

Selection into Credit Markets: Evidence from Agriculture in Mali

September 2020

Lori Beaman, Dean Karlan, Bram Thuysbaert, and Christopher Udry¹

Abstract

We use a two-stage experiment on agricultural lending in Mali to test whether selection into lending is predictive of heterogeneous returns to capital. Understanding this heterogeneity, and the selection process which reveals it, is critical for guiding modelling of credit markets in developing countries, as well as for policy. We find such heterogeneity: returns to capital are higher for farmers who borrow than for those who do not. In our first stage, we offer loans in some villages and not others. In the second stage, we provide cash grants to a random subset of all farmers in villages where no loans were offered, and to a random subset of the farmers who do not borrow in villages where loans were offered. We estimate seasonal returns to the grant of 130% for borrowers, whereas we find returns near zero for the sample representative of those who had recently not borrowed. We also provide evidence that there are some farmers – particularly those that are poor at baseline – that have high returns but do not receive a loan.

JEL: D21, D92, O12, O16, Q12, Q14

Keywords: credit markets; agriculture; returns to capital

¹ Lori Beaman: l-beaman@northwestern.edu, Northwestern University; Dean Karlan: karlan@northwestern.edu, Northwestern University, IPA, J-PAL, and NBER; bram.thuysbaert@ugent.be, Ghent University; and Christopher Udry: christopher.udry@northwestern.edu, Northwestern University. Paper previously circulated as “Self-selection into Credit Markets: Evidence from Agriculture in Mali”. The authors thank partners Save the Children and Soro Yiriwaso for their collaboration. Thanks to Yann Guy, Pierrick Judeaux, Henriette Hanicotte, Nicole Mauriello, Diego Santa Maria, and Aissatou Ouedraogo for excellent research assistance and to the field staff of Innovations for Poverty Action – Mali office. We thank Dale Adams and Alex W. Cohen for helpful comments. All errors and opinions are our own.

1 Introduction

The return to investment in productive activities depends on a myriad of influences, reflecting both the realization of risk and underlying heterogeneity in the characteristics, effort, and constraints of producers. Some of this variation may be apparent to outside observers; much may not. Some of this variation may be apparent to producers themselves; some may not. A primary role of financial markets is to help capital flow to the highest return activities.

In a two-stage randomized controlled trial of loans and grants for low-income farmers in rural Mali, we demonstrate positive selection into borrowing with respect to marginal returns to capital. We designed a two-stage protocol specifically to test whether returns to capital are heterogeneous and sufficiently predictable that high return agents receive loans. The sample consists of likely liquidity constrained farmers in rural Mali, a capital-poor economy not well integrated into global financial markets. In stage one (the loan stage), a microcredit organization (Soro Yiriwaso, “Soro”) identified 198 villages that were within their expansion plans but which they had not previously entered. Soro then offered group-liability loans to all women farmers in 88 villages, randomly selected from the 198 villages. In these loan treatment villages, some farmers choose, or are chosen by their peers, to borrow via group liability loans under a community association. In stage two of the experiment (the cash grant stage), after first waiting for households and the associations to make their loan decisions from stage one, we announced and immediately gave cash grants (40,000 FCFA, about US\$140) to a random subset of households that did not borrow in the loan villages and of all households in the no-loan villages.

The first stage effectively creates two samples over which we compare the returns to the stage two cash grants: 88 “loan villages” (where we measure returns to the cash grant for individuals who did not borrow) and 110 “no-loan” villages (where we measure returns to the cash grant for all individuals, i.e. those who would have borrowed had they been offered a loan as well as those who would not have borrowed). Comparing the average returns in these two samples allows us to test an important selection question: do those who do not borrow have lower average returns to a grant than the implied returns to a grant among farmers who did borrow?

We find large average increases in investment and agricultural profits for the non-selected population (i.e., grant recipients vs. non-grant-recipients in no-loan villages). Specifically, the cash grants in no-loan villages led to a significant increase in land being cultivated (8.7%, se=3.3%), fertilizer use (18%, se=5%), and overall input expenditures (16%, se=4%). These households also experienced an increase in the value of their agricultural output and in gross

profit² by 13% (se=4%) and 12% (se=5%), respectively. Thus, we observe a statistically significant and economically meaningful increase in investments in cultivation and gross profit from relaxing capital constraints. This impact on gross profit even persists after an additional agricultural season. In this environment, therefore, capital constraints are limiting investments in cultivation.³

However, we find low, indeed zero, average returns to the cash grants for those who did *not* borrow (i.e., the difference between randomly receiving a grant and not among non-borrowers in loan villages). In loan villages, households given grants did not earn any higher gross profit from the farm than households not provided grants. This contrasts sharply with households given grants in the no-loan villages who had large increases in gross profit relative to those not provided grants. Therefore, we conclude that households which borrowed, and were thus selected out of the sample frame in loan villages, had higher marginal returns than those who did not borrow. The differences in the impact of the grants between households who would borrow and those who do not are substantial. We calculate that among borrowing households, farm output would have increased by US\$222 (se=120) and farm gross profit by US\$183 (se=96) had those households received grants. In contrast, we estimate that among households who do not borrow, receipt of the grant generates only US\$25 of additional output and US\$1.04 additional gross profit (neither being statistically significantly different from zero).

Thus, putting the findings from the two samples together, we infer that farmers with particularly high returns to capital are much more likely to select – or be selected – into borrowing. This implies that some of the variation in returns is predictable *ex ante*, and that farmers are aware of this heterogeneity in expected returns.

Although 93% of non-borrowing households report farming as their primary source of income, perhaps non-borrowers did not invest in farming because they had higher return opportunities elsewhere. To examine this, we also look at other outcomes such as livestock ownership and

² We do not have a complete profit measure, and use instead the term “gross profit” for agricultural revenue net of most, but not all, expenses. Importantly, the value of family and unpaid labor is not subtracted. See section 2.3.

³ The increase in investment contingent upon receipt of the grant is sufficient to reject neoclassical separation, but not to demonstrate the existence of binding capital constraints. For example, in models akin to Banerjee and Duflo (2012) with an upward-sloping supply of credit for each farmer, a capital grant could completely displace borrowing from high-cost lenders, lower the opportunity cost of capital to the farmer and induce greater investment even though the farmer could have borrowed more from the high cost lender and thus was not capital constrained in a strict sense. However, there is no evidence that these grants lowered total borrowing. We therefore refer to capital market imperfections that could cause investment responses to cash grants simply as credit constraints.

small business operations. However, we do not find evidence of grant recipients in loan villages investing the cash in alternative activities more than their counterparts in no-loan villages.

Farmers with high returns to grants are differentially selected into borrowing from Soro. But how efficient is this selection? In particular, are there women with high return investment opportunities who do not borrow? To examine this, we compare the distribution of returns in no-loan villages (thus a representative sample of everyone) to loan villages (thus only to those selected *out* from borrowing, either by themselves or their peers). In no-loan villages we find no correlation between baseline gross profit and marginal returns to the grant. In the loan villages, however, baseline gross profits are *negatively* correlated with marginal returns to the grant. More specifically, those with higher baseline gross profit have close to zero marginal returns to the grant, whereas those with low baseline gross profit have positive marginal returns to the grant. We find that high marginal return, low baseline gross profit farmers are under-represented among borrowers, suggesting that there is a subset of particularly poor women who face higher borrowing frictions than other farmers.

We also exploit a machine learning algorithm (Athey & Imbens, 2016; Athey, Tibshirani, & Wager, 2019; Athey & Wager, 2019; Wager & Athey, 2018) to detect heterogeneity and estimate conditional average treatment effects (CATEs). CATEs trained in the no-loan villages show a high density of farmers with high baseline profits and high CATEs. When the causal forest is trained in the loan villages, however, these farmers are notably less represented. Farmers who have both high marginal returns and high baseline profits are much more likely to be borrowers, while farmers with high CATEs and low baseline profits are more likely to be non-borrowers.

The heterogeneity in returns to loans that we discover is consistent with Meager (2020), which uses Bayesian hierarchical modeling of the quantiles of response to seven different microcredit interventions with RCTs to show evidence of strongly positive returns for a small set of borrowers, but near zero returns to borrowing for the large majority. Crépon et al. (2020) also finds a great deal of heterogeneity in the returns to loans (and grants) among microentrepreneurs in Egypt. Our finding that farmers are aware of these heterogenous returns is similar to that of Hussam et al. (2020), which finds that businesses (in their case, nonfarm enterprises in urban India) have widely varying marginal returns to grants, and that entrepreneurs themselves and community members are able to distinguish between those with relatively high and low returns. In a different setting (enterprise business plan competitions in Nigeria and in Ghana), McKenzie (2018; 2015), McKenzie and Sansone (2019), and Fafchamps and Woodruff (2017) provide evidence of the difficulty in predicting who will be the most successful, although average estimated returns are high.

Our experiment also speaks to three additional questions important to academia and policy: First, do loans generate different investment behavior than grants? Second, what is the impact of a microlending program that targets farmers (as compared to the more standard microenterprise focus of microlenders)? Third, are the impacts of the cash grants persistent after seven years?

First, on comparing grants to loans, about 21% of households in our sample received loans (in loan villages), which is typical of other microcredit contexts, but of course far below the 100% take-up rate of the grants. The average loan size was 32,000 FCFA (US\$113). Like the grants, offering loans led to an increase in investments in cultivation, particularly fertilizer, insecticides and herbicides, and an increase in agricultural output. We do not detect, however, a statistically significant increase in gross profit. Our treatment on the treated estimates of the impact of borrowing on the cultivation activities and harvests of those who borrowed are large and consistent in magnitude with our entirely separate estimates of the impact of grants on borrowers. Therefore, it does not appear that the lending process leads to dramatically different behavior on the part of farmers than cash grants. This is consistent with Crépon et al. (2020).

Second, underlying our experiment is an estimate of the impact of an agriculture microcredit program: we find high returns, particularly when compared to experiments estimating the impact of microcredit designed for entrepreneurship.⁴ High average returns to agricultural investment could emerge when farmers lack capital and face credit and savings constraints. Microcredit organizations have attempted to relieve credit constraints, but most microcredit lenders focus on small or micro business entrepreneurial financing. Furthermore, the typical microcredit loan requires frequent, small repayments and therefore does not facilitate investments in agriculture, where income comes as a lump sum once or twice a year (see Fink, Jack, & Masiye, 2018 for an experiment demonstrating the importance of this timing issue for farmers; see Karlan & Mullainathan, 2007 for a discussion of this). By contrast, the loan product studied here is designed for farmers by providing capital at the beginning of the planting season and requiring repayment as a lump sum at harvest. Maitra et al. (2020) also finds positive impacts from an agricultural microcredit program on farm value-added in India for one version of the program,

⁴ The evidence from traditional microcredit, targeting micro enterprises, is more mixed; some randomized evaluations find an increase in investment in self-employment activity (Angelucci, Karlan, & Zinman, 2015; Crépon, Devoto, Duflo, & Pariente, 2015) while others do not (Attanasio, Augsburg, De Haas, Fitzsimons, & Harmgart, 2015; Augsburg, De Haas, Harmgart, & Meghir, 2015; Banerjee, Duflo, Glennerster, & Kinnan, 2015; Karlan & Zinman, 2011; Tarozzi, Desai, & Johnson, 2015). See Banerjee, Karlan and Zinman (2015) and Meager (2019) for an overview of the above seven studies. Rarely have randomized evaluations of microcredit found an increase in the profitability of small businesses as a result of access to microcredit, at least at the mean or median. These limited results from microcredit come despite evidence that the marginal returns to capital can be quite high for micro-enterprises (de Mel, McKenzie, & Woodruff, 2008).

though not for a version which targeted the program differently. However, lending may not be sufficient to induce investments in the presence of other constraints. Farmers may be constrained by a lack of insurance (Karlan, Osei-Akoto, Osei, & Udry, 2013), have time inconsistent preferences (Duflo, Kremer, & Robinson, 2011), or face high costs of acquiring inputs (Suri, 2011).

These loan impact results are in stark contrast to a long history of failed agricultural credit programs (Adams, 1971), which often were implemented as subsidized government programs and thus plagued by politics (Adams, Graham, & Von Pischke, 1984). In the expansion of microcredit in the 1980s and onward, we had seen several mostly simultaneous shifts: group instead of individual lending (de Quidt, Fetzer, & Ghatak, 2012; although now this trend is reversing, e.g. see Giné & Karlan, 2014); high frequency repayment instead of one-time balloon payments (see Field, Pande, Papp, & Rigol, 2013 for an important test, demonstrating the potential benefits to delayed-start repayment); nongovernment (and now for-profit) lending instead of government; and, enterprise targeted loans instead of agricultural (Armendariz de Aghion & Morduch, 2010; Karlan & Morduch, 2009). The loan impact component of this study tests a new model of agricultural credit with group lending, balloon payment, and nonprofit management (with little to no subsidy).

