them. Column 1 in Panel C shows that instead of borrowing more at higher interest rates, treatment borrowers were less likely to borrow from the agent during the three years of the experiment. The average interest rate charged by the agent also did not change.

Thus, our evidence does not indicate that the agent extracted side-payments from treated borrowers by engaging in a larger volume of transactions, charging higher prices for inputs sold or paying lower prices for outputs purchased from the borrowers, using Control 1 households as the benchmark. It appears likely that the TRAIL treatment households retained control over the program benefits that accrued to them. These results also cast doubt on the hypothesis that the agent gave extra concessions or useful advice about output sales or input purchases to TRAIL borrowers, than he gave to others whom he recommended but who did not receive TRAIL loans.

4.4.2 Year-Specific Effects

The TRAIL and GBL treatment effects on agricultural output estimated in Table 6 were averages across the three years of the intervention. It is informative to examine how these effects varied across years. As Figure 3 shows, the point estimates of the TRAIL treatment effects on potato acreage and the rate of return from potato cultivation were positive and higher than the corresponding GBL effects in each year of the intervention.

As the left panel of Figure 3 shows, the TRAIL treatment effects on potato acreage are statistically significant in each of the three years. They are also similar over the three years of the experiment. The point estimate for the corresponding effect in GBL increased monotonically over the years, although it was not statistically significant for the first two years and was borderline significant only in the third year. However statistically there is no evidence to suggest that the GBL treatment effects on potato acreage increased over the three years of the experiment. The differences in the GBL treatment effects were 0.023 acres in Year 2 versus Year 1 (p-value = 0.381), 0.006 acres in Year 3 versus Year 2 (p-value = 0.796) and 0.029 acres in Year 3 versus Year 1 (p-value = 0.331).

The right panel of Figure 3 shows that for TRAIL borrowers, the estimate of rates of return on cultivation (aggregated over the 4 major crop categories) increased from year 1 to year 2 and then remains roughly similar in year 3. The increase from year 1 to year 2 was statistically significant, although the difference between year 1 and year 3 was not. For GBL borrowers, we see an increase in the estimated rate of return from a negative value in year 1 to a positive value in year 2 and then a decline to a negative value in year 3 again. However none of these changes are significantly different from zero.[52]

5 Conclusion

The problem of identifying creditworthy borrowers and ensuring repayment in the absence of collateral has made agricultural finance in developing countries notoriously cost-ineffective for formal financial institutions. While microcredit has famously solved these problems by leveraging local information and local enforcement capability, recent interventions have shown that it generally does not increase borrowers' incomes or production.

The trader-agent intermediated lending (TRAIL) scheme attempts to overcome this problem by delegating the selection of borrowers for individual liability loans to local lenders or traders who are experienced at doing business with farmers in the local community. We compared the outcomes of this scheme with the outcomes of a traditional group loan treatment. Borrowers in the TRAIL scheme were particularly successful at increasing potato cultivation and output.

[52]The statistical significance of all comparisons of rates of return is inferred from 90% confidence intervals of the differences of 2000 cluster-bootstrapped estimates, constructed using Hall's percentile method.

Their farm incomes increased significantly, without any offsetting decline in income from any other source.

Consistent with other recent microcredit experiments, GBL borrowers' outcomes did not change appreciably. This was despite the fact that both the TRAIL and GBL loans were provided at below-market-average interest rates, had repayment durations that matched local crop cycles, and included insurance against local yield and price shocks. This makes it unlikely that these loan features were primarily responsible for the success of the TRAIL scheme. Instead, we argue that the TRAIL scheme induced agents to select higher ability borrowers than the borrowers who self-selected into the GBL scheme. To test this prediction of our model, we impose a Cobb-Douglas functional form on the farmers' production function and use this to estimate ability for each farmer in the sample, in the spirit of Olley and Pakes (1996) and Levinsohn and Petrin (2003). We then verify that more able households paid lower interest rates on the informal market, that recommended borrowers in the TRAIL scheme were more able than self-selected borrowers in the GBL scheme, and that treatment effects of the TRAIL loans were larger for more able households. We find also that selection differences can explain about one-third of the estimated difference in average treatment effects of the two schemes on farm value-added. The remainder of the difference may be caused by differences in borrower incentives: GBL loans could have had a smaller impact on borrowing, cultivation and farm incomes because the joint liability tax raised the effective interest rate. Other factors such as differential scope for learning or social capital may also have played a role. These need to be investigated further in future research.

Loan take-up rates were higher in the TRAIL scheme, suggesting that the scheme had larger *ex ante* effects on the welfare of the borrowers who were offered these loans. Consistent with our theoretical predictions, TRAIL loans were repaid at the same high rate as GBL loans. At the same time, the costs of administering the TRAIL scheme were lower. We found no evidence that TRAIL agents siphoned off the benefits that accrued to borrowers.

These results suggest that further experimentation with alternative versions of TRAIL may be fruitful. Future research also needs to examine a number of issues that have not been addressed this paper. These include the external validity of these results in other regions and contexts, the effects of scaling the scheme up to more borrowers per village, financial sustainability, distributive impacts on farm incomes, impacts on the empowerment of women or other disadvantaged social groups, and impacts on household consumption and liquidity management.

Our paper contributes to the policy debate on ways to promote financial inclusion of the rural poor in the developing world. Several countries have attempted to expand financial services in rural areas by employing local agents, but with limited success.[53] The TRAIL scheme is also related to a lending approach that India's central bank has been promoting recently, where "banking facilitators" are recruited from within the local communities to select and monitor borrowers on behalf of formal banks (Srinivasan, 2008). To our knowledge no rigorous evaluation of that approach has been carried out so far. The findings from our study could inform policy makers and central bank officials involved in the design of such schemes.

[53] Agents have been employed to intermediate financial services in Thailand (Onchan, 1992), Philippines (Floro and Ray, 1997), Bangladesh (Maloney and Ahmad, 1988), Malaysia (Wells, 1978), Indonesia (Fuentes, 1996) and Senegal (Warning and Sadoulet, 1998).

References

Attanasio, O., B. Augsburg, R. De Haas, E. Fitzsimons, and H. Harmgart (2015): "The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia," *American Economic Journal: Applied Economics*, 7(1), 90 – 122.

Banerjee, A. V., D. Karlan, and J. Zinman (2015): "Six Randomized Evaluations of Microcredit: Introduction and Further Steps," *American Economic Journal: Applied Economics*, 7(1), 1 – 21.

Beaman, L., D. Karlan, B. Thuysbaert, and C. Udry (2015): "Selection into Credit Markets: Evidence from Agriculture in Mali," Discussion paper, Yale University.

Besley, T., and S. Coate (1995): "Group lending, repayment incentives and social collateral," *Journal of Development Economics*, 46(1), 1 – 18.

Feigenberg, B., E. Field, and R. Pande (2013): "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance," *Review of Economic Studies*, 80(4), 1459 – 1483.

Field, E., R. Pande, J. Papp, and N. Rigol (2013): "Does the Classic Microfinance Model Discourage Entrepreneurship Among the Poor? Experimental Evidence from India," *American Economic Review*, 103(6), 2196 – 2226.

Fischer, G. (2013): "Contract Structure, Risk Sharing and Investment Choice," *Econometrica*, 81(3), 883 – 939.

Floro, M. S., and D. Ray (1997): "Vertical Links Between Formal and Informal Financial Institutions," *Review of Development Economics*, 1(1), 34 – 56.

Fuentes, G. (1996): "The Use of Village Agents in Rural Credit Delivery," *Journal of Development Studies*, 33(2), 188 – 209.

Ghatak, M. (1999): "Group lending, local information and peer selection," *Journal of Development Economics*, 60, 27 – 50.

——— (2000): "Screening by the company you keep: joint liability lending and the peer selection effect," *Economic Journal*, 110(465), 601 – 631.

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

Hochberg, Y. (1988): "A sharper Bonferroni procedure for multiple tests of significance," *Biometrika*, 75(4), 800 – 802.

Kaboski, J. P., and R. M. Townsend (2011): "A structural evaluation of a large-scale quasi-experimental microfinance initiative," *Econometrica*, 79(5), 1357–1401.

Karlan, D., and J. Zinman (2011): "Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation," *Science*, 332(6035), 1278 – 1284.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75, 83 – 119.

Levinsohn, J., and A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317 – 341.

Maloney, C., and A. B. Ahmad (1988): *Rural Savings and Credit in Bangladesh*. University Press Ltd., Dhaka Bangladesh.

Mitra, S., D. Mookherjee, M. Torero, and S. Visaria (2016): "Asymmetric Information and Middleman Margins: An Experiment with Indian Potato Farmers," Discussion paper, Mimeo, Hong Kong University of Science and Technology.

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263 – 1297.

Onchan, T. (1992): "Informal Rural Finance in Thailand," in *Informal Finance in Low-Income Countries*, ed. by D. W. Adams, and D. Fitchett. Westview Press, Boulder, CO.

Srinivasan, N. (2008): *Microfinance India. State of the Sector Report.* SAGE.

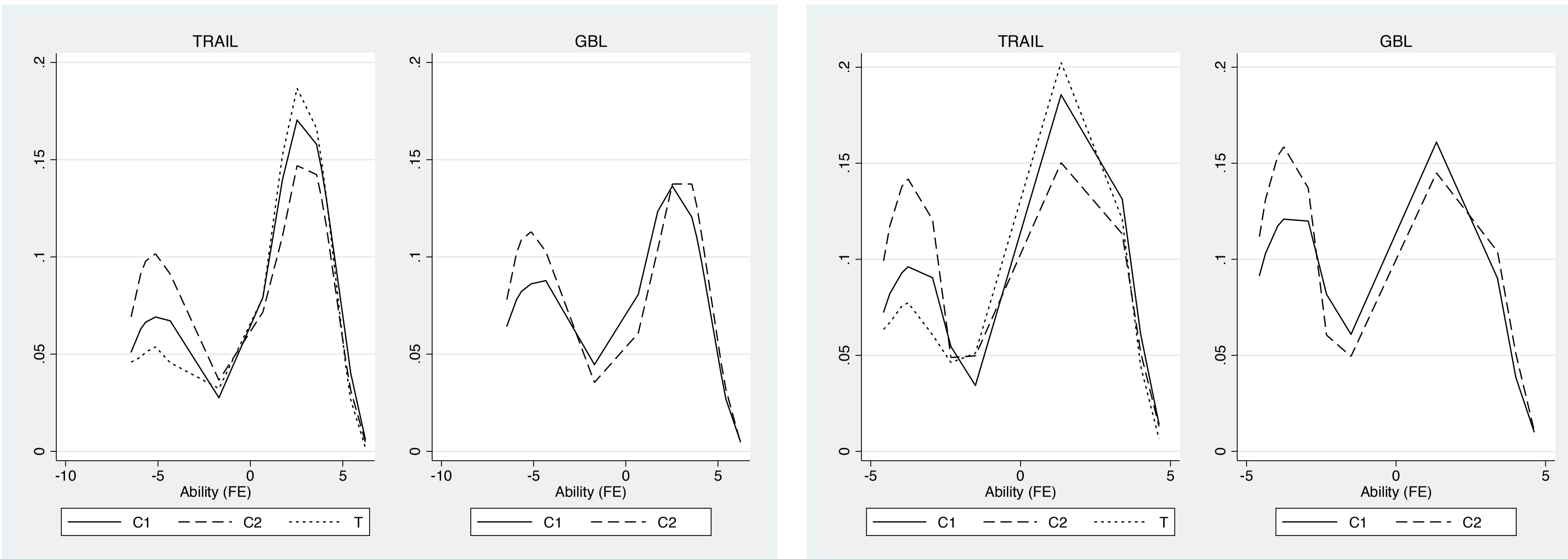
Warning, M., and E. Sadoulet (1998): "The Performance of Village Intermediaries in Rural Credit Delivery under Changing Penalty Regimes: Evidence from Senegal," *Journal of Development Studies*, 35(1), 115 – 138.