Third, we conducted a follow-up survey in 2017, almost seven years after the grants, to measure their long-term effects. We find no evidence that the grants had a persistent effects over this extended period, which was marked by political upheaval and systematic changes in cropping patterns, as well as the highly variable seasonal rainfall typical of the West African semiarid tropics.

2 The Experimental Design and Data

2.1 The Experimental Design

Agriculture in most of Mali, and in all of our study area, is exclusively rain fed. Evidence from nearby Burkina Faso suggests that income shocks translate into consumption volatility (Kazianga & Udry, 2006), so improved credit markets can have important welfare consequences from both increasing average production and insulating consumption from output volatility. The main crops grown in the area include millet/sorghum, maize, cotton (mostly grown by men), and rice and groundnuts (mostly grown by women). At baseline, about 40% of households were using fertilizer⁵, and 51% were using other chemical inputs (herbicides, insecticide).

⁵ The government of Mali introduced heavy fertilizer subsidies in 2008. The price of fertilizer was fixed to 12,500 FCFA (US\$44) per 50 kg of fertilizer. This constituted a 20% to 40% subsidy, depending on the type of fertilizer and

The sample consists of 198 villages identified by Soro as villages that they had not previously entered but that were within their expansion plans. The villages are located in two *cercles* (an administrative unit larger than the village but smaller than a region) in the Sikasso region of Mali.⁶

Figure 1 presents the design.

Stage One: Loans

Soro, a Malian microcredit organization and affiliate of Save the Children (an international nongovernmental organization based in the United States), marketed, financed, implemented, and serviced the loans. After a baseline survey was completed (see below), we randomly assigned the 198 villages to either loan (88 villages) or no-loan (110 villages) status using a re-randomization technique ensuring balance on key variables.⁷ This stage one randomization was done at the village level (because that is how Soro marketed and implemented loans).

Soro offered their standard agricultural loan product, called *Prêt de Campagne*, in the 88 loan villages. This product is given exclusively to women, but naturally money may be fungible within the household. Unlike most microloan products, the loan is designed specifically for farmers: loans are dispersed at the beginning of the agricultural cycle in May–July and repayment is required after harvest. The loan is administered to groups of women organized into village associations, and each individual woman then receives an informal contract with their village association. Qualitative interviews with members outside the study villages, prior to the

year. Initial usage of the subsidy was low in rural areas initially but has grown over time, helping to explain the increase in input expenses we observe in our data from baseline to endline (Druihe & Barreiro-Huré, 2012).

⁶ Bougouni and Yanfolila are the two *cercles*, both in the northwest portion of the region and within the expansion zone of Soro. The sample was determined by randomly selecting 198 villages from the 1998 Malian census that met three criteria: (1) were within the planned expansion zone of Soro Yiriwaso, (2) were not being serviced by Soro Yiriwaso, and (3) contained at least 350 individuals (i.e., sufficient population to generate a lending group).

⁷ First, we ran a loop with a set number of iterations that randomized villages to either loan or no-loan in each iteration, and then we selected the random draw that minimized the t-values for all pairwise orthogonality tests. This is done because of the difficulties stratifying using a block randomization technique with this many baseline and continuous variables. For village-level randomization of stage one loans, we used the following: village size, whether the village was all Bambara (the dominant ethnic group in the area), distance to a paved road, distance to the nearest market, percent of households with a plough, percent of women with a plough, frequency of fertilizer use among women in the village, literacy rate, and distance to the nearest health center. For household-level randomization of stage two grants, after first stratifying on stage one village loan status, we used the following: whether the household was part of an extended family; whether the household was polygamous; an index of the household's agricultural assets, other assets, and per capita food consumption; and, the primary female respondent's land size, fertilizer use, and plough access. See Bruhn and McKenzie (2009) for a more detailed description of the randomization procedure.

intervention, revealed that the application process is informal with few administrative records at the village level. For example, there are records of neither loan applications nor denials. Nor is a record kept of more subtle, informal processes of “application” or “denial”, such as women who discuss the possibility of joining the group to get a loan but who are discouraged from joining (such data would have been helpful for ascertaining the extent of peer versus self-selection, for instance). The size of the group is not constrained by the lender; a group could add a member without decreasing the size of loan each woman received. The size of the loan to each woman is also determined through an informal, iterative process. Repayment is tracked only at the group level, and nominally there is joint liability. On average there are about 30 women per group and typically one, though up to three, associations per village. This is a limited liability environment since these households have few assets and the legal environment of Mali would make any formal recourse on the part of the bank nearly impossible. However, given that loans are administered through community associations, the social costs of default could be quite high. We observe no defaults over the two agricultural cycles during which we were collaborating with Soro.⁸

Soro offered loans in the loan villages for two years, the 2010 and 2011 agricultural seasons. The average loan size in 2010 was 32,000 FCFA (US\$113).⁹ The annual interest rate is 25% plus 3% in fees and a mandatory savings rate¹⁰ of 10%.

Women who borrowed are represented by the far-left box in Figure 1.

Stage Two: Grants

Grants worth 40,000 FCFA (US\$140) were distributed by Innovations for Poverty Action (IPA), and with no stated relationship to the loans or to Soro, to about 1,600 female survey respondents in May and June of 2010 (i.e., planting time).

In the 110 no-loan villages, households were randomly selected to receive grants and—to parallel the loans—a female household member was always the direct recipient. This corresponds to the boxes on the right side of Figure 1. US\$140 is a large grant; average input expenses, in the absence of the grant, were US\$196 and the value of agricultural output was US\$522. The size of the grant was chosen to approximate the average loan size provided by Soro, though *ex post* the

⁸ This is not atypical for Soro. In an assessment conducted by Save the Children in 2009, 0% of Soro’s overall portfolio for this loan product was at risk (more than 30 days overdue) in years 2004-2006, rising to only 0.7% in 2007.

⁹ We use the 2011 PPP exchange rate with the Malian FCFA (284 FCFA per USD) throughout the paper.

¹⁰ The mandatory savings are removed from the loan at the time of disbursement and held at the MFI.

grant is slightly larger on average than the loans. In no-loan villages, we also provided some grants to a randomly selected set of men, but we exclude those households from the analysis.¹¹

In the 88 loan villages, grant recipients were randomly selected among survey respondents who did not take out a loan (see Figure 1).¹² We attempted to deliver grants at the same time in all villages, but administrative delays on the loan side meant that most grants were delivered first in no-loan villages, and there is an average difference of 20-days between when grants were received by households in no-loan villages and their counterparts in loan villages. We discuss the implications of this delay in section 3.2.3.

In order to minimize the possibility of dynamic incentives not to borrow, we informed recipients that the grants were a one-time grant, not an ongoing program, and also distributed an additional 104 grants (one or two per village) to loan village women not in our sample. It was therefore not obvious to those in the study that borrowing precluded someone from being a grant recipient.

2.2 Identification

We focus on agricultural outcomes, so consider agricultural gross profit $Q(\text{loan}, \text{grant})$. $\{Q(0,0), Q(0,1), Q(1,0)\}$ represent the set of possible gross profits in year 1 of households in our sample.¹³ $Q(0,0)$ is a random variable representing potential profit if the household neither borrows nor receives a grant; $Q(1,0)$ and $Q(0,1)$ are similarly defined for households who receive a loan but not a grant, and for those who receive a grant but not a loan, respectively.¹⁴ The joint distribution of potential outcomes is $F(Q(0,0), Q(0,1), Q(1,0))$, and the three marginal distributions are denoted $F_{NG}(Q(0,0))$, $F_G(Q(0,1))$ and $F_B(Q(1,0))$.

¹¹ The grants to men are intended for a separate paper analyzing household dynamics and bargaining, and we do not consider them useful for the analysis here since the loans were only given to women.

¹² We determined who took out a loan by matching names and basic demographic characteristics from the loan contracts between the client and Soro, which Soro shared with us on an ongoing basis. There were a few cases (67) where Soro allowed late applications for loans and households received both a grant and a loan. The majority (41 out of 67) of these cases occurred because there were multiple adult women in the household, and one took out a loan and another received a grant. We include controls for these households. The results are similar if these observations are excluded.

¹³ This is a minor adaptation of the standard potential outcomes notation building on Rubin (1974); Heckman (1992, 1997); Imbens and Angrist (1994); Angrist et al. (1996); Heckman et al. (1997).

¹⁴ There is a fourth logically possible potential outcome, $Q(1,1)$ for households who both borrow and receive a grant, but this is irrelevant in our context because no one who receives a loan is ever assigned to the grant treatment.

Define $G \in \{0,1\}$ and $L \in \{0,1\}$ as random variables that designate a household's status in the grant treatment arm and in a loan treatment village, respectively. Furthermore, define B as a binary variable representing the observed loan take-up outcome for each household. It is useful to introduce potential treatments $B(1)$ and $B(0)$. Since households in no-loan villages do not borrow, $B = B(0) = 0$. $B(1) = 1$ if the household would borrow if located in a loan village, and $B(1) = 0$ if the household would not borrow if located in a loan village. Therefore, we can write

$$B = B(1)L. \quad (1)$$

Furthermore, define the effect on profit of receiving a grant (without a loan) as $\Delta_G Q \equiv Q(0,1) - Q(0,0)$. Our first goal is to identify the expected value of the effect on profit of receiving a grant for households for which $B(1) = 1$ versus those for which $B(1) = 0$, that is $\mathbb{E}(\Delta_G Q | B(1) = 1)$, and $\mathbb{E}(\Delta_G Q | B(1) = 0)$.

Similarly, define the effect on profit of borrowing without a grant as $\Delta_B Q \equiv Q(1,0) - Q(0,0)$. Our second goal is to identify the expected treatment effect of borrowing on those who would borrow if loans were available: $\mathbb{E}(\Delta_B Q | B(1) = 1)$. The two-stage randomization provides identification of these expected treatment effects.

The first stage randomization of villages ensures

$$\{Q(0,0), Q(0,1), Q(1,0), B(0), B(1)\} \perp L. \quad (2)$$

The second stage randomization of grants across the random sample when $L = 0$ and across non-borrowers when $L = 1$ ensures

$$\{Q(0,0), Q(0,1), Q(1,0), B(0), B(1)\} \perp G | L = 0 \quad (3)$$

$$\{Q(0,0), Q(0,1)\} \perp G | (B = 0, L = 1). \quad (4)$$

There is 100% take-up of the offer of a grant, so in our sample of the full population of no-loan villages we observe

$$Q | (L = 0) = (1 - L)[Q(0,1)G + Q(0,0)(1 - G)]$$

and in our sample of non-borrowers in loan villages, we observe

$$Q | (L = 1) = L(1 - B)[Q(0,1)G + Q(0,0)(1 - G)]$$

Therefore, (2) and (3) imply that data from the full population of no-loan villages can be used to identify the conditional marginal distributions

$$F_G(Q(0,1)|L = 0, G = 1) = F_G(Q(0,1)|L = 0, G = 0) = F_G(Q(0,1)) \quad (5)$$

and

$$F_{NG}(Q(0,0)|L = 0, G = 0) = F_{NG}(Q(0,0)|L = 0, G = 1) = F_{NG}(Q(0,0)). \quad (6)$$

Similarly, (2) and (4) imply that data from the population of non-borrowers in loan villages can be used to identify the conditional marginal distributions (7) and (8) from the profits of those who receive and do not receive a grant, respectively:

$$\begin{aligned} F_G(Q(0,1)|B = 0, L = 1, G = 1) &= F_G(Q(0,1)|B = 0, L = 1, G = 1) \\ &= F_G(Q(0,1)|B = 0) \end{aligned} \quad (7)$$

$$\begin{aligned} F_{NG}(Q(0,0)|B = 0, L = 1, G = 0) &= F_{NG}(Q(0,0)|B = 0, L = 1, G = 0) \\ &= F_{NG}(Q(0,0)|B = 0). \end{aligned} \quad (8)$$

Moreover, (2) implies that data from the population of borrowers in loan villages can be used to identify the conditional marginal distribution

$$\begin{aligned} F_B(Q(1,0)|B(1) = 1, L = 1) &= F_B(Q(1,0)|B(1) = 1, L = 0) \\ &= F_B(Q(1,0)|B(1) = 0). \end{aligned} \quad (9)$$

The loan village population provides an estimate of $\mathbb{P}(B = 1|L = 1) = \mathbb{P}(B = 1)$, which together with (5) and (7) and (6) and (8) provides

$$\begin{aligned} F_G(Q(0,1)|B = 1) &= \frac{F_G(Q(0,1)) - F_G(Q(0,1)|B = 0)(1 - \mathbb{P}(B = 1))}{\mathbb{P}(B = 1)} \\ F_{NG}(Q(0,0)|B = 1) &= \frac{F_{NG}(Q(0,0)) - F_{NG}(Q(0,0)|B = 0)(1 - \mathbb{P}(B = 1))}{\mathbb{P}(B = 1)}. \end{aligned} \quad (10)$$

From (3) and (4), we have

$$\begin{aligned} \mathbb{E}(Q(0,1)|B = 0, L = 1, G = 1) - \mathbb{E}(Q(0,0)|B = 0, L = 1, G = 0) \\ = \mathbb{E}(Q(0,1) - Q(0,0)|B = 0, L = 1) \equiv \mathbb{E}(\Delta_G Q|B = 0, L = 1) \end{aligned} \quad (11)$$

and

$$\begin{aligned} \mathbb{E}(Q(0,1)|L = 0, G = 1) - \mathbb{E}(Q(0,0)|L = 0, G = 0) \\ = \mathbb{E}(Q(0,1) - Q(0,0)|L = 0) \equiv \mathbb{E}(\Delta_G Q|L = 0). \end{aligned} \quad (12)$$

Equation (11), along with $\mathbb{P}(B = 1|L = 1) = \mathbb{P}(B = 1)$, can be estimated from the loan villages and (12) can be estimated with data from the no-loan villages. Equations (2)–(4) then imply that we can identify three average treatment effects of immediate interest:

$$\begin{aligned}\mathbb{E}(\Delta_G Q) &= \mathbb{E}(\Delta_G Q|L = 0) \\ \mathbb{E}(\Delta_G Q|B = 0) &= \mathbb{E}(\Delta_G Q|B = 0, L = 1) \\ \mathbb{E}(\Delta_G Q|B = 1) &= \frac{\mathbb{E}(\Delta_G Q) - \mathbb{E}(\Delta_G Q|B = 0)}{\mathbb{P}(B = 1)} + \mathbb{E}(\Delta_G Q|B = 0)\end{aligned}\tag{13}$$

These are the average effects of receiving a grant (without a loan) amongst the general population, amongst those who would not borrow if they were in a loan village, and amongst those who would borrow if they were in a loan village. We provide estimates of these three in section 3.2.

Our second goal is to identify $E(\Delta_B Q|B(1) = 1)$. We have already noted that (2), (3) and (4) imply that $E(Q(0,0)|B(1) = 0)$ is identified from data on the profits of non-borrowers who do not receive a grant in loan villages, $E(Q(1,0)|B(1) = 1)$ is identified from data on the profits of borrowers in loan villages, and that $E(Q(0,0))$ is identified from average profit of those who do not receive a grant in no-loan villages. Moreover, from (2), data from the loan villages identifies $\mathbb{P}(B = 1|L = 1) = \mathbb{P}(B(1) = 1)$. Then, in parallel with (13)

$$\mathbb{E}(Q(0,0)|B = 1) = \frac{\mathbb{E}(Q(0,0)) - \mathbb{E}(Q(0,0)|B = 0)}{\mathbb{P}(B = 1)} + \mathbb{E}(Q(0,0)|B = 0)\tag{14}$$

(14) and (2) imply that we can identify the average treatment effect on the treated of borrowing:

$$\begin{aligned}\mathbb{E}(Q(1,0)|L = 1, B = 1) - \mathbb{E}(Q(0,0)|B(1) = 1) \\ &= \mathbb{E}(Q(1,0)|B(1) = 1) - \mathbb{E}(Q(0,0)|B(1) = 1) \\ &= \mathbb{E}(\Delta_B Q|B(1) = 1).\end{aligned}\tag{15}$$

Section 5 presents intent-to-treat (ITT) treatment effects of residing in a loan village.

Note that we needed no assumption about whether farmers make the same investment decisions with a loan than with a grant in order to identify (15) and (13). We can test whether loan recipients are those with high returns to grants (whether or not they get large returns to the loans). We will, however, discuss the possibility that investment decisions could differ in section 4, when we discuss whether we observe an overall efficient allocation of capital.

2.3 Data

A baseline survey was conducted in January–May 2010. A first follow-up survey was conducted after the first year of treatment and the conclusion of the 2010 agricultural season¹⁵ in January–May 2011; a second follow-up survey was conducted after the second year of treatment and the conclusion of the 2011 agricultural season in January–May 2012; and a third follow-up survey was conducted seven years after the initial grant distribution in January–May 2017. The four rounds used similar survey instruments, which covered a large set of household characteristics and socioeconomic variables, with a strong focus on agricultural data including cultivated area, input use and production output at both the individual and the household level.

Throughout the paper we refer to “gross profit” as a key outcome variable. We do not have a complete measure of profits. Gross profit is the value of agricultural output net of most, but not all, expenses. Specifically, gross profit is the value of harvest (whether sold, stored or consumed) minus the cost of fertilizer, manure, herbicide, insecticide, hired labor, cart and traction animal expenses (rental or maintenance), and seed expenses (although valuing last year’s seeds at zero). We do not subtract either the value of unpaid labor (own, family or other) or the implicit rental value of land used, because both the labor and land markets are too thin to provide reliable guidance on these values. We will, however, examine the use of these inputs directly.

We also collected data on food and non-food expenses of the household as well as on financial activities (formal and informal loans and savings) and livestock holdings.¹⁶ The food expenditure module asked about consumption of over 50 food items over the previous seven days. We calculate prices using village-level reports in all sample villages. We use these sample-wide prices to convert consumption of all items into expenditure. It is important to note that there is a lot of consumption seasonality in Mali (Beaman, Karlan, & Thuysbaert, 2014). Our measure of food expenditure reflects consumption in the post-harvest season only.

2.4 Randomization Balance Check and Attrition

We conduct several tests to verify that we cannot reject the orthogonality of treatment assignment to baseline characteristics and attrition. Appendix Table 1 examines baseline

¹⁵ We also conducted an “input survey” on a sub-sample of the sample frame right after planting in the first year (September-October 2010), in order to collect more accurate data on inputs such as seeds, fertilizer and other chemicals, labor and equipment use. This input survey covered a randomly selected two-thirds of our study villages (133 villages) and randomly selected half of the households (stratifying by treatment status) to obtain a sub-sample of 2,400 households. We use the input survey if conducted, and we use the end of season survey if not. We also control for timing of the collection of the data in all relevant specifications.

¹⁶ The survey instruments are all available upon request.

characteristics across three comparisons: (i) loan to no-loan villages; (ii) grant to no-grant households in no-loan villages; and (iii) grant to no-grant households in loan villages. Few covariates are individually statistically significantly different across the three comparisons, and an aggregate test in which we regress assignment to treatment on the set of 11 covariates fails to reject orthogonality for each of the three comparisons (p-value of 0.26, 0.91 and 0.67, respectively, reported at the bottom of the table).

Our attrition rate is low at approximately one percent each round.¹⁷

3 Selection into loans and the return to cash grants

3.1 Observable characteristics of borrowers versus non-borrowers in loan villages

Take-up of the loans, determined by matching names from administrative records of Soro with our sample, was 21% in the first agricultural season (2010–11) and 22% in the second (2011–2012). Despite the similarity in overall take-up numbers, there is turnover in clients. About 65% of clients who borrowed in year 1 took out another loan in year 2. This overall take-up figure is similar to other evaluations of group-based microcredit focusing on small enterprise (for analysis of randomized evaluations of group-based microcredit, see Angelucci, Karlan, & Zinman, 2015; Attanasio, Augsburg, De Haas, Fitzsimons, & Harmgart, 2015; Banerjee, Duflo, Glennerster, & Kinnan, 2015; and for a summary discussion of these studies, see Banerjee, Karlan, & Zinman, 2015; Crépon, Devoto, Duflo, & Pariente, 2015; Tarozzi, Desai, & Johnson, 2015).

Table 1 provides descriptive statistics from the baseline on households who choose to take out loans in loan villages, compared to non-clients in those villages. We provide information on the household as a whole, as well as the primary female respondent and primary male respondent. There is a striking pattern of selection into loan take-up: households that invest more in agriculture, and have higher agricultural output and gross profits are more likely to take out a loan. Borrowers also have more agricultural assets and livestock. Figure 2 demonstrates that this holds across the whole distribution. Women in households who borrow are also more likely to own a business and are more “empowered” by three metrics: they have higher intra-household

¹⁷ Despite the low attrition rate, we report differential attrition tests in Appendix Table 2. We compare the same groups as in Table A1, from baseline to the first follow-up and to the endline. For each of the three comparisons, we fail to reject that attrition rates are on average the same in the compared groups for both follow-up years. In a regression of attrition on the nine covariates, treatment status, and the interaction of nine covariates and treatment status, we fail to reject orthogonality for all six regressions (results on bottom row of Appendix Table 2).

decision-making power, are more socially integrated, and are more engaged in community decisions.¹⁸ Households that borrow also have higher consumption at baseline than non-clients.

3.2 Experimental results on returns to grants in loan and no-loan villages

Next, we present the estimated returns to capital of receiving a grant amongst the general population, amongst those who would not borrow if they were in a loan village, and amongst those who would borrow if they were in a loan village (equation 13). To isolate the role of selection into loans, we focus on the first year of the experiment.¹⁹ Table 2 shows the estimates from the following regression using the first follow-up data on farm investments and output.

$$Y_{ijt} = \alpha + \beta_1 grant_i + \beta_2 grant_i \cdot loan_j + X_{ijt}\pi + \lambda_j + \epsilon_{ijt} \quad (16)$$

where $grant_i$ indicates individual i received a grant in May–June 2010, and $loan_j$ indicates that the MFI offered loans in village j . We include additional baseline controls (X), which include the baseline value of the dependent variable y_0 ²⁰ plus its interaction with village type (loan village / no-loan village) and the baseline variables used in the re-randomization routine (listed in the notes of table 2). λ_j are village fixed effects. β_1 and β_2 are the primary coefficients of interest. β_1 is the effect of the cash grant on the outcome Y_{ijt} in the no-loan villages, i.e., the average effect of the cash grant in a sample of the full population. β_2 shows the differential impact of the grant on the outcome Y_{ijt} for the loan village households that did not borrow. Standard errors are clustered at the village level. We also provide randomization inference p-values (Young, 2019)

¹⁸ All three of these variables are indices, normalized by the no-grant households in no-loan villages. The household decision-making index includes questions on how much influence the woman has on decisions in the following domains: food for the household, children’s schooling expenses, their own health, her own travel within the village, and economic activities such as fertilizer purchases and raw materials for small business activities. The community action index includes questions on the frequency she speaks with different village leaders, and different types of participation in village meetings and activities. The social capital index includes questions about seven other randomly selected community members from our sample and whether the respondent knows the person, is in the same organization, would engage in informal risk-sharing and transfers with the person, and topics of their discussions (if any).

¹⁹ The second-year data is more difficult to interpret. In loan villages, a different set of households borrowed in year 2 than in year 1. In particular, we observe a positive, though modest, treatment effect of receiving a grant on taking out a loan in year 2. The impact of the grant in year 2 in loan villages is therefore a combination of mechanisms and does not isolate selection. The results in year 2 are shown in section 6.1.

²⁰ In cases where the observation is missing a baseline value, we instead give the lagged variable a value of -9 and also include an indicator for a missing value.

that account for both the re-randomization routine used to assign treatment status and multiple comparisons within families of outcomes (details discussed in table notes).

Table 2 shows the estimates from this regression for a variety of cultivation outcomes (inputs along with harvest output and gross profits), and Table 3 shows the analogous estimates for non-cultivation outcomes such as livestock, enterprise, consumption, and female empowerment.

3.2.1 Agriculture

Columns (1)–(8) of Table 2 examine agricultural inputs and crop choice. We first focus on the first row of coefficients, β_1 , which captures the impact of the grant in no-loan villages. We find that households who received a grant in no-loan villages cultivated more land than those who did not (0.18 ha, $se=0.07$). This is approximately an 8.7% increase (control mean=2.07) compared to households who did not receive a grant in no-loan villages. Households also allocate their land to a different crop mix: column (2) shows that 0.07 more hectares ($se=0.02$) are dedicated to growing rice and groundnuts, which are cash crops in the area. The grant also induced an increased use of the plough (6 percentage points, $se=1$), the quantity of seeds used (5 kg, $se=2.1$), and in hired labor days (2.7 days, $se=0.8$). While 2.7 days over the entire agricultural season is a small number, these households use little hired labor: the mean in the control in 2011 was only 17 days. We observe no change in family labor. Fertilizer and other chemical inputs increased by 18% (US\$21, $se=6$). The agricultural inputs and crop choice variables in columns (1)–(7) are grouped together as a family of outcomes for the randomization- c p values (Young 2019). The adjusted p -values are qualitatively similar to our simple tests. Moreover, the omnibus test indicates a statistically significant ($p<.001$) experimental effect.