Wells, R. J. G. (1978): "An Input Credit Programme for Small Farmers in West Malaysia," *Journal of Administration Overseas*, 17, 4 – 16.

Figure 1: Density Functions of Estimated Ability for Treatment, Control 1 and Control 2 households

First stage: Log(Quantity of potatoes produced)

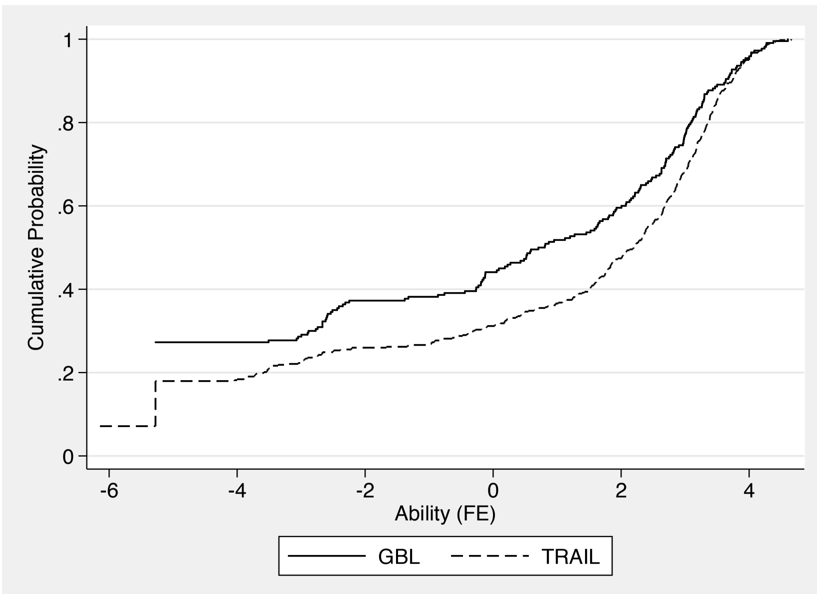
First stage: Log(Acreage under potato cultivation)



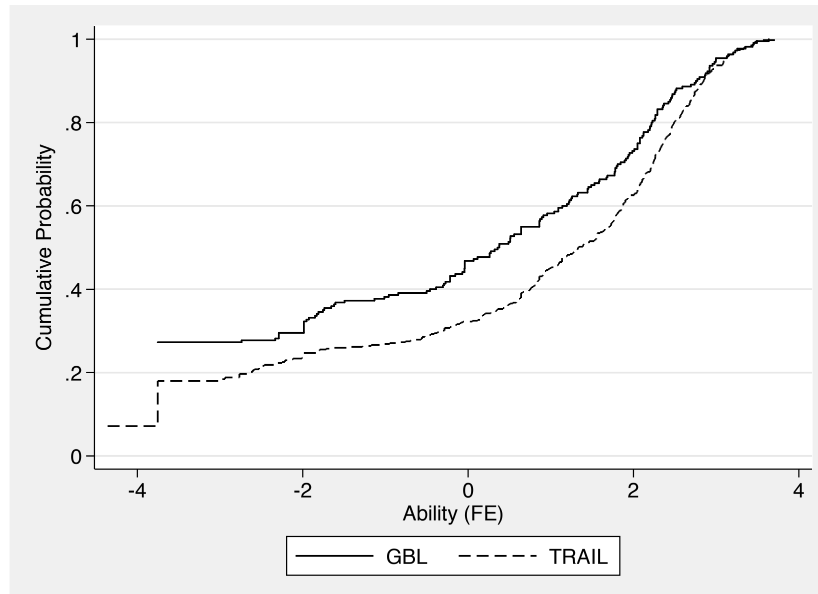
Notes:
Ability estimates for each household in TRAIL villages and Control 1 and Control 2 households in GBL villages are constructed from the household fixed effects from regressions shown in Table 9 columns 3 and 6. The ability index is not estimated for GBL treated households because their effective cost of borrowing depends on the pattern of assortative matching within groups, so that the formula for estimating ability does not have a closed-form solution. See the discussion on page 19 of the text.

Figure 2: Cumulative Distribution Functions of Estimated Ability for Selected Households in TRAIL and GBL villages

First stage: Log(Quantity of potatoes produced)



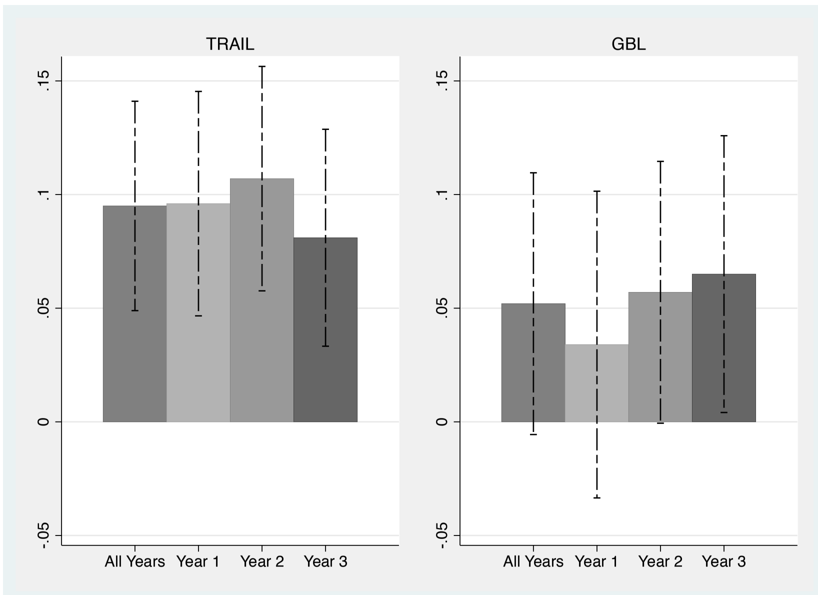
First stage: Log(Acreage under potato cultivation)



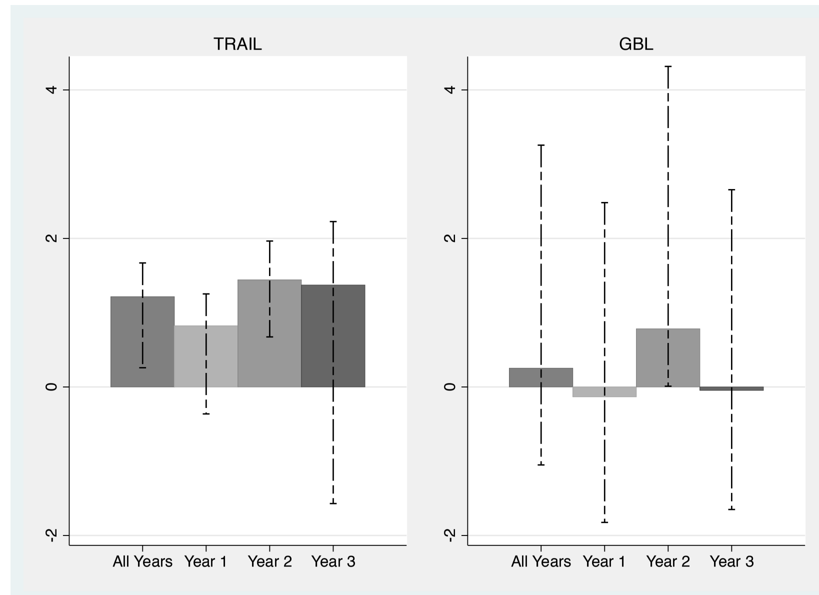
Notes:
Ability estimates for each household in TRAIL villages and Control 1 and Control 2 households in GBL villages are constructed from the household fixed effects from regressions shown in Table 9 columns 3 and 6. The ability index is not estimated for GBL treated households because their effective cost of borrowing depends on the pattern of assortative matching within groups, so that the formula for estimating ability does not have a closed-form solution. See the discussion on page 19 of the text.

Figure 3: Year-Specific Effects on Potato Acreage and Rate of Return

Potato Acreage



Rate of Return



Notes:
The values represent the estimated treatment effects from regressions following equation (30) in the text. In the left panel, the vertical axis measures the treatment effect on acres devoted to potato cultivation. In the right panel the vertical axis measures the rate of return on value-added across all four crop categories, computed as the ratio of the treatment effect on value-added to the treatment effect on the cost of cultivation. The dashed lines show the 90% confidence intervals. In the right panel, standard errors are cluster bootstrapped with 2000 replications, and the confidence intervals are constructed according to Hall's percentile method.

Table 1: Terms of TRAIL and GBL loans relative to other recent interventions

	TRAIL (1)	GBL (2)	Summary: Six evaluations [Table 1]BanerjeeKarlanZinman2015 (3)
Liability	Individual	Group (Joint)	Group (4), Individual (1), Both (1)
Interest Rate	18% APR	18% APR	12–27% APR (Mexico = 110% APR)
Market Interest Rate	24% APR	24% APR	16–47% APR (Mexico = 145% APR)
Loan Length	4 months	4 months	3–18 months
Repayment Frequency	4 months	4 months	Weekly/Bi-monthly/Monthly
Group Size	–	5	3–50
Collateralized	No	No	Yes (3), No (3)
Dynamic Incentives	Yes	Yes	Yes

Notes:
Columns 1 and 2 summarize the contractual terms of the TRAIL and GBL loans; Column 3 is adapted from Table 1 in Banerjee, Karlan, and Zinman (2015) and summarizes the contractual terms of the microloans offered in the six evaluations reported in the AEJ Symposium (2015).

Table 2: Balance of Characteristics across Treatment Categories

Panel A: Village Characteristics

	TRAIL (1)	GBL (2)	Difference (3)
Number of households	297.59	388.50	-90.91
	(48.06)	(80.36)	
Percent households electrified	0.60	0.59	0.01
	(0.06)	(0.05)	
Has primary school	0.77	0.79	-0.02
	(0.09)	(0.08)	
Has primary health centre	0.27	0.21	0.06
	(0.10)	(0.08)	
Has bank branch	0.14	0.17	-0.03
	(0.07)	(0.08)	
Has pucca road	0.27	0.42	-0.14
	(0.10)	(0.10)	

Panel B: Household Characteristics

	All		TRAIL			GBL	
		Treatment	Control 1	Difference	Treatment	Control 1	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	
Non Hindu	0.182	0.163	0.171	-0.008	0.131	0.118	0.013
	(0.008)	(0.025)	(0.025)		(0.022)	(0.022)	
Low caste	0.404	0.374	0.385	-0.010	0.520	0.459	0.061
	(0.011)	(0.032)	(0.032)		(0.033)	(0.034)	
High caste	0.414	0.463	0.444	0.018	0.349	0.423	-0.073
	(0.011)	(0.033)	(0.033)		(0.032)	(0.033)	
Landholding	0.464	0.448	0.454	-0.006	0.354	0.395	-0.040
	(0.009)	(0.027)	(0.025)		(0.025)	(0.026)	
Male headed household	0.941	0.987	0.991	-0.005	0.930	0.895	0.035
	(0.005)	(0.008)	(0.006)		(0.017)	(0.021)	
Head Education: Primary Schooling	0.420	0.520	0.487	0.033	0.432	0.427	0.005
	(0.011)	(0.033)	(0.033)		(0.033)	(0.033)	
Joint Significance ($\chi^2(5)$)			1.04			5.41	

Notes:
Panel A uses village census data collected in 2007 by Mitra, Mookherjee, Torero, and Visaria (2016). Panel B uses household survey data from the current study and restricts the sample to households with at most 1.5 acres of land. Column 1 includes Treatment, Control 1 and Control 2 households. Columns 2 and 5 include only Treatment households. Columns 3 and 6 include only Control 1 households. Standard errors are in parentheses. ‡: $\chi^2(5)$. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table 3: Credit Market Characteristics Before Experiment

	All Loans (1)		Agricultural Loans (2)	
Household had borrowed	0.67		0.59	
Total Borrowing†	6352	(10421)	5054	(8776)
Proportion of Loans by Source‡				
Traders/Money Lenders	0.63		0.66	
Family and Friends	0.05		0.02	
Cooperatives	0.24		0.25	
Government Banks	0.05		0.05	
Annualized Interest Rate by Source (percent)				
Traders/Money Lenders	24.93	(20.36)	25.19	(21.47)
Family and Friends	21.28	(14.12)	22.66	(16.50)
Cooperatives	15.51	(3.83)	15.70	(2.97)
Government Banks	11.33	(4.63)	11.87	(4.57)
Duration by Source (days)				
Traders/Money Lenders	125.08	(34.05)	122.80	(22.43)
Family and Friends	164.08	(97.40)	183.70	(104.25)
Cooperatives	323.34	(90.97)	327.25	(87.74)
Government Banks	271.86	(121.04)	324.67	(91.49)
Proportion of Loans Collateralized by Source				
Traders/Money Lenders	0.02		0.01	
Family and Friends	0.04		0.07	
Cooperatives	0.79		0.78	
Government Banks	0.81		0.83	

Notes:
Statistics are reported for all sample households in TRAIL and GBL villages with at most 1.5 acres of land. All characteristics are for loans taken by the households in Cycle 1. Program loans are not included. For the interest rate summary statistics loans where the principal amount is reported equal to the repayment amount are not included. To arrive at representative estimates for the study area, Treatment and Control 1 households are assigned a weight of $\frac{30}{N}$ and Control 2 households are assigned a weight of $\frac{N-30}{N}$, were N is the total number of households in their village. †: Total borrowing = 0 for households that do not borrow. ‡: Proportion of loans in terms of value of loans at the household level. All proportions are computed only over households that borrowed. Standard deviations are in parentheses.

Table 4: Selected Crop Characteristics

	Sesame (1)	Paddy (2)	Potatoes (3)
Cultivate the crop (%)	0.46	0.67	0.62
	(0.006)	(0.006)	(0.006)
Acreage (acres)	0.21	0.47	0.31
	(0.004)	(0.007)	(0.005)
Harvested quantity (kg)	145.5	1191.26	5387.76
	(2.70)	(17.05)	(79.74)
Cost of production (Rs)	335.05	2985.55	7556.46
	(8.15)	(53.52)	(142.30)
Price (Rs/kg)	30.7	10.3	4.7
	(0.169)	(0.097)	(0.027)
Revenue (Rs)	1636.38	5561.95	13600.5
	(38.37)	(102.77)	(256.34)
Value added (Rs)	1300.47	2636.47	5986.28
	(33.73)	(69.93)	(151.43)
Value added per acre (Rs/acre)	6530.38	6596.34	18139.33
	(82.31)	(109.82)	(296.79)

Notes:
Statistics are annual averages over the 3-year study period, reported for all sample households in TRAIL and GBL villages with at most 1.5 acres of land. To arrive at representative estimates for the study area, Treatment and Control 1 households are assigned a weight of $\frac{30}{N}$ and Control 2 households are assigned a weight of $\frac{N-30}{N}$, were N is the total number of households in the village. Standard errors are in parentheses.

Table 5: Program Impacts: Treatment Effects on Total Borrowing

	All Agricultural Loans (₹) (1)	Non Program Agricultural Loans† (₹) (2)	Index of dependent variables[II] (3)
TRAIL	5050***	-364.60	0.242***
	(767.50)	(630.90)	(0.082)
Hochberg p-value			0.025
Mean TRAIL Control 1	5590	5590	
% Effect TRAIL	90.34	-6.52	
GBL	3732***	-157.8	0.176*
	(845.10)	(600.2)	(0.092)
Hochberg p-value			0.368
Mean GBL Control 1	4077	4077	
% Effect GBL	91.54	-3.87	
Sample size	6,204	6,204	

Notes:
Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. Regressions also control for the gender and educational attainment, caste and religion of the household head, household's landholding, a set of year dummies and an information village dummy. % Effect: Treatment effect as a percentage of the mean of the relevant Control 1 group. [II]: In column 3 the dependent variable is an index of z-scores of the outcome variables in the panel; the p-values for treatment effects in this column are computed according to Hochberg (1988)'s step-up method to control for the family-weighted error rate across all index outcomes. †: Non-Program loans refer to loans from sources other than the TRAIL/GBL schemes. The complete regression results are in Table A-5. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table 6: Program Impacts: Treatment Effects in Agriculture

Panel A: Potatoes

	Cultivate (%) (1)	Land planted (Acres) (2)	Harvested quantity (Kg) (3)	Cost of Production (₹) (4)	Revenue (₹) (5)	Value Added (₹) (6)	Imputed Profit‡ (₹) (7)	Index of dependent variables[II] (8)
TRAIL Treatment	0.047	0.095***	975.371	1909.738***	4011.624***	2109.242***	1939.494***	0.198***
	(0.032)	(0.028)	(301.124)	(718.799)	(1186.538)	(621.037)	(591.339)	(0.057)
Hochberg p-value								0.003
Mean TRAIL Control 1	0.715	0.333	3646.124	8474.628	14285.467	5739.479	4740.893	
% Effect TRAIL	6.56	28.46	26.75	22.53	28.08	36.75	40.91	
GBL Treatment	0.053	0.052	514.435	1601.298*	2343.964	714.137	553.708	0.111
	(0.044)	(0.035)	(395.082)	(877.219)	(1729.723)	(918.671)	(866.430)	(0.081)
Hochberg p-value								0.861
Mean GBL Control 1	0.620	0.251	2761.127	5992.080	11014.286	4997.446	4018.796	
% Effect GBL	8.59	20.79	18.63	26.72	21.28	14.29	13.78	
Sample Size	6210	6210	6210	6210	6210	6210	6210	

Continued ...

Notes:
Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. Regressions also control for the gender and educational attainment, caste and religion of the household head, household's landholding, a set of year dummies and an information village dummy. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. II: In column 8 the dependent variable is an index of z-scores of the outcome variables in the panel; the p-values for treatment effects in this column are computed according to Hochberg (1988)'s step-up method to control for the family-weighted error rate across all index outcomes. The complete regression results are in Table A-6. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table 6 (*Continued*): Program Impacts: Treatment Effects in Agriculture

Panel B: Other Major Crops

		Sesame			Paddy			Vegetables	
	Land planted (Acres) (1)	Value Added (₹) (2)	Index of dependent variables^{II} (3)	Land planted (Acres) (4)	Value Added (₹) (5)	Index of dependent variables^{II} (6)	Land planted (Acres) (7)	Value Added (₹) (8)	Index of dependent variables^{II} (9)
TRAIL Treatment	0.044*	278.223*	0.096	0.036*	267.790	0.045	0.011	51.952	0.044
	(0.023)	(142.192)	(0.058)	(0.020)	(241.457)	(0.030)	(0.007)	(321.736)	(0.080)
Hochberg p-value			0.302			0.269			0.580
Mean TRAIL Control 1	0.266	1519.558		0.470	2556.755		0.015	889.229	
% Effect TRAIL	16.39	18.31		7.66	10.47		72.13	5.84	
GBL Treatment	0.003	-204.084	-0.041	0.011	213.527	-0.004	0.000	-323.404	-0.031
	(0.031)	(229.475)	(0.084)	(0.029)	(271.907)	(0.053)	(0.009)	(676.455)	(0.150)
Hochberg p-value			>0.999			0.943			>0.999
Mean GBL Control 1	0.193	1252.850		0.456	2336.837		0.022	1142.350	
% Effect GBL	1.46	-16.29		2.39	9.14		0.80	-28.31	
Sample Size	6210	6210		6210	6210		6210	6210	

Notes:
Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. Regressions also control for the gender and educational attainment, caste and religion of the household head, household's landholding, a set of year dummies and an information village dummy. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. ^{II}: In columns 3, 6 & 9, the dependent variables are indices of z-scores of the outcome variables related to that crop; the p-values for treatment effects in these columns are computed according to Hochberg (1988)'s step-up method to control for the family-weighted error rate across all index outcomes. The complete regression results corresponding to columns 1–2 are in Table A-7, to columns 4–5 are in Table A-8, and to columns 7–8 are in Table A-9. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Appendix A: Proofs

Proof of Lemma 2:

(i) Suppose l_i^T is nondecreasing in i. Take any pair of types satisfying $i > j$. Then $l_i^T \geq l_j^T$. Applying condition RC we obtain $\frac{f'(\bar{l}_i+l)}{f'(\bar{l}_j+l)}$ is increasing in l. Combining this with (1) which implies $\frac{f'(\bar{l}_i)}{f'(\bar{l}_j)} = \frac{\bar{\theta}_j}{\bar{\theta}_i}$, we therefore obtain $\bar{\theta}_i f'(\bar{l}_i + l) > \bar{\theta}_j f'(\bar{l}_j + l)$ for any $l > 0$.