The grants led to an overall increase in agricultural investment: column (8) shows that measured input expenses increased by US\$30 ($se=8$).²¹ Columns (9)–(10) report statistically significant and economically meaningful increases in output and gross profits: output increased by US\$66 ($se=19$) and gross profits increased by US\$39 ($se=16$), equivalent to 13% and 12% increases, respectively. Overall, we see statistically significant increases in investments and ultimately gross profits from relaxing capital constraints.²²

²¹ The value of land and the shadow wage of family labor cannot be estimated given the extremely thin land and labor market in this area. In addition, only seeds that were purchased in the market and rental costs of a plough are included in total input expenses. The value of seeds used from the previous year's harvest and the cost of using their own plough are also not included. See the notes in table 2 for more details.

²² We are not estimating the marginal product of capital as in de Mel et al. (2008) but instead the "total return to capital"—i.e., cash. Beaman et al. (2013) shows that labor inputs adjust along with agricultural inputs, making it impossible to separate the returns to capital from the returns to labor without an additional instrument for labor

Critically, the coefficient on Grant * Loan village (β_2) demonstrates striking heterogeneity in the returns to the cash grant between no-loan and loan villages. The β_2 coefficient shows that the selected sample of households who did not take out a loan do not experience the same positive returns when capital constraints are relaxed.

Column (1) shows that households in loan villages who did not take out a loan did not increase the amount of land they cultivated when randomly selected to receive a grant ($\beta_2 = -0.16$ ha, $se=0.10$ and the p-value of the test that the sum of β_1 and β_2 is zero is 0.80). The interaction terms for family labor and fertilizer/other chemical expenses are also negative (-6.9 days, $se=6.5$ and -US\$15, $se=9$, respectively). Households who received grants in loan villages did seem to increase some inputs, such as quantity of seeds and hired labor, although neither is statistically significant as shown in columns (2)–(6). Column (8) shows that total input expenses in loan villages increase in response to the grant by about US\$20 (p-value=0.02), which is not statistically different from the estimate in no-loan villages of US\$29. Note, however, that the inputs that are measured with the most precision—fertilizer and chemical expenses in column (7)—demonstrate a statistically significant difference in the impact of the grant on investment choices between loan and no-loan villages.

However, even though we observe increased inputs for the grant recipients in loan villages, we see no corresponding increase in either agricultural output or in gross profits. The β_2 interaction coefficient for output is similar in magnitude to β_1 but *negative* (-US\$41, $se=28$), offsetting the increase in output in no-loan villages (US\$66, $se=19$). The test that the sum of the two coefficients is different from zero is not rejected ($p=0.23$), indicating that the (intentionally) selected sample did not experience a statistically significant increase in output when given a grant. Similarly, the total effect on gross profits in loan villages is essentially zero (US\$1.04), which is not significantly different from zero ($p=0.95$) and fairly precisely measured. Thus while there is some evidence that households who did not take out loans used some of the grant to increase agricultural inputs, there is—in stark contrast to the random sample of households in no-loan villages—no evidence of increases in either agricultural output or gross profits.

These estimates imply that there is a great deal of heterogeneity in marginal returns to relaxing capital constraints across farmers, and that those who borrow are disproportionately those with high marginal returns. The return to the grant implied for would-be borrowers in no-loan villages

inputs. We are therefore capturing the total change in profits and investment behavior when capital constraints are relaxed.

is US\$131 (se=68) in additional gross profits per US\$100 of grant.²³ In contrast, the return for non-borrowers is close to zero.

The analysis indicates that households who do not borrow are those without high returns in agriculture to cash transfers. In contrast to the literature on health products, where much of the evidence points towards limited screening benefits from cost sharing (Ashraf, Berry, & Shapiro, 2010; Cohen & Dupas, 2010; Tarozzi et al., 2014), we find that the repayment liability leads to lower return households being screened out. Appendix A1 explores this in depth, and demonstrates that we are unable to predict either the returns to the grants or the heterogeneity in returns using baseline characteristics (see Table A5).

3.2.2 Other outcomes

Table 3 shows the estimates of equation (14) on non-agricultural outcomes. The most striking results are in columns (1) and (2): grant-recipient households in no-loan villages are more likely to own livestock (11 percentage points, se=1), and there is a large (US\$166, se=71) increase in the value of total livestock compared to no-grant households. This represents a 14% increase in the value of household livestock, and is slightly larger than the value of the grant itself. Recall we saw in Table 2 that households also spent an extra US\$30 on cultivation investments. The livestock value is measured several months after harvest; these results could indicate that households moved some of their additional farming profits into livestock post-harvest, or they may reflect measurement challenges.²⁴ We also see that the grant increased the likelihood in no-loan villages that a recipient household had a small enterprise (column (3); +4 percentage points, se=2, control group mean =0.83). Grant recipient households also consumed more, including 5.7% more food (column (4); US\$0.34 per day in adult equivalency, se=0.14, control group mean = 5.96) and 5.8% in non-food expenditures (column (5); US\$2.53 per month, se=1.39, control group mean = 43.81). Columns (6)–(8) show no statistically significant main effect of the grant on whether the household has any financial savings, education expenses or medical expenses.

The investment and spending patterns among grant recipient households in loan villages for the most part echo those described above in no-loan villages. Column (1) shows that while grant

²³ Calculated as $(\beta_1 - 0.79(\beta_1 + \beta_2))/(0.21) * (100/140)$. The average return in the entire village is β_1 . The take-up rate of loans is 21%, so 79% of households in the village would be non-borrowers and would have earned a return of $(\beta_1 + \beta_2)$. The return is then scaled to be per US\$100, so we divide by the grant size of US\$140.

²⁴ We may over-value recently purchased livestock. At the household level, we collected data on the quantity of animals. We use village-level reports of livestock prices to value livestock quantities for all households. Therefore, if recently purchased livestock are younger or smaller in treatment household, leading to a large estimated treatment effect.

recipients in loan villages were overall more likely to own livestock than their control counterparts, the magnitude of the effect is smaller than in the no-loan villages (interaction term is -4 percentage points, $se=2$). The remainder of the outcomes however show few differences.²⁵

Taken together, Table 3 shows that the grants benefited households in a variety of ways. However, we have no strong evidence that households in loan villages, who did not experience higher agricultural output and gross profits than in no-loan villages, used their grants to invest in alternative activities that offered higher-returns than cultivation.

3.2.3 Robustness

Timing of delivery of grants

One concern about our interpretation of the results is a timing issue: households received grants in loan villages on average 20 days later than in no-loan villages because of delays in the administration of the loans. If farmers in loan villages received grants too late in the agricultural cycle to make productive investments, we would erroneously conclude that there is positive selection into agricultural loans, since we would observe more investments and returns in no-loan villages than in loan villages. This is particularly a concern since we observe farmers increase the amount of land they farm, which is a decision which occurs very early in the agricultural cycle. In Appendix Table 3, we look at land cultivated (i.e., an investment decision made early in the process) and an index of all the agricultural outcomes and find no relationship with the timing of the grant among the grant-recipient households in no-loan villages. We look at two main specifications: one in which we include the date the grant was received linearly and squared, and a second in which the sample is split into the first half of the grant period and the second half (since most of the grants in the loan-available villages were distributed in the second half). In both cases we control for whether this was the team's first visit to the village (rather than a revisit).²⁶

²⁵ The only outcome which suggests potential heterogeneity in behavior between loan and no-loan villages is medical expenses, in column (9). Medical expenses (in the last 30 days) are marginally significantly higher in loan grant households (US\$5.01, $se=2.55$), since medical expenses may have declined (-US\$2.58, $se=1.87$) among grant recipients in no-loan villages. The total effect in loan villages is not statistically different from zero ($p=0.16$). This is a difficult outcome to interpret because (i) having more resources could mean a household is more able to treat illnesses, but (ii) having more resources could lead to higher preventative care, which should lower total medical expenses.

²⁶ Households who are revisited are those who were not available during the first visit to the village. They may be systematically different than households who are reached during a first visit.

Spillovers

It is possible that households that received neither grants nor loans were *indirectly* affected by the study interventions. Spillovers could be either positive (if grants or loans were shared) or negative (through general equilibrium effects on locally determined prices or competition over land). We do not have a perfect way to estimate such spillovers. We do, however, have data from an additional 69 villages in the same administrative units (cercles) as our study villages.²⁷ Appendix Table 4 shows that no-grant households in no-loan villages had similar agricultural practices to households in villages where we did no intervention. There are no statistically significant differences in hectares of land cultivated, suggesting that the increase in land cultivated among grant recipients was not zero-sum with households who did not get a grant. We also observe no statistically significant change in land cultivated with rice or groundnuts (column (2)). This is important since land used to grow rice, which needs to be in a flood plain, is more constrained than other types of land and is thus most likely to be crowded out by treated households. There are also no statistically significant differences in total input expenses, value of the harvest, and gross profits (columns (6)–(8)). The number of hired labor days (column (4)) is the one statistically significant difference: non-grant recipients in no-loan villages hired more labor by 3.5 laborer days (se=1.4). While this is precisely estimated and a point estimate comparable to main treatment effect in Table 2, recall that this is four man-days over the entire course of the agricultural season and therefore unlikely to have affected total output and gross profits. Column (9) suggests no statistically significant changes in equilibrium prices. This makes sense since villages in Mali are small. Households engage in market activities in local weekly markets, which bring multiple villages together (Ellis & Hine, 1998).

We note that this analysis cannot speak directly to the possibility of spillovers in loan villages. The dynamics of sharing a grant with others in a village in which loans are available may differ, and the direction is difficult to predict. There may be pressure to share or hide “free” money when others recently borrowed; on the other hand, those who needed capital would have received a grant and therefore grant recipients may share less.

²⁷ Our partner organization would only commit to not enter 110 villages, which serve as our no-loan villages. The villages we use as no-intervention villages were villages not used for the primary study, but the selection of villages into the experimental study sample was not explicitly randomized. For example, the no-intervention villages have larger average population size but fewer children per household than study villages. Also Soro Yiriswaso may have offered loans in up to 15 of the 69 villages in year 1. Removing those 15 villages leaves Appendix Table 4 qualitatively unchanged.

4 Efficient selection?

The experimental design provided us with a transparent method for showing that the impact of the grants on gross profits in the random sample of households is greater than their impact in the selected sample of non-borrowers. Soro loans are being directed towards households that—on average—have higher rates of return to grants. However, this observation raises an important question. Are the loans successfully allocated to all women with high return investment opportunities? There may be potential borrowers with projects that could generate high returns, but who do not receive loans. Concerns about the likelihood or costs of possible default, or about the risk of high expected return projects may mean that the loans are not reaching all farmers with high marginal returns.²⁸

We identify a set of potential borrowers with high marginal returns that do not borrow. Farmers with low baseline gross profits or low baseline consumption are less likely to borrow, conditional on their marginal returns to a grant, than farmers with higher baseline profits or consumption.

Credit transactions require a credible commitment to repay the loan; these poorer potential borrowers with high marginal returns may be unable or unwilling to make a credible repayment commitment. Alternatively, it may be that risk aversion may deter poorer farmers with high expected return projects from borrowing. These frictions may require a wedge (τ_i) between farmer i 's marginal return to a loan ($\Delta_L Q_i$) and the gross cost of funds to the lender (ρ) before i can borrow, which means that a farmer borrows if and only if $\Delta_L Q_i \geq \rho + \tau_i$. In the absence of repayment concerns, risk aversion or other transaction costs, $\tau_i = 0$ for all farmers, and all farmers with marginal returns to loans higher than ρ are offered and accept loans. However, any of these frictions could generate positive wedges between $\Delta_L Q_i$ and ρ for some farmers. For example, in an environment in which collateral is used to encourage repayment, τ_i might be (negatively) related to farmer i 's wealth (or her holdings of a particular asset). If insurance is incomplete, τ_i might be positively associated with a farmer's risk aversion.

A particularly simple example is provided by a limited liability constraint. Suppose a loan is repaid only to the extent that borrower income net of repayments is no lower than some minimum level \underline{c} . If the microcredit institution lends at an interest rate equal to its cost of funds, and must break even, then it will lend only to those farmers who can repay in full. In this case $\tau_i = \underline{c} - Q_i(0,0)$.²⁹

²⁸ The same concerns could mean that loans would not be used for the same high return projects as grants. This possibility is examined in Section 5.