Hence the output treatment effect for type i is

$$\begin{aligned}
\bar{\theta}_i[f(\bar{l}_i + l_i^T) - f(\bar{l}_i)] &\geq \bar{\theta}_i[f(\bar{l}_i + l_j^T) - f(\bar{l}_i)] \\
&= \bar{\theta}_i \int_0^{l_j^T} f'(\bar{l}_i + l)dl \\
&> \bar{\theta}_j \int_0^{l_j^T} f'(\bar{l}_j + l)dl \\
&= \bar{\theta}_j[f(\bar{l}_j + l_j^T) - f(\bar{l}_j)]
\end{aligned}$$

the output treatment effect for j.

(b) If the probability of crop success p_i is independent of i, equal to p, equations (1, 2) imply $\frac{f'(\bar{l}_i+l_i^T)}{f'(\bar{l}_i)}$ is independent of i. This implies $\log f'(\bar{l}_i + l_i^T) - \log f'(\bar{l}_i)$ is a constant. Differentiating this expression and setting equal to zero, we obtain

$$1 + \frac{l_i^{T'}}{\bar{l}_i'} = [-\frac{f''(\bar{l}_i)}{f'(\bar{l}_i)}][-\frac{f''(\bar{l}_i + l_i^T)}{f'(\bar{l}_i + l_i^T)}]^{-1} > 1 \tag{36}$$

owing to RC. Hence the loan treatment effect is rising in i. Applying part (a) we infer the output treatment effect is rising in i.

Finally consider the income treatment effect for i:

$$\begin{aligned}
\bar{\theta}_i[f(\bar{l}_i + l_i^T) - f(\bar{l}_i)] - p_i r_T l_i^T &= [\bar{\theta}_i f(\bar{l}_i + l_i^T) - p_i r_T l_i^T] - \bar{\theta}_i f(\bar{l}_i) \\
&\geq [\bar{\theta}_i f(\bar{l}_i + l_j^T) - p_i r_T l_j^T] - \bar{\theta}_i f(\bar{l}_i) \\
&= \bar{\theta}_i[f(\bar{l}_i + l_j^T) - f(\bar{l}_i)] - p_i r_T l_j^T \\
&= \bar{\theta}_i \int_0^{l_j^T} f'(\bar{l}_i + l)dl - p_i r_T l_j^T \\
&> \bar{\theta}_j \int_0^{l_j^T} f'(\bar{l}_j + l)dl - p_i r_T l_j^T
\end{aligned}$$

where the first inequality uses the property that type i chooses TRAIL loan size to maximize his post-intervention income, and has the option of choosing the loan size selected by type j. The last inequality again uses RC in the way described in the proof of (a) above. The expression in the last line above equals the income treatment effect for type j, less $(p_i - p_j)r^T l_j^T$. This last 'correction' term equals zero (approximately) when p_i does not vary (varies very little) with i.

(c) If productivity does not vary with i, the pre-intervention loan size and output are rising in i, but after the intervention do not vary with i. Hence the loan and output treatment effects are falling in i. QED

Table 7: Program Impacts: Effects on Farm Value Added and Rates of Return

	Farm Value Added (₹) (1)	Non-Agricultural Income (₹) (2)	Index of dependent variables[II] (3)	Rate of Return: Potato Cultivation (4)	Rate of Return: Farm Value Added (5)
TRAIL Treatment	2707.21***	-608.000	0.116***	1.10†	1.22†
	(739.46)	(4153.557)	(0.044)	(0.02)	(0.05)
Hochberg p-value			0.034		
Mean TRAIL Control 1	10705.02	40115.81			
% Effect TRAIL	25.29	-1.52			
GBL Treatment	400.176	-6092.631	-0.017	0.45	0.25
	(1069.61)	(4959.88)	(0.047)	(1.10)	(3.04)
Hochberg p-value			>0.999		
Mean GBL Control 1	9729.48	45645.10			
% Effect GBL	4.11	-13.35			
TRAIL vs GBL p-value	0.077	0.393			
TRAIL vs GBL (90% CI)				[-1.410, 1.418]	[-2.484, 2.407]
Sample Size	6,210	6,210			

Notes:
Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. Regressions also control for the gender and educational attainment, caste and religion of the household head, household's landholding, a set of year dummies and an information village dummy. The full set of results corresponding to columns 1 and 2 are in Table A-10. The rate of return is the ratio of the treatment effect on value-added to the treatment effect on cost. [II]: In column 3 the dependent variable is an index of z-scores of the outcome variables in the panel following Kling, Liebman, and Katz (2007); p-values for this regression are reported using Hochberg (1988)'s step-up method to control the FWER across all index outcomes. In columns 1 and 2, the standard errors in parentheses are clustered at the hamlet level. In columns 4 and 5, the numbers in parentheses are the averages of cluster bootstrapped standard errors with 2000 replications. † indicates that the 90 percent confidence interval of bootstrapped estimates constructed according to Hall's percentile method does not include zero. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table 8: Loan Performance

	Repayment (1)	Take up (2)	Continuation (3)
Panel A: Sample Means			
TRAIL	0.958	0.856	0.805
	0.005	(0.008)	(0.009)
GBL	0.954	0.746	0.691
	(0.006)	(0.011)	(0.011)
Difference	0.003	0.110***	0.114***
	(0.008)	(0.014)	(0.014)
Panel B: Regression Results			
TRAIL	0.006	0.117*	0.116*
	(0.020)	(0.067)	(0.067)
Constant	0.989***	0.838***	0.827***
	(0.015)	(0.053)	(0.053)
Mean GBL	0.956	0.747	0.694
Sample Size	2,406	3,226	3,512

Notes:
The sample consists of household-cycle level observations of Treatment households in TRAIL and GBL villages. The dependent variable in column 1 takes value 1 if a borrowing household repaid a loan taken in the cycle within 30 days of the due date, and that in columns 2 and 3 takes value 1 if the household took the program loan. In column 1 the sample consists of households that had taken a program loan in that cycle, in column 2 it consists of households that were eligible to take the program loan in that cycle, and in column 3 it consists of all households that were eligible to receive program loans in Cycle 1. In Panel B, treatment effects are computed from regressions that follow equation 31 in the text. Standard errors in parentheses are clustered at the hamlet level. †: Difference between mean in TRAIL and mean in GBL. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 9: Estimating Ability. First Stage Regressions

Dependent variable:	**Log(Quantity of potatoes produced)**			**Log(Acreage under potato cultivation)**		
	Random Effects (2)	OLS (3)	Fixed Effects (1)	Random Effects (5)	OLS (6)	Fixed Effects (4)
Landholding	2.003***	2.365***		1.588***	1.835***	
	(0.251)	(0.235)		(0.182)	(0.169)	
Non Hindu household	-1.818***	-0.443		-1.268***	-0.294	
	(0.326)	(0.360)		(0.238)	(0.257)	
Low caste household	-0.510**	-0.422**		-0.342**	-0.323**	
	(0.227)	(0.213)		(0.166)	(0.154)	
Male headed household	0.175	0.192		0.122	0.143	
	(0.367)	(0.347)		(0.260)	(0.245)	
Education of head: at least primary	-0.055	0.002		-0.022	0.010	
	(0.159)	(0.144)		(0.115)	(0.103)	
Occupation of head: cultivation	2.699***	2.443***		1.932***	1.749***	
	(0.203)	(0.175)		(0.147)	(0.126)	
Household size	0.175***	0.195***		0.124***	0.141***	
	(0.040)	(0.039)		(0.029)	(0.028)	
Year 2	-0.456***	-0.456***	-0.453***	-0.289***	-0.289***	-0.283***
	(0.073)	(0.073)	(0.065)	(0.052)	(0.052)	(0.046)
Year 3	-0.374***	-0.374***	-0.381***	-0.248***	-0.248***	-0.250***
	(0.076)	(0.076)	(0.070)	(0.053)	(0.054)	(0.049)
Constant	2.898***	3.637***	5.554***	-4.967***	-4.348***	-2.976***
	(0.417)	(0.783)	(0.040)	(0.297)	(0.575)	(0.028)
Sample size	4,833	4,833	6,204	4,833	4,833	6,204
Number of households	1,613		2,068	1,613		2,068
δ			0.951			0.965
			[0.873, 1.034]			[0.891, 1.045]
K			0.567			0.454
			[0.115, 1.035]			[0.118, 0.794]

Notes:
Standard errors in parentheses are clustered at the hamlet level. In columns 1 and 4, the estimating sample includes all sample households in TRAIL and GBL villages with at most 1.5 acres of land. In columns 2, 3, 5 and 6 the estimating sample include Control 1 and Control 2 households in TRAIL and GBL villages with at most 1.5 acres of land. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$. The numbers in square brackets denote the 90 percent confidence interval of the TRAIL–GBL difference in rate of return, computed using Hall's percentile method with 2000 replications. † indicates that the 90 percent confidence interval of bootstrapped estimates does not include zero.

Table 10: Ability Differences and Patterns of Selection into the TRAIL scheme

Ability estimates from:	**Log(Quantity of potatoes produced)** (1)	**Log(Acreage under potato cultivation)** (2)
Selected	0.390	0.278
	(0.286)	(0.205)
Own Clientele	0.947**	0.715**
	(0.463)	(0.340)
Selected × Own Clientele	0.174	0.071
	(0.547)	(0.394)
Constant	0.213	0.183
	(0.656)	(0.477)
Total Effects		
Selected:		
Own Clientele	1.723**	1.247**
	(0.704)	(0.513)
Other Clientele	0.602	0.461
	(0.655)	(0.478)
Not Selected:		
Own Clientele	1.159	0.898
	(0.784)	(0.572)
Other Clientele	0.213	0.183
	(0.656)	(0.477)
Difference Estimates		
Selected: Own v. Other Clientele	1.121***	0.786***
	(0.322)	(0.236)
Own Clientele: Selected v. Not Selected	0.564	0.349
	(0.488)	(0.353)
Sample Size	1,032	1,032

Notes:
Coefficients are reported from regressions that follow equation 32 in the text. The dependent variable is ability estimates constructed from household fixed effects, as reported in columns 1 and 4 of Table 9. The regressions also control for village fixed effects. The estimating sample includes all sample households in TRAIL villages with at most 1.5 acres of land. Standard errors in parentheses are clustered at the hamlet level. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 11: Relationship between ability and interest rate paid on informal loans

Ability estimates from:	**Log(Quantity of potatoes produced)** (1)	(2)	**Log(Acreage under potato cultivation)** (3)	(4)
$\hat{Q}_1$	0.348***	0.328***	0.348***	0.328***
	(0.076)	(0.077)	(0.076)	(0.078)
$\hat{Q}_2$	0.246***	0.234***	0.250***	0.239***
	(0.020)	(0.022)	(0.020)	(0.023)
$\hat{Q}_3$	0.237***	0.222***	0.218***	0.209***
	(0.021)	(0.020)	(0.019)	(0.019)
$\hat{Q}_4$	0.210***	0.211***	0.221***	0.218***
	(0.010)	(0.018)	(0.012)	(0.018)
Village Fixed Effects	No	Yes	No	Yes
Sample Size	661	661	661	661
Difference Estimates				
$\hat{Q}_1 - \hat{Q}_2$	0.103	0.0940	0.0988	0.0895
	(0.089)	(0.088)	(0.087)	(0.086)
$\hat{Q}_1 - \hat{Q}_3$	0.111	0.105†	0.130†	0.119†
	(0.086)	(0.084)	(0.087)	(0.086)
$\hat{Q}_1 - \hat{Q}_4$	0.139†	0.117†	0.127†	0.111
	(0.084)	(0.086)	(0.085)	(0.086)
$\hat{Q}_2 - \hat{Q}_3$	0.00878	0.0111	0.0313	0.0295
	(0.036)	(0.028)	(0.033)	(0.028)
$\hat{Q}_2 - \hat{Q}_4$	0.0361	0.0230	0.0286	0.0211
	(0.028)	(0.025)	(0.029)	(0.026)
$\hat{Q}_3 - \hat{Q}_4$	0.0273	0.0119	-0.00273	-0.00843
	(0.027)	(0.024)	(0.026)	(0.022)
Average interest rate paid by Control 1 households				
Both schemes:		0.242		
TRAIL		0.226		
GBL		0.260		

Notes:
The dependent variable is the average annualized interest rate paid on informal production loans from traders, moneylenders and family and friends, as reported in Cycle 1. The estimating sample includes all Control 1 and Control 2 households in TRAIL and GBL villages with at most 1.5 acres who had borrowed from traders, moneylenders and family and friends in Cycle 1. Loans where the principal amount is reported equal to the repayment amount are not included. Standard errors in parentheses are clustered at the hamlet level. Standard errors are averages of cluster bootstrap standard errors from 2000 replications. † indicates that the 90 percent confidence interval of bootstrap estimates constructed according to Hall's percentile method does not include zero.

Table 12: Heterogeneous Treatment Effects by Ability in the TRAIL Scheme

Ability estimates from:	**Log(Quantity of potatoes produced)**			**Log(Acreage under potato cultivation)**		
	Total Borrowing (1)	Farm Revenue (2)	Farm Value Added (3)	Total Borrowing (4)	Farm Revenue (5)	Farm Value Added (6)
Treatment Effects by Quartile						
$\hat{Q}_1$	656†	971	723	668	1030	755
	(610.69)	(2119.77)	(1352.14)	(612.53)	(2126.99)	(1358.77)
$\hat{Q}_2$	2832†	-2384	-2238	3222†	-763	-912
	(1562.77)	(3101.69)	(1993.64)	(1576.05)	(2705.65)	(1629.61)
$\hat{Q}_3$	6327†	8123†	4039†	5657†	5628†	2397†
	(1266.39)	(2029.75)	(1406.82)	(1364.59)	(2451.27)	(1820.57)
$\hat{Q}_4$	9474†	14022†	8487†	8614†	11404†	7198†
	(2728.61)	(4675.39)	(2938.84)	(2653.91)	(4798.55)	(2941.62)
Differences in Treatment Effects by Quartile						
$\hat{Q}_2 - \hat{Q}_1$	2176†	-3355	-2960	2554†	-1793	-1667
	(1710.89)	(3982.57)	(2546.51)	(1730.28)	(3697.93)	(2276.35)
$\hat{Q}_3 - \hat{Q}_1$	5672†	7152†	3317†	4989†	4598†	1642†
	(1421.79)	(2933.36)	(1970.81)	(1489.25)	(3257.88)	(2302.24)
$\hat{Q}_4 - \hat{Q}_1$	8819†	13051†	7764†	7946†	10374†	6444†
	(2853.91)	(5140.47)	(3193.03)	(2789.50)	(5272.96)	(3208.94)
$\hat{Q}_3 - \hat{Q}_2$	3495†	10507†	6277†	2435†	6391†	3309†
	(1913.40)	(3612.24)	(2424.46)	(1957.65)	(3649.24)	(2507.55)
$\hat{Q}_4 - \hat{Q}_2$	6642†	16407†	10725†	5392†	12167†	8110†
	(3217.19)	(5710.95)	(3773.64)	(3174.47)	(5533.51)	(3472.78)
$\hat{Q}_4 - \hat{Q}_3$	3147	5900	4448†	2957	5775	4801
	(2697.75)	(4997.28)	(3310.25)	(2726.12)	(5403.31)	(3561.57)
Sample size	3,093	3,093	3,093	3,093	3,093	3,093

Notes:
Treatment effects and differences in treatment effects are presented. Estimating sample includes all sample households in TRAIL villages with at most 1.5 acres of land. Regressions control for year dummies and information village dummy. Standard errors are averages of cluster bootstrap standard errors from 2000 replications. † indicates that the 90 percent confidence interval of cluster bootstrap estimates constructed according to Hall's percentile method does not include zero.

Table 13: Decomposition of Average Effect on Farm Value Added by Ability

	TRAIL (1)	GBL (2)	Difference (TRAIL - GBL) (3 = 1-2)	Treatment Effect (4)	Difference × Treatment Effect (5 = 3 × 4)
Ability estimates from: log(quantity of potatoes produced)					
$\hat{Q}_1$	0.18	0.27	-0.09	722.6	-67.25
$\hat{Q}_2$	0.24	0.28	-0.04	-2238	93.01
$\hat{Q}_3$	0.30	0.25	0.05	4039	187.96
$\hat{Q}_4$	0.28	0.20	0.09	8487	747.66
% of Average Effect due to Ability					35.91
Ability estimates from: log(acreage under potato cultivation)					
$\hat{Q}_1$	0.18	0.27	-0.09	754.8	-70.25
$\hat{Q}_2$	0.25	0.29	-0.04	-911.9	36.32
$\hat{Q}_3$	0.28	0.24	0.04	2397	97.02
$\hat{Q}_4$	0.29	0.20	0.09	7198	665.27
% of Average Effect due to Ability					27.21

Notes:
Columns 1 and 2 present the fraction of selected borrowers in TRAIL and GBL respectively who belonged to each estimated ability quartile, and column 3 presents the difference between the two. Column 4 presents the TRAIL treatment effects on farm value-added from Table 12. The last row in each panel shows the percentage of the average treatment effect difference between the TRAIL and GBL schemes that can be explained by the Selection Effect, as per equation 29 in the text.

Table 14: Treatment Effects for Transactions with TRAIL agent

Panel A: Input Purchase†

	Purchased from agent	Agent's share in purchases	Input Price (Rs/unit)					Index of input prices[II]
			Fertilizer	Outside Seed	Pesticide	Power tiller	Water	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TRAIL Treatment	0.002	0.000	0.136	2.099*	-32.41	-29.11***	109.80	0.023
	(0.014)	(0.011)	(0.929)	(1.131)	(48.30)	(4.854)	(109.80)	(0.026)
Hochberg p-value								0.773
Mean Control 1	0.0813	0.0620	15.77	24.82	536.8	211.2	72.30	
Sample Size	17,928	17,784	2,908	2,394	3,830	1,983	1,822	

Panel B: Output Sales†

	Sold to agent	Agent's share in sales	Output Price (Rs/kg)			Index of output prices[II]
			Potato	Paddy	Sesame	
	(1)	(2)	(3)	(4)	(5)	(6)
TRAIL Treatment	0.020	0.027	-0.024	0.401	0.010	0.001
	(0.029)	(0.028)	(0.141)	(0.285)	(0.516)	(0.005)
Hochberg p-value						0.846
Mean Control 1	0.192	0.152	4.566	10.13	30.59	
Sample Size	4,303	4,098	2,026	791	1,280	

Panel C: Borrowing†

	Borrowed from Agent	Agent's share in borrowing	APR
	(1)	(2)	(3)
TRAIL Treatment	-0.076**	-0.036***	0.011
	(0.038)	(0.012)	0.043
Mean Control 1	0.161	0.0489	0.139
Sample Size	1,960	1,960	5,468

Notes:
The regressions are run on household-year level data for sample households with at most 1.5 acres of land in TRAIL villages. In Panel C only borrowing for agricultural purposes is considered. †: Purchased inputs from, sold output to or borrowed from agent during the survey period. II: In column 8 in Panel A and Column 6 in Panel B the dependent variables are indices of z-scores of input prices and output price respectively, following KlingLiebmanKatz2007; p-values for this regression are reported using Hochberg (1988)'s step-up method to control the FWER across the two indices. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-1: Robustness of Results 1. Program Impacts: Treatment Effects on Total Borrowing

	Alternative Clustering		Restricting to Recommended Households		All Loans	
	All	Non-Program†	All	Non Program†	All	Non-Program†
	(1)	(2)	(3)	(4)	(5)	(6)
TRAIL	5049***	-365.50	5015***	-357.10	6829***	-285.90
	(989.50)	(581.40)	(759)	(625.10)	(1249)	(1121)
Mean Control 1	5590	5590	5590	5590	7523	7523
% Effect TRAIL	90.32	-6.54	89.71	-6.39	90.77	-3.80
GBL	3731***	-158.50)	3787***	-134.60	4556***	-952
	(1000)	(716.70)	(843.80)	(599.9)	(1040)	(853.80)
Mean Control 1	4077	4077	4077	4077	6005	6005
% Effect GBL	91.51	-3.89	92.89	-3.30	75.87	-15.85
Sample Size	6,210	6,210	2,733	2,733	6,204	6,204