²⁹ The limited liability constraint is $Q_i(1,0) - \rho \geq \underline{c}$. Substitute $\Delta_L Q_i \equiv Q_i(1,0) - Q_i(0,0)$. Appendix A3 describes the constrained efficient allocation of loans by a provider faced with a zero-profit constraint in a limited liability

In the left panel of Figure 3, the horizontal curve E defines the boundary in $(Q(0,0), \Delta_L Q)$ between those who borrow and those who do not in an efficient allocation assigning credit to all farmers with a sufficiently profitable investment opportunity. Farmer i with values of $(Q_i(0,0), \Delta_L Q_i) \in NB$ does not borrow because her returns are too low. In the right panel, the curve C defines the boundary in an allocation constrained by limited liability concerns. The set of values of $(Q_i(0,0), \Delta_L Q_i) \in NB$ such that a farmer does not borrow is expanded because of the limited liability constraint, and there are high return farmers who inefficiently do not borrow.

Empirical evidence of inefficient selection

We can now consider the consequences of this constraint to the efficient allocation of loans for the observed distribution of gross profits. In the no-loan villages, where grants were given to a random sample of the population, we observe $Q_i(0,0) \equiv Q_i^0$ and $Q_j(0,1) \equiv Q_j^G$ for farmers i and j randomly selected into the no grant and grant treatment groups, respectively. Recall that $\Delta_G Q_i \equiv Q_i^G - Q_i^0$ is the return to the cash grant. Let $h(\Delta_G Q, Q_i^0)$ denote the joint density of $\Delta_G Q$ and Q_i^0 in the population of our study area. For the purposes of this section, we suppose that $\Delta_G Q$ is a monotonic function of $\Delta_L Q$: the same farmers who have a high marginal return to the grant have a high marginal return to a loan. As noted above, we examine the relationship between these marginal returns in section 5. Given our randomization, the distributions of Q_j^G and Q_i^0 simply reflect draws from the full density $h(\cdot)$. The left panel of Figure 4 depicts these distributions empirically.³⁰ As can be anticipated from our preceding results, the distribution of Q_j^G lies to the right of that of Q_i^0 over virtually the whole range.

In the loan villages, grants were given to a random sample of non-borrowers. Suppose that selection into borrowing is constrained by repayment concerns, that is, by C in Figure 3. In this case, the joint density of $\Delta_L Q$ and Q^0 in the population of non-borrowers is the truncated probability distribution

$$h^{NB}(\Delta_G Q, Q^0) = \frac{h(\Delta_G Q, Q^0)}{\text{prob}((\Delta_L Q, Q^0) \in NB)} \quad (17)$$

with support $(\Delta_L Q, Q^0) \in NB$. As can be seen in the right panel of Figure 4, the endowments of the approximately 80% of the population who do not borrow differ from the overall population

environment, without the restriction that the interest rate is fixed at ρ , illustrating the general result that potential default can generate farmer-specific wedges between marginal returns and the cost of funds.

³⁰ Note that this is the same sample as we use in table 2, and therefore continues to exclude households who borrowed in loan villages.

in two ways. First is the presence of a large fraction of non-borrowers with relatively high gross profits (>\$500), but approximately zero marginal return from the grant. This pattern is consistent with an efficient allocation: farmers who have low returns to capital do not borrow and therefore show up in this sample. Second is the presence of a large fraction of non-borrowers with high marginal productivity but low gross profits. This feature corresponds to exclusion of potential defaulters. We infer that the realizations of output are determined by an allocation constrained by repayment concerns so that non-borrower endowments are drawn from $h^{NB}(\Delta_L Q, Q^0)$. This suggests there are some high return potential borrowers who do not receive capital, highlighting imperfect efficiency.

Correlations between observable characteristics of borrowers and non-borrowers and the return to grants are also informative of the nature of the selection process. We saw in Table 1 that there are observable characteristics that are strongly correlated with loan take-up. Consider any such attribute, Z , that we *a priori* expect might be negatively correlated with farmer-specific borrowing frictions τ_i . For example, baseline gross profits would be one such attribute. In Table 4, we report the results of estimating

$$Y_{ijt} = \alpha + \beta_1 grant_i + \beta_2 grant_i \cdot loan_j + \gamma_1 grant_i \cdot Z_{ijt} + \gamma_2 Z_{ijt} + \delta_1 grant_i \cdot Z_{ijt} \cdot loan_j + X_{ijt}\pi + \lambda_j + \epsilon_{ijt} \quad (18)$$

where we use a specification that includes the interaction term Grant * Z * Loan village. This additional interaction permits us to examine whether the correlation between Z and the marginal return to the grant is different for the general population (γ_1) than for a selected population of non-borrowers ($\gamma_1 + \delta_1$). This helps illuminate whether the underlying allocation mechanism is efficient or characterized by farmer-specific borrowing frictions. The lower τ_i associated with the higher value of Z reduces the likelihood that the farmer has been screened out of borrowing by concerns of default. Non-borrowers with higher values of Z are therefore more likely to have selected out of borrowing because they have low marginal productivity. Hence, among the population of non-borrowers in loan villages, higher values of Z are associated with lower values of $\Delta_G Q$, relative to the association in the population in general.

Column (1) of Table 4 examines the association between baseline gross profits and the marginal return to the grant in the overall population and in the selected sample of non-borrowers. In accord with borrowing frictions that decline with baseline profits, households in loan villages have a statistically significantly more negative correlation between baseline gross profits and the return to a grant than households in the overall population (Grant * Baseline gross profits * Loan village: -US\$0.18, se=0.07). This reflects a constraint to the allocation of loans to the most productive farmers. Differential wedges between the marginal productivity of a loan and the cost

of funds—from repayment concerns, risk aversion or other farmer-specific frictions—generate a positive correlation between baseline gross profits and loan take-up.

In columns (2)–(4), we report the estimates of equation (19) for three additional characteristics of households that are positively associated with loan take-up and plausibly farmer-specific borrowing frictions: baseline value of livestock holdings, baseline food consumption per capita (in USD), and baseline non-food expenditure per capita (in USD). Column (2) reports the results for baseline value of livestock holdings. The differential returns to the grant for the general population (those in no-loan villages) and non-borrowers (those in loan villages) does not differ for those with higher versus lower baseline livestock holdings (column (2), -US\$0.015, se=0.013). Thus, this provides no evidence in support of the hypothesis that farmers with low livestock holdings are subject to higher borrowing frictions. Next we examine the same but for food consumption (column (3)) and non-food expenditures (column (4)), hypothesizing that these may be strongly positively correlated with a household’s permanent income (and hence negatively with τ_i). Here we do find statistically significant differences, in which the differential returns to the grant for the general population relative to the non-borrowers is lower for those with both higher food consumption (-US\$23, se=6) and non-food consumption (-US\$1.61, se=0.61).

To capture a multifaceted Z, we exploit a machine learning algorithm to estimate heterogeneity in treatment effects (Athey & Imbens, 2016; Athey et al., 2019; Wager & Athey, 2018). The causal tree algorithm of Athey and Imbens (2016) extends the basic intuition of decision trees like those used in random forests by selecting splits in order to maximize heterogeneity in treatment effects across leaves (less a penalty for the variance of treatment and control outcomes in each leaf). This approach, based on ensemble of decision trees, provides estimates of conditional average treatment effects (CATEs) for each household. We implement the generalized random forest method (Athey et al., 2019; Athey & Wager, 2019; Wager & Athey, 2018) using the R package *grf* version 0.10.4 (Athey et al., 2019; Tibshirani et al., 2018). See Appendix A2 for details on implementation of the causal forest methodology.

We can estimate predicted treatment effects (CATEs) using either an algorithm trained on no-loan villages only or on loan villages only. Appendix A1 shows that following the method by Chernozhukov et al. (2018), there is robust evidence of heterogeneity in grant treatment effects among the selected sample in loan villages, and little evidence of observable heterogeneity in no-loan villages. Table 5 explores whether the baseline characteristics which are associated with high CATEs are the same in both models. Table 5, column (1) shows that in the general population of no-loan villages, households with high CATEs have *higher* baseline gross profits, more food and non-food consumption, more livestock, and more landholdings. In contrast, column (2) shows that in the sample of non-borrowers in loan villages, households with high CATEs have lower

baseline gross profits, lower baseline food consumption and non-food expenditure, lower livestock values and smaller land holdings.

The comparison between columns (1) and (2) is striking as six out of the eight characteristics have the opposite sign in their correlation with predicted treatment effects in the two models. This is further evidence that the allocation of loans is not based on marginal productivity alone. Among the selected sample in the loan villages who did not borrow, we see that those who are less poor—as proxied by having higher food and non-food consumption—have lower returns. These are households that would be less likely to default, or to be less risk averse—they have a lower τ . They are not allocated loans because they have low returns. In the full sample in column (1), we see a positive correlation between baseline food and non-food consumption and predicted returns. This is consistent with Table 1, where borrowers tended to be less poor than non-borrowers. To square this with column (2), the model would suggest that the less poor households with expected high returns borrow, and left the sample that we used to train the model in the loan villages. Those that remain are the less poor households with low anticipated returns, generating the negative correlation in column (2).

Figure 5 demonstrates visually the effects of constraints based on repayment concerns on the joint distribution of baseline gross profits and the return to grants. The x-axis is the quantiles of baseline gross profits, while the y-axis is the quantiles of the predicted treatment effects (CATEs). Figure 5a reports the results of the causal forest trained on and estimated in the no-loan village sample. The highest density of observations is in the upper right, and there is an apparent positive correlation between baseline gross profits and the estimated CATE of a grant. Figure 5b reports the results of the causal forest trained on and estimated in the sample of non-borrowers in the loan villages. The high-baseline profit and high CATE quadrant of the population is much less represented: these are households that demand loans and are able to borrow.

Average agricultural returns to the grants for non-borrowers in loan villages are zero, as shown in column (10) of Table 2, while they are on average high for the random sample in no-loan villages. However, Table 5 demonstrates that average agricultural returns to grants for non-borrowers with low values of baseline profits, baseline food consumption, or baseline non-food consumption are large. Indeed, Appendix Table 7 shows that among non-borrowers in the first tercile of the distribution of baseline food and non-food consumption, average returns to the grant are at least as high as the average returns in no-loan villages. We refer to these households as “the most poor”. Thus, it appears that among the most poor, there are households with high returns to grants that are not borrowers, implying an inefficient allocation of loans.

Would these most poor farmers use the loans in a similar way to the grants? We cannot observe the returns to the grant for any individual farmer, of course. But we do observe the *ex-post* gross

profits of grant recipients. Among the most poor households, it is not possible to reject the hypothesis that the distribution of profits among those who receive grants in the loan villages (the non-borrowers) is the same as for those who receive grants in the no-loan villages. However, this may be due to low power. Among these most poor households, the median, second tercile and third quartile of the distributions of profits among those nonborrowers who receive grants is greater than or equal to those of the distribution of profits among grant recipients in no-loan villages (although none of the differences is statistically significant).³¹ The distribution of observed profits for grant recipients among the most poor, therefore, is consistent with the existence of high return households among the non-borrowers. There is no evidence of selection of high return farmers into borrowing amongst the most poor; all of the selection is occurring among the less poor. There are farmers with high returns who do not borrow.

5 Impact of the loans

We also show our estimates of the intent-to-treat (ITT) effects of being offered an agricultural loan on the same set of outcomes already discussed in section 3. In this analysis, we exclude all grant recipients, from both loan and ineligible villages. We use the following specification:

$$Y_{ijt} = \alpha + \beta_1 loan_j \cdot I\{t = 2011\} + \beta_2 loan_j \cdot I\{t = 2012\} + X_{ijt}\pi + \epsilon_{ijt} \quad (19)$$

where (X) includes the baseline value of the dependent variable y_0 , *cercle* (an administrative unit above a village and below a region) fixed effects, and the village stratification controls listed in the notes of the Table 2. The specification uses probability weights to account for the sampling strategy, which depends on take-up in the loan villages. See notes in table 6 for details.

Table 6 and Appendix Table 8 show the ITT estimates for agricultural outcomes and broader outcomes, respectively. In Table 6, we observe an increase in input expenditures on family labor days (8.6, se=4.8) and in fertilizer and other chemicals expenses (US\$14, se=7); total input expenses rose by US\$20 (se=9) in villages offered loans. Land cultivated also increases but is not statistically significant at conventional levels (0.08 ha, se=0.06). The value of the harvest rose by US\$34 (se=20), but we do not measure a statistically significant increase in gross profits (US\$19, se=16).