Notes:
Treatment effects presented. †: Non-Program loans are loans from sources other than the TRAIL or GBL schemes. Columns 1, 2, 5 and 6 use data from all sample households with at most 1.5 acres of land in TRAIL and GBL villages. Columns 3 and 4 use data from Treatment and Control 1 households in the sample with at most 1.5 acres of land in TRAIL and GBL villages. Columns 1–4 restricts loans for agricultural purposes from moneylenders, traders and family and friends. Columns 5 and 6 uses all loans irrespective of stated purpose. In columns 3–6 standard errors in parentheses are clustered at the para (hamlet) level while those in columns 1 and 2 are clustered at the village level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-2: Robustness of Results 2. Program Impacts: Treatment Effects in Agriculture. Alternative Clustering

	Cultivate (%) (1)	Acreage (Acres) (2)	Harvested quantity (Kg) (3)	Cost of Production (₹) (4)	Revenue (₹) (5)	Value Added (₹) (6)	Imputed Profit‡ (₹) (7)
				Potato			
TRAIL Treatment	0.047*	0.095***	975.041***	1908.985***	4010.208***	2108.584***	1938.840***
	(0.025)	(0.024)	(249.701)	(632.620)	(1060.698)	(546.374)	(534.952)
Mean TRAIL	0.715	0.333	3646.124	8474.628	14285.467	5739.479	4740.893
% Effect TRAIL	6.564	28.452	26.742	22.526	28.072	36.738	40.896
GBL Treatment	0.053	0.052*	514.011	1600.336**	2342.149	713.290	552.870
	(0.048)	(0.030)	(331.594)	(735.408)	(1479.915)	(833.048)	(775.843)
Mean GBL	0.620	0.251	2761.127	5992.080	11014.286	4997.446	4018.796
% Effect GBL	8.586	20.779	18.616	26.708	21.265	14.273	13.757
Sample Size	6216	6216	6216	6216	6216	6216	6216
				Sesame			
TRAIL Treatment	0.035	0.044**	9.640	25.846	304.917*	278.171**	179.145
	(0.025)	(0.021)	(5.850)	(44.772)	(159.456)	(133.339)	(112.534)
Mean TRAIL	0.581	0.266	81.624	436.910	1957.498	1519.558	1080.800
% Effect TRAIL	6.024	16.383	11.810	5.916	15.577	18.306	16.575
GBL Treatment	-0.024	0.003	-5.452	16.776	-188.692	-204.157	-129.467
	(0.043)	(0.030)	(8.156)	(32.762)	(218.261)	(203.706)	(197.176)
Mean GBL	0.484	0.193	60.848	258.878	1513.138	1252.850	866.288
% Effect GBL	-4.959	1.453	-8.960	6.480	-12.470	-16.295	-14.945
Sample Size	6,216	6,216	6,216	6,216	6,216	6,216	6,216
				Paddy			
TRAIL Treatment	-0.005	0.036**	22.214	212.371	471.201**	267.780	135.235
	(0.025)	(0.014)	(21.934)	(142.091)	(177.856)	(189.045)	(142.456)
Mean TRAIL	0.744	0.470	569.726	2889.838	5398.490	2556.755	93.133
% Effect TRAIL	-0.703	7.667	3.899	7.349	8.728	10.473	145.207
GBL Treatment	0.000	0.011	-36.517	-74.850	114.611	213.507	-120.627
	(0.029)	(0.022)	(57.881)	(281.160)	(437.564)	(277.301)	(194.767)
Mean GBL	0.689	0.456	672.894	3225.745	5513.227	2336.837	183.163
% Effect GBL	-0.069	2.403	-5.427	-2.320	2.079	9.137	-65.858
Sample Size	6216	6216	6216	6216	6216	6216	6216
				Vegetables			
TRAIL Treatment	0.000	0.011	27.623	81.354	137.182	51.969	-10.670
	(0.019)	(0.006)	(25.466)	(79.376)	(222.004)	(162.359)	(131.075)
Mean TRAIL	0.080	0.015	142.823	307.071	1207.642	889.229	664.507
% Effect TRAIL	0.329	72.131	19.340	26.494	11.359	5.844	-1.606
GBL Treatment	0.010	0.000	1.118	21.933	-308.116	-323.393	-396.598
	(0.019)	(0.006)	(39.512)	(134.271)	(524.917)	(393.604)	(351.505)
Mean GBL	0.112	0.022	135.893	404.919	1564.029	1142.350	853.062
% Effect GBL	9.079	0.794	0.823	5.417	-19.700	-28.309	-46.491
Sample Size	6216	6216	6216	6216	6216	6216	6216

Notes:
Treatment effects presented. Sample includes all sample households with at most 1.5 acres of land in TRAIL and GBL villages. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Standard errors in parentheses are clustered at the village level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-3: Robustness of Results 3. Program Impacts: Treatment Effects in Agriculture. Recommended/Formed Group.

	Cultivate (%) (1)	Acreage (Acres) (2)	Harvested quantity (Kg) (3)	Cost of Production (₹) (4)	Revenue (₹) (5)	Value Added (₹) (6)	Imputed Profit‡ (₹) (7)
				Potato			
TRAIL Treatment	0.048	0.094***	973.523***	1908.809***	4014.051***	2113.447***	1942.967***
	(0.031)	(0.028)	(302.107)	(721.690)	(1185.867)	(619.217)	(589.612)
Mean TRAIL	0.715	0.333	3646.124	8474.628	14285.467	5739.479	4740.893
% Effect TRAIL	6.650	28.360	26.700	22.524	28.099	36.823	40.983
GBL Treatment	0.052	0.053	515.496	1593.643*	2344.036	720.522	559.976
	(0.044)	(0.035)	(393.250)	(878.323)	(1724.660)	(908.683)	(855.015)
Mean GBL	0.620	0.251	2761.127	5992.080	11014.286	4997.446	4018.796
% Effect GBL	8.334	21.040	18.670	26.596	21.282	14.418	13.934
Sample Size	2,733	2,733	2,733	2,733	2,733	2,733	2,733
				Sesame			
TRAIL Treatment	0.036	0.044**	9.743	25.535	309.539*	283.120**	183.070
	(0.033)	(0.023)	(6.821)	(44.606)	(172.207)	(143.565)	(125.542)
Mean TRAIL	0.581	0.266	81.624	436.910	1957.498	1519.558	1080.800
% Effect TRAIL	6.264	16.461	11.936	5.844	15.813	18.632	16.938
GBL Treatment	-0.025	0.004	-5.741	17.993	-190.955	-207.767	-133.134
	(0.043)	(0.031)	(9.638)	(38.547)	(252.920)	(226.176)	(200.260)
Mean GBL	0.484	0.193	60.848	258.878	1513.138	1252.850	866.288
% Effect GBL	-5.141	2.017	-9.436	6.950	-12.620	-16.584	-15.368
Sample Size	2,733	2,733	2,733	2,733	2,733	2,733	2,733
				Paddy			
TRAIL Treatment	-0.004	0.035*	22.047	208.594	471.752*	272.167	134.595
	(0.032)	(0.021)	(31.055)	(179.907)	(281.557)	(241.453)	(131.514)
Mean TRAIL	0.744	0.470	569.726	2889.838	5398.490	2556.755	93.133
% Effect TRAIL	-0.538	7.447	3.870	7.218	8.739	10.645	144.519
GBL Treatment	-0.002	0.010	-37.354	-86.749	92.495	200.575	-116.557
	(0.039)	(0.028)	(67.640)	(352.057)	(442.096)	(270.836)	(228.690)
Mean GBL	0.689	0.456	672.894	3225.745	5513.227	2336.837	183.163
% Effect GBL	-0.290	2.193	-5.551	-2.689	1.678	8.583	-63.636
Sample Size	2,733	2,733	2,733	2,733	2,733	2,733	2,733
				Vegetables			
TRAIL Treatment	0.000	0.010	25.763	78.973	126.624	43.898	-16.404
	(0.021)	(0.007)	(45.535)	(105.759)	(419.095)	(322.147)	(237.174)
Mean TRAIL	0.080	0.015	142.823	307.071	1207.642	889.229	664.507
% Effect TRAIL	-0.051	70.509	18.038	25.718	10.485	4.937	-2.469
GBL Treatment	0.011	0.000	1.154	27.972	-287.323	-308.686	-383.217
	(0.036)	(0.009)	(61.136)	(179.261)	(856.439)	(670.745)	(576.998)
Mean GBL	0.112	0.022	135.893	404.919	1564.029	1142.350	853.062
% Effect GBL	10.182	2.208	0.849	6.908	-18.371	-27.022	-44.922
Sample Size	2,733	2,733	2,733	2,733	2,733	2,733	2,733

Notes:
Treatment effects presented. Sample includes all recommended/self selected (Treatment and Control 1) households with at most 1.5 acres of land in TRAIL and GBL villages. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Standard errors in parentheses are clustered at the para (hamlet) level. $^{***}: p < 0.01$, $^{**}: p < 0.05$, $^{*}: p < 0.1$.

Table A-4: Effects of Treatment on Treated. Potato

	Cultivate	Land Planted	Harvested Quantity	Cost of Production	Revenue	Value Added Profit	Imputed Value Added	Farm Income	Non-Agricultural
	(%)	(Acres)	(Kg)	(₹)	(₹)	(₹)	(₹)	(₹)	(₹)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
TRAIL Treatment	0.052	0.103***	1061.189***	2079.845***	4375.197***	2304.347***	2118.617***	2958.415***	-743.708
SE	(0.034)	(0.030)	(327.758)	(782.565)	(1285.807)	(672.217)	(639.760)	(808.777)	(4408.361)
Mean TRAIL Control 1	0.715	0.333	3646.124	8474.628	14285.467	5739.479	4740.893	8324.306	40115.808
% Effect TRAIL	7.24	30.91	29.10	24.54	30.63	40.15	44.69	35.539	-1.854
GBL Treatment	0.063	0.065	632.252	1953.564*	2874.560	884.420	687.635	498.116	-8019.467
SE	(0.054)	(0.042)	(480.127)	(1071.692)	(2107.997)	(1110.926)	(1045.470)	(1286.937)	(6078.544)
Mean GBL Control 1	0.620	0.251	2761.127	5992.080	11014.286	4997.446	4018.796	7741.833	45645.104
% Effect GBL	10.21	25.80	22.90	32.60	26.10	17.70	17.11	6.434	-17.569
Sample Size	2,733	2,733	2,733	2,733	2,733	2,733	2,733	2,733	2,733

Notes:
IV regression results presented. Assignment to treatment used an an instrument for loan take-up. Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for Treatment and Control 1 households in TRAIL and GBL villages with at most 1.5 acres of land. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Regressions also control for gender and educational attainment of household head, household caste and religion and landholding, a set of year dummies and an information village dummy. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-5: Program Impacts: Treatment Effects on Total Borrowing

	All Loans (₹) (1)	Non Program Loans[†] (₹) (2)
TRAIL	1,314.487**	1,258.715**
	(597.577)	(597.740)
TRAIL × Control 1	-216.505	-162.695
	(623.478)	(616.692)
TRAIL × Treatment	4,833.918***	-527.303
	(631.596)	(491.950)
GBL × Control 1	324.865	254.498
	(674.202)	(674.256)
GBL × Treatment	4,056.411***	96.717
	(802.940)	(556.996)
Landholding	9,078.801***	8,275.706***
	(702.236)	(696.243)
Non Hindu household	-3,287.325***	-2,772.558***
	(643.155)	(585.505)
Low caste household	-1,692.419***	-1,658.744***
	(536.831)	(462.049)
Male headed household	1,746.050***	1,349.010***
	(502.763)	(498.760)
Household head: Completed Primary Schooling	-4.810	66.508
	(425.426)	(392.054)
Constant	-84.527	795.939
	(702.412)	(644.860)
Treatment Effect		
TRAIL	5050***	-364.60
	(767.50)	(630.90)
Mean TRAIL Control 1	5590	5590
% Effect TRAIL	90.34	-6.52
GBL	3732***	-157.8
	(845.10)	(600.2)
Mean GBL Control 1	4077	4077
% Effect GBL	91.54	-3.87
Recommendation/Group Formation Effect		
TRAIL Recommendation	-216.50	-162.7
	(623.50)	(616.7)
GBL Group Formation	324.90	254.50
	(674.20)	(674.30)
Sample size	6,204	6,204

Notes:
Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. Regressions also control for a set of year dummies and an information village dummy. % Effect: Treatment effect as a percentage of the mean of the relevant Control 1 group. [†]: Non-Program loans refer to loans from sources other than the TRAIL/GBL schemes. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-6: Program Impacts: Treatment Effects in Potato Cultivation

	Cultivate (%) (1)	Land planted (Acres) (2)	Harvested quantity (Kg) (3)	Cost of production (₹) (4)	Revenue (₹) (5)	Value added (₹) (6)	Imputed profit[‡] (₹) (7)
TRAIL	0.057*	0.032	292.763	781.354	526.844	-263.985	-290.129
	(0.030)	(0.027)	(304.252)	(684.127)	(1,222.338)	(617.008)	(580.964)
TRAIL × Control 1	0.069*	0.01	137.647	446.827	433.107	-19.9	-99.093
	(0.035)	(0.026)	(293.854)	(705.030)	(1,186.249)	(600.392)	(565.336)
TRAIL × Treatment	0.116***	0.104***	1,113.018***	2,356.565***	4,444.731***	2,089.342***	1,840.401***
	(0.031)	(0.024)	(264.634)	(615.135)	(1,037.906)	(524.430)	(491.336)
GBL × Control 1	0.06	-0.002	-39.003	-360.801	-641.264	-260.993	-406.701
	(0.037)	(0.033)	(381.224)	(804.040)	(1,685.321)	(945.242)	(897.473)
GBL × Treatment	0.114***	0.050*	475.432	1,240.496*	1,702.70	453.143	147.007
	(0.035)	(0.027)	(299.069)	(703.861)	(1,249.871)	(616.922)	(569.248)
Landholding	0.364***	0.471***	5,172.909***	10,555.661***	21,918.658***	11,262.024***	10,370.559***
	(0.028)	(0.031)	(351.564)	(795.920)	(1,424.911)	(718.648)	(682.919)
Non Hindu household	-0.246***	-0.126***	-1,456.438***	-3,666.870***	-6,134.404***	-2,471.695***	-1,937.662***
	(0.041)	(0.032)	(353.577)	(791.591)	(1,424.167)	(705.968)	(649.551)
Low caste household	-0.091***	-0.067***	-785.399***	-2,350.417***	-3,662.629***	-1,299.159**	-1,201.942**
	(0.027)	(0.024)	(275.179)	(619.846)	(1,080.364)	(516.179)	(480.608)
Male headed household	0.173***	0.089***	994.458***	2,193.331***	3,533.360***	1,319.783**	925.504*
	(0.039)	(0.025)	(280.924)	(604.000)	(1,176.181)	(594.393)	(545.037)
Household head: Completed Primary Schooling	-0.032*	0.001	26.736	318.901	134.955	-198.726	-5.649
	(0.018)	(0.017)	(190.199)	(433.805)	(751.643)	(371.166)	(346.968)
Constant	0.372***	0.018	327.659	332.504	-421.257	-811.724	-926.514
	(0.048)	(0.033)	(377.208)	(816.221)	(1,567.057)	(826.290)	(781.899)
Treatment Effects							
TRAIL Treatment	0.047	0.095***	975.371	1909.738***	4011.624***	2109.242***	1939.494***
	(0.032)	(0.028)	(301.124)	(718.799)	(1186.538)	(621.037)	(591.339)
Mean TRAIL Control 1	0.715	0.333	3646.124	8474.628	14285.467	5739.479	4740.893
% Effect TRAIL	6.56	28.46	26.75	22.53	28.08	36.75	40.91
GBL Treatment	0.053	0.052	514.435	1601.298*	2343.964	714.137	553.708
	(0.044)	(0.035)	(395.082)	(877.219)	(1729.723)	(918.671)	(866.430)
Mean GBL Control 1	0.620	0.251	2761.127	5992.080	11014.286	4997.446	4018.796
% Effect GBL	8.59	20.79	18.63	26.72	21.28	14.29	13.78
Recommendation/Group Formation Effects							
TRAIL Recommendation	0.069	0.010	137.647	446.827	433.107	-19.900	-99.093
	(0.035)	(0.026)	(293.854)	(705.030)	(1186.249)	(600.392)	(565.336)
GBL Group Formation	0.060	-0.002	-39.003	-360.801	-641.264	-260.993	-406.701
	(0.037)	(0.033)	(381.224)	(804.040)	(1685.321)	(945.242)	(897.473)
Sample Size	6210	6210	6210	6210	6210	6210	6210

Notes: Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Regressions also control for a set of year dummies and an information village dummy. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-7: Program Impacts: Treatment Effects in Sesame Cultivation

	Cultivate (%) (1)	Land planted (Acres) (2)	Harvested quantity (Kg) (3)	Cost of production (₹) (4)	Revenue (₹) (5)	Value added (₹) (6)	Imputed profit‡ (₹) (7)
TRAIL	0.076**	0.042*	13.575*	34.763	216.493	180.985	103.803
	(0.035)	(0.021)	(7.615)	(38.950)	(175.758)	(148.649)	(127.428)
TRAIL × Control 1	0.087**	0.036*	6.924	82.209**	220.684	138.585	86.299
	(0.035)	(0.022)	(7.105)	(36.323)	(171.621)	(152.051)	(135.914)
TRAIL × Treatment	0.122***	0.079***	16.567**	108.064**	525.663***	416.808**	265.513*
	(0.035)	(0.020)	(7.703)	(42.835)	(199.094)	(175.922)	(160.609)
GBL × Control 1	0.087**	0.03	7.1	-29.721	177.119	205.714	109.877
	(0.038)	(0.029)	(9.517)	(35.876)	(249.834)	(226.649)	(202.945)
GBL × Treatment	0.063*	0.033	1.652	-12.931	-11.49	1.631	-19.497
	(0.034)	(0.023)	(7.225)	(37.340)	(169.964)	(144.475)	(124.771)
Landholding	0.369***	0.361***	101.112***	456.443***	2,619.484***	2,160.211***	1,824.131***
	(0.025)	(0.024)	(8.423)	(39.195)	(202.611)	(179.670)	(166.549)
Non Hindu household	-0.219***	-0.110***	-43.415***	-190.726***	-981.758***	-790.540***	-586.309***
	(0.040)	(0.024)	(8.564)	(37.854)	(225.407)	(197.268)	(174.952)
Low caste household	-0.046	-0.035*	-24.264***	-87.270**	-570.446***	-483.747***	-471.244***
	(0.032)	(0.019)	(6.397)	(34.847)	(148.469)	(123.300)	(104.570)
Male headed household	0.159***	0.058***	20.140***	60.999**	417.114***	357.395***	160.861
	(0.030)	(0.017)	(5.775)	(29.267)	(146.048)	(124.006)	(104.290)
Household head: Completed Primary Schooling	-0.028	-0.006	-0.845	55.517**	-31.173	-87.309	-17.108
	(0.019)	(0.012)	(4.247)	(23.008)	(102.636)	(89.738)	(79.863)
Constant	0.206***	-0.019	11.853	105.689**	-182.827	-287.901	-235.148
	(0.042)	(0.026)	(8.489)	(46.796)	(214.331)	(181.757)	(157.018)
Treatment Effects							
TRAIL Treatment	0.035	0.044**	9.643	25.855	304.979*	278.223**	179.214
	(0.033)	(0.023)	(6.738)	(44.185)	(170.527)	(142.192)	(124.683)
Mean TRAIL Control 1	0.581	0.266	81.624	436.910	1957.498	1519.558	1080.800
% Effect TRAIL	6.02	16.39	11.81	5.92	15.58	18.31	16.58
GBL Treatment	-0.024	0.003	-5.449	16.790	-188.605	-204.084	-129.374
	(0.044)	(0.031)	(9.768)	(39.016)	(256.021)	(229.475)	(203.585)
Mean GBL Control 1	0.484	0.193	60.848	258.878	1513.138	1252.850	866.288
% Effect GBL	-4.97	1.46	-8.95	6.49	-12.46	-16.29	-14.93
Recommendation/Group Formation Effects							
TRAIL Recommendation	0.087**	0.036*	6.924	82.209**	220.684	138.585	86.299
	(0.035)	(0.022)	(7.105)	(36.323)	(171.621)	(152.051)	(135.914)
GBL Group Formation	0.087**	0.030	7.100	-29.721	177.119	205.714	109.877
	(0.038)	(0.029)	(9.517)	(35.876)	(249.834)	(226.649)	(202.945)
Sample Size	6210	6210	6210	6210	6210	6210	6210