Loans have to be repaid, while grants do not. Concerns about the costs of default or risk could deter borrowers from investing in the highest return activities; loan recipients to use loans differently from the way in which they use grants, and to realize different returns for loans than grants. The selection effect we have identified, in which women with high agricultural returns to

³¹ These results are available upon request.

grants are strongly selected into borrowing, may not imply that these same women have high agricultural returns to loans. We calculate the Treatment on the Treated estimates for year 1 for the sub-population who take up loans.³² Compared to the estimate of the impact of the grant from table 2, we do not reject the hypothesis that the per US\$100 dollar effects of grants and loans are the same for any of the agricultural outcomes.³³ Taken as a whole, the grants and loans are having similar effects on agricultural inputs and outcomes.³⁴

Appendix Table 8 demonstrates that overall, the microcredit agricultural loans did not have broad impacts beyond agriculture. We do not detect an impact on outcomes such as food and non-food consumption, whether the household has a small business, or educational expenses.³⁵ We observe a large but imprecisely estimated impact on livestock (columns (1)–(2)). We do find a statistically significant *reduction* in medical expenses (column (9), -US\$5.03, se=1.64). We are not, however, able to document any corresponding increase in preventative health care expenditures.

These results on the impact of loans stand in stark contrast to the recent experimental literature on the impact of entrepreneurially focused credit (see Angelucci et al., 2015; Attanasio et al., 2015; Augsburg, De Haas, Harmgart, & Meghir, 2015; Banerjee, Duflo, et al., 2015, and an overview in Banerjee, Karlan, et al., 2015; in contrast, Breza & Kinnan, 2018 finds noticeable general equilibrium effects as a consequence of a state-wide shutdown of the microcredit market; Crépon et al., 2015; Karlan & Zinman, 2011; Tarozzi et al., 2015). Analysis pooling these studies using a Bayesian hierarchical model, however, unravels evidence of positive treatment effect at higher quantiles, even though the average treatment effect is a fairly precise null (Meager, 2019, 2020). An earlier agricultural lending literature also documented institutional

³² See table notes of Table 6.

³³ The standard errors are calculated using a bootstrap routine: the difference in the impact of the grant and loan is estimated for 1,000 draws of households (with replacement), with probability weights for households calculated in each bootstrap sample for the loan impact estimation.

³⁴ Note that we do not remove the cost of the loan, i.e. interest payments, from gross profits. The true difference in take home profits between the grant and loan would be larger. We do not include the interest because the goal is to see if the behavior of farmers, in terms of investments and the associated agricultural output, differs between the grants and the loans. We see that there is no evidence that the fact that they must pay interest leads to different investment choices.

³⁵ Columns (9)–(11) of Appendix Table 8 further shows no detectable effect on women’s decision-making power within the household, women’s involvement in community decisions, or women’s social capital. This is similar to the existing evaluations of microcredit (one exception is Angelucci et al., 2015; finding no impact on these measures: Attanasio et al., 2015; Augsburg et al., 2015; Banerjee, Duflo, et al., 2015; Crépon et al., 2015). Soro Yiriwaso did not have any explicit component of the program emphasizing women’s empowerment.

failures, typically with high default rates (Adams, 1971; Adams et al., 1984), although a newer study in Zambia finds positive impacts from agricultural loans, similar to those found here (Fink et al., 2018).

The impact estimates are also promising from the perspective of the microcredit institution: repayment was 100%, and the retention to the following year (65%) is on par with typical client retention rates for sustainable, entrepreneurially focused microcredit operations.

6 Persistent effects of grants

We focus first on the impact of the grants in year 2 and then on the impact at the longer-term follow-up in year 7.

Agriculture

We observe a persistent increase in output and gross profits in the 2011–2012 agricultural season (year 2) from the grant given in 2010. In Panel A of Table 7, column (8) shows that output is higher in grant recipient households by US\$52 (se=23) and column (9) demonstrates that gross profit was higher by US\$49 (se=17). This is striking since we do not observe grant-recipient households spending more on inputs *that we can easily measure* in column (8) (US\$1.10, se=10.45). Recall that there are a number of inputs, such as land, seeds used from the previous year's harvest, and family labor, that we cannot value. Columns (2)–(4) provide evidence that grant recipients continued to make different investments than the control group. Grant recipients in no-loan villages planted 6.5% more land with rice and peanut crops in year 2. Rice and peanuts are high value crops. Grant recipients in no-loan villages were also 4.9% more likely to use a plough during land preparation (4 pp, se=1), and used 6.8% more seeds (6.1 kg, se=2.6).

We show the estimates of the interaction term of Grant * Loan village in year 2 in Table 7, but the interpretation of the results is challenging. In the second year of the experiment, the MFI offered loans again. Only about half of households who took out a loan in year 1 took out another loan. There were also households who did not borrow in year 1 who chose to borrow in year 2. Moreover, households who randomly received a loan in year 1 are more likely to receive a loan in year 2. With the caveats in mind, we see a similar negative interaction term on gross profits in column (10) of Panel A as in year 1 (-US\$40, se=24). The lower gross profits may be a result of higher input use: column (8) shows that, in loan villages, grant-recipient households spent more on agricultural inputs (US\$30, se=17) than control households in 2012.

Other outcomes

Appendix Table 9 shows the persistent impacts of the grant in year 2 on non-agricultural outcomes. Columns (1) and (2) demonstrate that grant-recipient households are more likely to

own livestock (9 percentage points, $se=2$) and continue to hold more livestock assets (US\$184, $se=102$) than control households in no-loan villages. They are also more likely to own a business (3 percentage points, $se=1$).³⁶ There is no significant increase in food consumption in year 2 (US\$0.24, $se=0.19$) but monthly non-food expenditure does increase (US\$3.89, $se=2.13$). Households are also more likely to have financial savings (3.3 percentage points, $se=1.9$). Columns (7)-(8) show that there continues to be no measurable impact on educational expenses (US\$0.39, $se=3.76$), or medical expenses (-US\$0.72, $se=1.82$).

Appendix Table 9 also shows that, similar to year 1, there is no evidence of households in no-loan villages using grants differently to those in loan villages across this set of non-agricultural outcomes (livestock ownership, owning a small business, and consumption) in year 2.

Longer-term follow-up

In 2017, almost seven years after the grants were distributed, we conducted another round of data collection, interviewing 5,560 of the original sample households. Panel B of Table 7 shows no evidence of a persistent effect of the grant on the key agricultural outcomes analyzed in the paper. The time period between 2012 and 2017 was a tumultuous time in Mali. There was a military coup in March 2012, followed by a French military intervention in the north of the country until 2014 (all of which were factors in why there was a large gap in our field work between the second and seven year follow-ups). Second, unrelated to the political instability, there was an expansion in cotton cultivation in the Segou region of Mali. From 2007 to 2010, it is estimated that between 200 and 244 million tonnes of cotton were produced per year. In 2017, that figure had risen to 703 million tonnes (USAID, 2018). The increase largely came from an increase in the land dedicated to cotton cultivation. The state-owned Malian Textile Development Company (CMDT), which was re-structured starting in late 2010, provides fertilizer and credit to cotton farmers. This change in cultivation patterns could easily wash out any long-term benefits from a single cash transfer many years prior.

Note that we did not analyze if there is a difference in agricultural outcomes between loan and no-loan villages since our partner organization Soro was unable to provide any information on whether loans were disbursed in the treatment and/or control villages between 2012 and 2017.

7 Conclusion

Capital constraints are binding for at least some farmers in Southern Mali, and agricultural lending with balloon payments (i.e., with cash flows matched to those of the intended productive activity) can increase investments in agriculture. This is an important policy lesson since the

³⁶ In results available from the authors, business profits increase by 18% (US\$41, $se=19$) in year 2.

majority of microcredit has focused on small enterprise lending, and the typical microcredit loan contract—where clients must start repayment after a few weeks—is ill-suited to agriculture. In Mali, for example, Soro Yiriwaso is among very few microcredit organizations with a product specially designed for agriculture, despite the fact that the vast majority of households in rural Mali depend on agriculture for a sizeable part of their livelihood. Given the lackluster average estimated impact of entrepreneurial microcredit (Banerjee, Karlan, et al., 2015; Meager, 2019), our results could serve as a beacon for microcredit lenders looking to shift their model towards a product that generates higher average returns for borrowers without increasing default. Naturally, further experimentation would be fruitful in order to test, for example, whether each of the three changes from the more “normal” microcredit model (group liability, agricultural focus, balloon repayment) was necessary.

These results are also important for policy analysis and program evaluation. The random choice of communities into which to enter by the lender enables us to estimate ITT effects of the lending program, avoiding strong assumptions on the selection process. Our results provide evidence of quantitatively important selection on unobserved variables, which has methodological implications for impact evaluation. Had we matched borrowers to non-borrowers on observable characteristics, e.g. a quasi-experimental approach, to assess the impact of lending to farmers, we would have overestimated the impact of credit, since conditional on an unusually wide range of observed characteristics those who borrow have substantially higher returns to capital than those who do not borrow.

There are also lessons relevant for the targeting of social programs. Cash transfer programs are often means-tested and recent work suggests that both community targeting, where community members rank-order households to identify the poor, and ordeal mechanisms can be an effective way of generating screening on wealth/income in developing countries (Alatas, Banerjee, Hanna, Olken, & Tobias, 2012; Alatas et al., 2013). We look at a price-based screening mechanism, since agricultural loans charge a positive interest rate that induces selection. In a different agricultural setting, Jack (2013) finds that a willingness to accept mechanism can induce self-selection among landholders in Malawi, leading to improved project success for tree planting. Similarly, Maitra et al. (2020) examines alternative mechanisms for hiring agents to manage loans to farmers, and finds more impact on farmers when the agents had prior experience lending and transacting with farmers. We find that the lending process is a mechanism that generates positive selection so farmers who benefit the most from relaxing capital constraints are more likely to choose to borrow.

We find that the returns to capital in cultivation are heterogeneous and that higher marginal-return farmers self-select into borrowing more so than low marginal-return farmers. But there is

also a set of high marginal return, extremely poor households that are unable to borrow. This has important implications for models of credit markets, as well as social policy that aims to relax liquidity constraints for the most vulnerable. In particular, our results provide rigorous empirical evidence for systematic selection into contracts, which is embedded in several models (Buera, 2009; e.g., Evans & Jovanovic, 1989; Moll, 2014) but which has lacked clear empirical evidence. As recognized by Banerjee et al. (2015) and Kaboski and Townsend (2011), our results highlight the need to incorporate heterogeneity of returns in credit market models.

References

- Adams, D. W. (1971). Agricultural Credit in Latin America: A Critical Review of External Funding Policy. *American Journal of Agricultural Economics*, 53, 163–172.
- Adams, D. W., Graham, D. H., & Von Pischke, J. D. (Eds.). (1984). *Undermining rural development with cheap credit*. Boulder: Westview Press.
- Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., & Tobias, J. (2012). Targeting the Poor: Evidence from a Field Experiment in Indonesia. *The American Economic Review*, 102, 1206–1240.
- Alatas, V., Banerjee, A., Hanna, R., Olken, Benjamin, Purnamasari, R., & Wai_Poi, M. (2013). *Self-Targeting: Evidence from a Field Experiment in Indonesia*.
- Angelucci, M., Karlan, D., & Zinman, J. (2015). Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco. *American Economic Journal: Applied Economics*, 7, 151–182.
- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91, 444–455. JSTOR.
- Armendariz de Aghion, B., & Morduch, J. (2010). *The Economics of Microfinance* (2nd ed.). Cambridge, MA: MIT Press.
- Ashraf, N., Berry, J., & Shapiro, J. M. (2010). Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia. *American Economic Review*, 100, 2383–2413.
- Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113, 7353–7360.
- Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47, 1148–1178.
- Athey, S., & Wager, S. (2019). Estimating Treatment Effects with Causal Forests: An Application. *Observational Studies*, 5, 36–51.
- Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., & Harmgart, H. (2015). The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia. *American Economic Journal: Applied Economics*, 7, 90–122.
- Augsburg, B., De Haas, R., Harmgart, H., & Meghir, C. (2015). The Impacts of Microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics*, 7, 183–203.
- Banerjee, A., Breza, E., Duflo, E., & Kinnan, C. (2015). Do credit constraints limit entrepreneurship? Heterogeneity in the returns to microfinance. *Working Paper*.