Notes: Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. ‡: Imputed profit = Value Added − shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Regressions also control for a set of year dummies and an information village dummy. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-8: Program Impacts: Treatment Effects in Paddy Cultivation

	Cultivate (%) (1)	Land planted (Acres) (2)	Harvested quantity (Kg) (3)	Cost of production (₹) (4)	Revenue (₹) (5)	Value added (₹) (6)	Imputed profit‡ (₹) (7)
TRAIL	-0.034	-0.016	-47.293	-275.251	-390.783	-82.75	-27.327
	(0.025)	(0.026)	(47.688)	(233.848)	(408.812)	(268.409)	(151.034)
TRAIL × Control 1	0.091***	0.015	26.824	87.853	112.528	5.425	-54.62
	(0.029)	(0.022)	(35.814)	(183.485)	(316.830)	(223.642)	(118.265)
TRAIL × Treatment	0.086***	0.051**	49.034	300.108	583.611*	273.215	80.825
	(0.030)	(0.021)	(31.896)	(184.719)	(303.178)	(234.779)	(117.951)
GBL × Control 1	0.04	0.04	153.415**	504.43	485.573	3.818	76.057
	(0.029)	(0.030)	(75.249)	(378.657)	(512.717)	(267.850)	(219.305)
GBL × Treatment	0.039	0.051*	116.894**	429.441	600.05	217.345	-44.305
	(0.031)	(0.029)	(57.721)	(307.952)	(432.811)	(231.699)	(166.892)
Landholding	0.461***	0.894***	1,095.056***	4,936.404***	9,176.493***	4,355.661***	965.438***
	(0.025)	(0.035)	(58.297)	(306.641)	(502.942)	(310.850)	(162.100)
Non Hindu household	-0.093***	0.034	74.091	529.416	653.812	246.985	287.07
	(0.030)	(0.036)	(78.212)	(416.421)	(578.789)	(311.277)	(193.235)
Low caste household	-0.068***	0.013	2.892	-233.103	-29.732	244.56	102.271
	(0.023)	(0.021)	(39.289)	(199.890)	(294.968)	(188.775)	(114.158)
Male headed household	0.203***	0.126***	217.535***	1,007.446***	1,883.556***	857.514***	-65.749
	(0.037)	(0.026)	(39.833)	(186.046)	(335.587)	(197.871)	(77.607)
Household head: Completed Primary Schooling	-0.042**	-0.033*	-78.276**	-59.974	-389.578	-317.842*	166.684**
	(0.017)	(0.019)	(31.112)	(161.223)	(288.158)	(175.847)	(80.034)
Constant	0.420***	0.009	-462.507***	104.135	-666.993	-710.122**	-545.365***
	(0.046)	(0.033)	(62.141)	(278.700)	(499.964)	(319.610)	(175.018)
Treatment Effects							
TRAIL Treatment	-0.005	0.036*	22.210	212.254	471.083*	267.790	135.445
	(0.032)	(0.020)	(30.817)	(178.716)	(280.807)	(241.457)	(131.079)
Mean TRAIL Control 1	0.744	0.470	569.726	2889.838	5398.490	2556.755	93.133
% Effect TRAIL	-0.71	7.66	3.90	7.35	8.73	10.47	145.43
GBL Treatment	-0.001	0.011	-36.521	-74.989	114.480	213.527	-120.362
	(0.039)	(0.029)	(68.446)	(354.916)	(447.467)	(271.907)	(227.270)
Mean GBL Control 1	0.689	0.456	672.894	3225.745	5513.227	2336.837	183.163
% Effect GBL	-0.07	2.39	-5.43	-2.33	2.08	9.14	-65.71
Recommendation/Group Formation Effects							
TRAIL Recommendation	0.091***	0.015	26.824	87.853	112.528	5.425	-54.620
	(0.029)	(0.022)	(35.814)	(183.485)	(316.830)	(223.642)	(118.265)
GBL Group Formation	0.040	0.040	153.415**	504.430	485.573	3.818	76.057
	(0.029)	(0.030)	(75.249)	(378.657)	(512.717)	(267.850)	(219.305)
Sample Size	6210	6210	6210	6210	6210	6210	6210

Notes: Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Regressions also control for a set of year dummies and an information village dummy. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-9: Program Impacts: Treatment Effects in Vegetable Cultivation

	Cultivate (%) (1)	Land planted (Acres) (2)	Harvested quantity (Kg) (3)	Cost of production (₹) (4)	Revenue (₹) (5)	Value added (₹) (6)	Imputed profit‡ (₹) (7)
TRAIL	0.013	0.003	92.447	87.622	257.174	155.847	174.656
	(0.023)	(0.005)	(58.992)	(103.217)	(432.459)	(332.506)	(268.585)
TRAIL × Control 1	-0.014	-0.004	-38.028	-71.467	-26.997	52.419	35.741
	(0.016)	(0.003)	(34.800)	(71.290)	(236.179)	(174.366)	(144.961)
TRAIL × Treatment	-0.013	0.007	-10.403	9.881	110.163	104.37	25.055
	(0.018)	(0.007)	(54.886)	(108.466)	(407.183)	(310.625)	(243.261)
GBL × Control 1	0.043	0.009	75.01	173.209	819.137	631.206	537.885
	(0.032)	(0.008)	(56.237)	(167.108)	(819.597)	(642.212)	(562.809)
GBL × Treatment	0.054**	0.009**	76.132**	195.138**	511.006*	307.802	141.274
	(0.021)	(0.004)	(31.890)	(87.212)	(278.071)	(210.586)	(164.347)
Landholding	0.046***	0.015***	160.812**	194.991**	1,220.125***	1,007.720***	834.731***
	(0.016)	(0.005)	(62.363)	(80.159)	(381.881)	(305.442)	(267.465)
Non Hindu household	-0.088***	-0.019***	-212.183***	-378.125***	-1,870.716***	-1,468.966***	-1,169.962***
	(0.025)	(0.006)	(61.072)	(115.947)	(480.626)	(371.842)	(303.036)
Low caste household	-0.067***	-0.014***	-171.913***	-354.617***	-1,648.337***	-1,273.133***	-1,059.710***
	(0.024)	(0.005)	(51.936)	(102.421)	(453.177)	(352.217)	(287.474)
Male headed household	0.065***	0.012***	75.320***	265.743***	641.904***	367.626**	207.3
	(0.013)	(0.003)	(26.894)	(52.538)	(217.433)	(169.123)	(133.128)
Household head: Completed Primary Schooling	-0.021**	-0.005	-60.686**	-101.688**	-511.050***	-402.102***	-280.270**
	(0.010)	(0.003)	(28.471)	(47.914)	(189.477)	(149.930)	(134.469)
Constant	0.036*	0.003	-15.724	55.99	139.705	90.951	43.941
	(0.020)	(0.005)	(46.817)	(83.638)	(324.532)	(251.147)	(206.444)
Treatment Effects							
TRAIL Treatment	0.000	0.011	27.625	81.348	137.159	51.952	-10.686
	(0.021)	(0.007)	(45.780)	(106.058)	(418.599)	(321.736)	(236.536)
Mean TRAIL Control 1	0.080	0.015	142.823	307.071	1207.642	889.229	664.507
% Effect TRAIL	0.33	72.13	19.34	26.49	11.36	5.84	-1.61
GBL Treatment	0.010	0.000	1.122	21.929	-308.131	-323.404	-396.611
	(0.036)	(0.009)	(61.427)	(180.672)	(862.988)	(676.455)	(582.580)
Mean GBL Control 1	0.112	0.022	135.893	404.919	1564.029	1142.350	853.062
% Effect GBL	9.08	0.80	0.83	5.42	-19.70	-28.31	-46.49
Recommendation/Group Formation Effects							
TRAIL Recommendation	-0.014	-0.004	-38.028	-71.467	-26.997	52.419	35.741
	(0.016)	(0.003)	(34.800)	(71.290)	(236.179)	(174.366)	(144.961)
GBL Group Formation	0.043	0.009	75.010	173.209	819.137	631.206	537.885
	(0.032)	(0.008)	(56.237)	(167.108)	(819.597)	(642.212)	(562.809)
Sample Size	6210	6210	6210	6210	6210	6210	6210

Notes: Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. ‡: Imputed profit = Value Added – shadow cost of labour. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Regressions also control for a set of year dummies and an information village dummy. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.

Table A-10: Program Impacts: Treatment Effects in Aggregate Farm Value-Added and Non Agricultural Incomes

	Farm Value-Added (₹) (1)	Non-Agricultural Income (₹) (2)
TRAIL	-9.903	1,167.37
	(751.40)	(4,619.44)
TRAIL × Control 1	176.529	-11,159.343***
	(777.903)	(3,686.649)
TRAIL × Treatment	2,883.735***	-11,767.343***
	(703.472)	(4,211.561)
GBL × Control 1	579.745	-4,744.83
	(1,106.787)	(5,088.544)
GBL × Treatment	979.921	-10,837.461**
	(747.682)	(4,600.405)
Landholding	18,785.617***	3,415.37
	(932.284)	(5,374.352)
Non Hindu household	-4,484.216***	5,675.42
	(904.270)	(4,559.521)
Low caste household	-2,811.479***	1,324.12
	(584.828)	(3,610.341)
Male headed household	2,902.318***	-4,912.29
	(857.462)	(10,315.913)
Household head: Completed Primary Schooling	-1,005.980**	-260.836
	(507.548)	(3,031.750)
Constant	-1,718.80	45,738.771***
	(1,065.959)	(8,109.731)
Treatment Effects		
TRAIL Treatment	2707.206***	-608.000
	(739.438)	(4153.557)
Mean TRAIL Control 1	10705.022	40115.808
% Effect TRAIL	25.29	-1.52
GBL Treatment	400.176	-6092.631
	(1069.611)	(4959.881)
Mean GBL	9729.483	45645.104
% Effect GBL	4.11	-13.35
TRAIL vs GBL p-value	0.077	0.393
Recommendation/Group Formation Effects		
TRAIL Recommendation	176.529	-11159.343***
	(777.903)	(3686.649)
GBL Group Formation	579.745	-4744.830
	(1106.787)	(5088.544)
Sample Size	6,210	6,210

Notes:
Treatment effects are computed from regressions that follow equation (30) in the text and are run on household-year level data for all sample households with at most 1.5 acres of land. % Effect: Treatment effect as a percentage of the Mean of Control 1 group. Regressions also control for a set of year dummies and an information village dummy. Standard errors in parentheses are clustered at the hamlet level. $^{***}: p < 0.01, ^{**}: p < 0.05, ^{*}: p < 0.1$.