- Banerjee, A., & Duflo, E. (2012). Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program. *M.I.T. Working Paper*.
- Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The Miracle of Microfinance? Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics*, 7, 22–53.
- Banerjee, A., Karlan, D., & Zinman, J. (2015). Six Randomized Evaluations of Microcredit: Introduction and Further Steps. *American Economic Journal: Applied Economics*, 7, 1–21.
- Beaman, L., Karlan, D., & Thuysbaert, B. (2014). *Saving for a (not so) Rainy Day: A Randomized Evaluation of Savings Groups in Mali*. Retrieved from <http://www.nber.org/papers/w20600>
- Beaman, L., Karlan, D., Thuysbaert, B., & Udry, C. (2013). Profitability of fertilizer: Experimental evidence from female rice farmers in Mali. *American Economic Review Papers & Proceedings*.
- Breza, E., & Kinnan, C. (2018). Measuring the equilibrium impacts of credit: Evidence from the Indian microfinance crisis. *Working Paper*.
- Bruhn, M., & McKenzie, D. (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics*, 1, 200–232.
- Buera, F. J. (2009). A dynamic model of entrepreneurship with borrowing constraints: Theory and evidence. *Annals of Finance*, 5, 443–464.
- Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I. (2018). Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments. *NBER Working Paper*, 24678.
- Cohen, J., & Dupas, P. (2010). Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment *. *Quarterly Journal of Economics*, 125, 1–45.
- Crépon, B., Devoto, F., Duflo, E., & Pariente, W. (2015). Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco. *American Economic Journal: Applied Economics*, 7, 123–150.
- Crépon, B., El Komi, M., & Osman, A. (2020). Is It Who You Are or What You Get? Comparing the Impacts of Loans and Grants for Microenterprise Development. *Working Paper*. Retrieved from <https://www.adam-osman.com/wp-content/uploads/2020/05/Loans-vs-Grants.pdf>
- Davis, J. M. V., & Heller, S. B. (2017). Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs. *American Economic Review*, 107, 546–550.
- Davis, J. M. V., & Heller, S. B. (2019). Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs. *The Review of Economics and Statistics*, 1–47.
- de Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to Capital in Microenterprises: Evidence from a Field Experiment. *Quarterly Journal of Economics*, 123, 1329–1372.
- de Quidt, J., Fetzer, T., & Ghatak, M. (2012). Group Lending Without Joint Liability. *London School of Economics Working Paper*.
- Druihe, Z., & Barreiro-Huré, J. (2012). Fertilizer subsidies in sub-Saharan Africa. *FAO ESA Working Paper*, No 12-04.

- Duflo, E., Kremer, M., & Robinson, J. (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review*, *101*, 2350–2390.
- Ellis, S. D., & Hine, J. L. (1998). The Provision of Rural Transport services. *Sub-Saharan Africa Transport Policy Program Working Paper*, *37*, 70.
- Evans, D. S., & Jovanovic, B. (1989). An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. *Journal of Political Economy*, *97*, 808–827.
- Fafchamps, M., & Woodruff, C. (2017). Identifying Gazelles: Expert Panels vs. Surveys as a Means to Identify Firms with Rapid Growth Potential. *The World Bank Economic Review*, lhw026.
- Field, E., Pande, R., Papp, J., & Rigol, N. (2013). Does the Classic Microfinance Model Discourage Entrepreneurship Among the Poor? Experimental Evidence from India. *American Economic Review*, *103*, 2196–2226.
- Fink, G., Jack, B. K., & Masiye, F. (2018). *Seasonal Liquidity, Rural Labor Markets and Agricultural Production* (Working Paper No. 24564). National Bureau of Economic Research.
- Giné, X., & Karlan, D. S. (2014). Group versus individual liability: Short and long term evidence from Philippine microcredit lending groups. *Journal of Development Economics*, *107*, 65–83.
- Heckman, J. (1992). Randomization and Social Policy Evaluation. In C. F. Manski & I. Garfinkel (Eds.), *Evaluating welfare and training programs* (pp. 201–230). Cambridge, Mass: Harvard University Press.
- Heckman, J. (1997). Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *The Journal of Human Resources*, *32*, 441.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, *64*, 605–654. JSTOR.
- Hussam, R., Rigol, N., & Roth, B. N. (2020). Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design In The Field. *Working Paper*.
- Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*, 467.
- Jack, B. K. (2013). Private information and the allocation of land use subsidies in Malawi. *American Economic Journal: Applied Economics*, *5*, 113–35.
- Kaboski, J. P., & Townsend, R. M. (2011). A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative. *Econometrica*, *79*, 1357–1406.
- Karlan, D., & Morduch, J. (2009). Access to Finance. In D. Rodrick & M. R. Rosenzweig (Eds.), *Handbook of Development Economics* (Vol. 5). Elsevier.
- Karlan, D., & Mullainathan, S. (2007). Rigidity in Microfinancing: Can One Size Fit All? *QFinance*. Retrieved from <http://www.qfinance.com/financing-best-practice/rigidity-in-microfinancing-can-one-size-fit-all?page=1>
- Karlan, D., Osei-Akoto, I., Osei, R. D., & Udry, C. R. (2013). Agricultural Decisions after Relaxing Credit and Risk Constraints. *Quarterly Journal of Economics*, *Forthcoming*. <https://doi.org/10.2139/ssrn.2169548>
- Karlan, D., & Zinman, J. (2011). Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation. *Science*, *332*, 1278–1284.

- Kazianga, H., & Udry, C. (2006). Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso. *Journal of Development Economics*, 79, 413–446.
- Maitra, P., Mitra, S., Mookherjee, D., & Visaria, S. (2020). Decentralized Targeting of Agricultural Credit Programs: Private versus Political Intermediaries. *National Bureau of Economic Research Working Paper*, 26730. <https://doi.org/10.3386/w26730>
- McKenzie, D. (2018). Can Business Owners Form Accurate Counterfactuals? Eliciting Treatment and Control Beliefs About Their Outcomes in the Alternative Treatment Status. *Journal of Business & Economic Statistics*, 36, 714–722.
- Mckenzie, D. J. (2015). *Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition* (Policy Research Working Paper Series No. 7391). The World Bank.
- McKenzie, D., & Sansone, D. (2019). Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria. *Journal of Development Economics*, 141, 102369.
- Meager, R. (2019). Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments. *American Economic Journal: Applied Economics*, 11, 57–91.
- Meager, R. (2020). Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature. *LSE Working Paper*.
- Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *American Economic Review*, 104, 3186–3221.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66, 688–701.
- Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica*, 79, 159–209.
- Tarozzi, A., Desai, J., & Johnson, K. (2015). The Impacts of Microcredit: Evidence from Ethiopia. *American Economic Journal: Applied Economics*, 7, 54–89.
- Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L., & Yoong, J. (2014). Micro-Loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India. *American Economic Review*, 104, 1909–1941.
- Tibshirani, J., Athey, S., Friedberg, R., Hadad, V., Miner, L., Wager, S., & Wright, M. (2018). grf: Generalized Random Forests (Beta). *ArXiv:1610.01271 [Econ, Stat]*. Retrieved from <https://github.com/grf-labs/grf> R package version 0.10.2
- USAID. (2018). *On the functioning of agricultural markets in Mali*. Retrieved from https://cdn.ymaws.com/www.andeglobal.org/resource/resmgr/research_library/2018-11_MIFP_Study_on_Agricu.pdf
- Wager, S., & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113, 1228–1242.
- Young, A. (2019). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. *The Quarterly Journal of Economics*, 134, 557–598.

Online appendix – not for publication

Appendix A1: Unobservable versus observable predictors of marginal returns

i. Predicting returns based on observable characteristics

Table 1 demonstrated that loan-takers are systematically different at baseline than those who do not take out loans on a number of characteristics, some which are likely to be important in cultivation: they have more land, spend more in inputs, and enjoy higher output and gross profits. Are these baseline characteristics enough to predict who could most productively use capital on their farm? Theoretically, the prediction is ambiguous: in many models those who have the highest returns are households who are the most credit constrained. But we observe that individuals who take out loans have on average *more* wealth in the form of livestock. It could be that those with lower returns to investments in cultivation instead invest in livestock. Several variables show that those who take-up loans are wealthier in general (more land, more livestock, higher consumption), and wealthier households may also have access to better technologies, like a plough, which could increase their returns to capital.

Here we examine whether the marginal returns from grants and the selection effect discussed above are predicted fully by characteristics observed in the baseline, or if there is additional selection that occurs based on unobservables.

We start by examining heterogeneity in returns by observable characteristics in no-loan villages only, in the unselected random sample of farmers. Columns (1) and (2) of Appendix Table 5 show that there is limited evidence of heterogeneity using the variables that we saw to be important in Table 1, including baseline gross profits, baseline land size, and baseline value of livestock. However, the estimates of the interaction terms with observable characteristics are very imprecise, and noise in the data may limit our power to detect heterogeneity. The exercise still demonstrates that it would be difficult for local NGOs or other policymakers to predict returns using easy-to-collect data.

Instead of relying on our intuition for choosing baseline characteristics, we also exploit a machine learning algorithm to estimate heterogeneity in treatment effects (Athey & Imbens, 2016; Athey et al., 2019; Wager & Athey, 2018). Researcher-chosen characteristics may (i) be subject to concerns about inference in light of multiple testing and simultaneously (ii) miss important heterogeneity which results from non-linear combinations of baseline characteristics. See Appendix A2 for details on the implementation of the causal forests algorithm.

In column (3) of Appendix Table 5, we assess heterogeneity using the predicted treatment effects from the algorithm trained on the no-loan village data only. As in Chernozhukov et al. (2018), Davis and Heller (2017) and Davis and Heller (2019), we examine how well the estimated treatment effects (CATEs) predict how gross profits vary with treatment. The point estimate is positive, but noisy (0.33, se=0.58). This is suggestive—but far from conclusive—evidence of heterogeneity in no-loan villages.

Columns (1)–(3) demonstrate that if we had only implemented a cash grant experiment in randomly selected villages, without the experimental design that allows us to compare returns to non-borrowers, we would not have concluded on the basis of the characteristics we observe that there is substantial heterogeneity in the returns to investments in cultivation.

We also estimate CATEs from the causal forests algorithm trained on the selected sample of non-borrowers in loan villages. Appendix Table 5, column (4) looks at this loan villages sub-sample. When we train a causal forest algorithm on this sub-sample, we find strong evidence of heterogeneous treatment effects. Grant * predicted causal effects is positive and significant at the 5% level (1.28, se=0.49).

ii. Does heterogeneity based on observables explain the heterogeneous treatment effects for borrowers and non-borrowers in the experiment?

Appendix Table 5 shows no strong evidence of heterogeneity based on observables in the agricultural returns to grants in the random sample of farmers in no-loan villages. We now explore the possibility that observable characteristics (which we have seen in Table 1 are correlated with loan take-up) can account for the lower return to grants of non-borrowers in loan villages. To explore whether the experiment induces selection not picked up by observable characteristics, we use a specification that interacts baseline characteristics (Z) with an indicator for receiving a grant:

$$Y_{ijt} = \alpha + \beta_1 grant_i + \beta_2 grant_i \cdot loan_j + \gamma_1 grant_i \cdot Z_{ijt} + \gamma_2 Z_{ijt} + X_{ijt}\pi + \lambda_j + \epsilon_{ijt} \quad (20)$$

We structure our analysis by sequentially increasing the controls we include in the regression, by first focusing on Z variables which would be fairly observable to microcredit institutions (MFIs), then including variables which would be fairly observable to the community and therefore may be included in peer screening mechanisms (as in group-lending). Finally, we include the predicted treatment effects from the causal forest model trained on the no-loan villages. This should be a robust synthesis of many covariates, and their interactions.

Appendix Table 6 shows our empirical specification (17) with gross profits as the outcome, with different baseline household-level controls. Column (1) is identical to Column (10) in Table 2 and is included for ease of comparison. Column (2) includes Z variables measured at baseline, and their interactions with grant receipt, that an MFI may be able to easily observe: the household's landholdings (in hectares), the value of their own livestock, gross profits, an indicator for whether the household has six or more adults (the 90th percentile), an indicator for the presence of an extended family, and the number of children in the household. Column (2) shows that the estimates of the differential effect of the grant in loan versus no-loan villages is reduced in magnitude slightly (-US\$33, se=22 compared to -US\$38 without controls) but is qualitatively unchanged. We show the coefficients from the interactions between some of these Z variables and grant receipt. Strikingly, higher baseline gross profits do not predict higher returns to the grant, on average. We also do not observe a statistically significant relationship between baseline livestock value and returns to the grant. However, larger households do benefit more from the grants than smaller households, and households with larger baseline landholdings have lower returns.

Column (3) adds in additional information that would likely be known within the community and thus usable in a peer lending screening process: the primary female respondent's intra-household decision-making power, her engagement in community decision-making and her social capital. In all specifications, the estimates on the differential impacts of the grants in loan versus no-loan villages are slightly smaller in magnitude but qualitatively similar.

Column (4) includes the predicted treatment effects from the causal forest algorithm trained on no-loan villages and then used to predict CATEs for the entire sample. This table uses data from both no-loan and loan villages, but we continue to see no meaningful heterogeneity in returns based on a model trained on the no-loan village data. It is also possible that we lack precision, either due to sample size or too much measurement error.

In section 4 and table 5, we show that a given Z characteristic—for example, gross profits—has a very different relationship with predicted treatment effects (CATEs) depending on whether the algorithm was trained on data from no-loan villages or from loan villages. Those with baseline higher gross profits had higher predicted treatment effects in no-loan villages, but lower predicted treatment effects in loan villages. We may not observe strong evidence of heterogeneous returns in the random sample in Table Appendix Table 5 because there is unobserved heterogeneity within households with similar observable characteristics, i.e. there are characteristics not observed in our data that drive the selection that we uncover through the experiment.

We therefore conclude that our estimates of selection effects are not driven by the rich set of observables we measure at the baseline, but by characteristics more difficult for outsiders to observe, such as land productivity, access to complementary inputs, or farmer skill.

Appendix A2: Causal forest estimates

We implement a generalized causal forest to estimate conditional average treatment effects (CATE) at the observation level. This method has two clear advantages over standard linear regression methods. First, it allows the researcher to consider a relatively high-dimensional set of observable characteristics that may influence the effectivity of the treatment. Second, it accounts for the potentially non-linear relationship between the treatment effect and the predictors.

The causal tree algorithm of Athey and Imbens (2016) selects splits in order to maximize heterogeneity in treatment effects across leaves, less a penalty for the variance of treatment and control outcomes in each leaf. They propose an “honest” approach for estimation, using only one half of the sample (the training sample) to determine and cross-validate the splits. Then, each observation in the second half of the sample (the estimation sample) is assigned to a terminal leaf according to its observable characteristics, and the predicted CATEs are calculated as the difference between the mean outcomes of treatment and control observations within each terminal leaf.

Wager and Athey (2018) builds on this method and propose a causal forest algorithm that assigns each individual observation the average of its predicted CATEs across a large number of trees. Under this approach, each tree is estimated through the honest method described above, but using only a random sub-sample drawn without replacement. Only a random fraction of the available covariates is made available when determining each split.

We employ the generalized causal forest method proposed by Athey et al. (2019), which adapts the Generalized Random Forests method to the estimation of CATEs. The algorithm has two basic steps. First, a causal forest is grown (with each tree based on a random sub-sample of the data, which is then split in half into a training sample to define leaves and an estimation sample to calculate CATEs). Second, each individual CATE is estimated using a set of kernel-based weights for all other observations in the sample. These weights are derived from the fraction of trees where each observation in the sample falls in the same terminal leaf as the target observation i .

Implementation

i. Preparing the dataset

Our sample for the estimation of the causal forests consists of all observations present at both the baseline and the first follow-up rounds of surveys. We estimate a different causal forest for the no-loan villages and the loan villages. The covariates are baseline net revenue, an indicator for the presence of an extended household, per capita food and non-food consumption, the value of livestock owned, area of land cultivated, the value of agricultural assets owned by the household, the total days of labor used, and the index of social capital.

ii. The algorithm

We implement the algorithm using the R package *grf* version 0.10.4 (Tibshirani et al., 2018). Following Athey and Wager (2019), and we allow the algorithm to tune the parameters through cross-validation using the “R-learner” objective function for heterogeneous treatment effects. This regularization method is not a standard cross-validation technique like “leave one out” or k-fold cross validation. It was developed by the authors specifically for generalized random forests. Intuitively, it picks random combinations of parameters to train multiple “mini forests”, then uses the out-of-bag predictions to estimate the objective function (the “R-objective”) for each forest, and picks the combination that minimizes it. This is explained in detail in section 1.3 of Athey & Wager (2019).

The parameters that are determined through this method are the number of variables considered during each split, minimum node size, the fraction of the sample drawn for the construction of each tree, the percentage of observations assigned to the training and the estimations samples, the split balance parameters, and whether empty leaves are pruned from the estimated trees. We used the “tune all” option in the algorithm (instead of manually selecting which parameters to tune) as done in the application in Athey and Wager (2019).

Regarding the number of trees in the forest, the documentation to the *grf* algorithm recommends “that users grow trees in proportion to the number of observations”. Davis and Heller (2019) use 100,000 trees. We tested different number of trees and noticed that the correlation between the predictions across different pairs of random seeds increases slightly with the number of trees in the forest until reaching 100,000 trees, after which it stabilizes. We verified that increasing the number of trees to 250,000, 500,000 or even 1,000,000 does not lead to meaningful changes in the distribution of the predictions or their stability. Therefore, we use 250,000 trees. The correlation between the predictions generated by different random seeds was consistently above 0.9 in the no-loan sample, and above 0.99 in the loan sample. The depth of the trees is

controlled by a parameter (`min.node.size`) in the algorithm and is tuned jointly with the other parameters listed above.

Overfitting

The `grf` algorithm uses honest estimation and the use of out-of-bag predictions to minimize the risk of overfitting. The goal is to avoid overfitting and allow for generalizability without giving up part of the sample when training the forest. Honesty is defined by Wager and Athey (2018) as “A tree is honest if, for each training example i , it only uses the response Y_i to estimate the within-leaf treatment effect or to decide where to place the splits, but not both.” Nevertheless, Davis and Heller (2017) demonstrate that overfitting can occur even with honest estimation. They propose out-of-bag predictions in addition to honest estimation to reduce the overfitting risk. In practice, this means that the prediction for a given observation is calculated using only trees that were not trained with that observation (or cluster, when using cluster-robust estimation as in our case, which we discuss below).

The `grf` package, released after Davis and Heller (2017), uses both out-of-bag predictions and honest estimation by default.

Clustered RCT design

Finally, we account for the fact that the observations in our sample are grouped in unevenly sized clusters (i.e., villages in our setting). In practice, this modifies the causal forest algorithm in two ways. First, the training and estimation samples for each tree are determined by selecting a random subset of clusters, and then drawing an equal number of observations from each cluster. Second, the out-of-bag predictions for each observation i are generated using only the trees where no observation in the training or estimation samples belongs to the same cluster as the target observation i .

Since some clusters in our study have a very small number of observations, we follow Athey and Wager (2019) and increase the number of observations to be drawn from each cluster for the training and estimation samples (the default is the size of the smallest cluster). This improves the stability of the tree-growing algorithm substantially, at the cost of using fewer observations from the clusters that are below this threshold. Considering that our sample size is relatively large compared to other field experiments, we decided to fix this parameter at the 25th percentile of the distribution of cluster sizes.

iii. Assessing treatment heterogeneity

In this subsection, we evaluate whether our generalized casual forest algorithm succeeded in identifying treatment heterogeneity. We conduct this analysis separately for the no-loan and loan villages.

Although the out-of-bag predictions from our model exhibit considerable variation, Athey and Wager (2019) warns that this does not necessarily rule out the possibility that the obtained estimates might just be noisy due to overfitting. Therefore, we follow Davis and Heller (2019) and compare how the predicted CATEs relate with the actual treatment effects. First, we group the observations in each sample into 20 bins according to their predicted CATE. Then we calculate the treatment effect for each bin, following the same specification as in our main results (i.e., we control for net revenue at baseline, village fixed effects, and stratification controls that are listed in the notes to Table 2). Finally, we plot the resulting treatment effect versus the mean predicted CATE per bin in Appendix Figure 1.

To provide a more robust test for heterogeneity, we employ a calibration test motivated by the best linear predictor of CATE method of Chernozhukov et al. (2018). Consider the no-loan villages. Let $B(Z)$ be the random forest predictor of $b_0(Z) \equiv E(Q(0,0)|Z)$, so $B(Z_{ijt})$ is the prediction from the random forest of the net output of a random household with characteristics Z_{ijt} that does not receive a grant or a loan. Similarly, let $S(Z)$ be the causal forest predictor of $s_0(Z) \equiv E(Q(1,0) - Q(0,0)|Z)$, so $S(Z_{ijt})$ is the predicted CATE for a household with characteristics Z_{ijt} . The probability of randomization into the grant treatment is $p(Z)$. We estimate

$$Y_{ijt} = \alpha_0 + \alpha_1 B(Z_{ijt}) + \beta_1 (grant_i - p(Z_{ijt})) + \beta_2 (grant_i - p(Z_{ijt})) (S(Z_{ijt}) - \bar{S}) + S(Z_{ijt}) + \epsilon_{ijt}$$

by weighted least squares using weights $\left(p(Z_{ijt}) (1 - p(Z_{ijt})) \right)^{-1}$. Chernozhukov et al. (2018) shows that rejecting the hypothesis that $\beta_2 = 0$ also rejects the hypothesis that there are no heterogeneous treatment effects, and implies that $S(Z)$ is a relevant predictor of that treatment effect heterogeneity.

In the loan villages, $B(Z)$ is the predictor of $b_l(Z) \equiv E(Q(0,0)|Z, B = 0)$, so $B(Z_{ijt})$ is the prediction from the random forest of the net output of a random household with characteristics Z_{ijt} that does not borrow when in a loan village, and who receives neither a grant nor a loan. Similarly, in the loan villages, $S(Z)$ is the causal forest predictor of $s_l(Z) \equiv E(Q(1,0) - Q(0,0)|Z, B = 0)$.

We find a coefficient for β_2 of -0.03 for the households in the no-loan sample, and a coefficient of 1.05 (p-value = 0.009) for the loan sample. We note that these findings are in line with the

patterns observed in Appendix Figure 1. Overall, the results suggest that the algorithm succeeded in finding meaningful heterogeneity for the loan sample. For the no-loan sample, on the other hand, the evidence is weak and inconclusive.

Appendix A3: Efficient allocation

An efficient allocation maximizes the gain in output from loans ($\eta \equiv \Delta_B Q$), net of the cost of capital to the lender (ρ). The efficient allocation is defined by the function $B(\eta_i, Q_i^0)$ chosen to maximize

$$\int_{\underline{\eta}}^{\bar{\eta}} \int_{\underline{Q^0}}^{\overline{Q^0}} B_i(\eta_i - \rho) h(\eta, Q^0) dQ^0 d\eta \quad (21)$$

where $h(\eta, Q^0)$ is the joint density of marginal returns to capital ($Q(1,0) - Q(0,0)$) and $Q(0,0)$ implied by the joint distribution of potential outcomes $F(Q(1,0), Q(0,1), Q(0,0))$ defined in section 3.

In this efficient allocation, $B_i(\eta_i, Q_i) = 1$ if $\eta_i \geq \rho$, and $B_i = 0$ otherwise.

However, suppose there is limited liability. Because of limited liability, the maximum repayment that the lender can obtain from a borrower i is $Q_i^1 - Q_i^0 = \eta_i$ if $\underline{c} \leq Q_i^0$, $Q_i^0 + \eta_i - \underline{c}$ if $Q_i^0 \leq \underline{c} \leq Q_i^0 + \eta_i$, and 0 if $Q_i^0 + \eta_i \leq \underline{c}$. The breakeven constraint of the lender, therefore, is

$$\begin{aligned} \int_{\underline{\eta}}^{\bar{\eta}} \int_{\underline{c}}^{\overline{Q^0}} B_i \eta_i h(\eta, Q^0) dQ^0 d\eta + \int_{\underline{\eta}}^{\bar{\eta}} \int_{\underline{c}-\eta_i}^{\underline{c}} B_i (Q_i^0 + \eta - \underline{c}) h(\eta, Q^0) dQ^0 d\eta \\ \geq \int_{\underline{\eta}}^{\bar{\eta}} \int_{\underline{Q^0}}^{\overline{Q^0}} B_i \rho_i h(\eta, Q^0) dQ^0 d\eta. \end{aligned} \quad (22)$$

The left hand side of the breakeven constraint is the revenue generated by the lending, which is equal to the full gain in output for farmers not subject to the limited liability constraint plus the constrained payments from those farmers subject to the limited liability constraint (which are zero for all farmers with $Q_i^0 + \eta_i < c$). The RHS is the cost of all loans. The constrained efficient allocation is the function $B(\eta_i, \theta_i)$ that maximizes (1) subject to the breakeven constraint (22).

If the breakeven constraint does not bind when $B_i = 1$ for all farmers i with $\eta_i \geq \rho$, and $B_i = 0$ for all farmers with $\eta_i < \rho$, then the unconstrained efficient allocation remains feasible. The

breakeven constraint may not bind at the unconstrained efficient allocation if the distribution of farmers is such that the surplus generated by farmers for whom limited liability does not bind is sufficient to cover the losses from borrowers who are (at least partially) defaulting. In this case

$$\int_{\rho}^{\bar{\eta}} \int_{\underline{c}+\rho-\eta_i}^{\bar{Q}^0} (\eta_i - \rho) h(\eta, Q^0) dQ^0 d\eta + \int_{\rho}^{\bar{\eta}} \int_{\underline{Q}}^{\underline{c}+\rho-\eta_i} (\max(Q_i^0 + \eta_i - \underline{c}, 0) - \rho) h(\eta, Q^0) dQ^0 d\eta \geq 0.$$

The first term is the surplus generated from high-return farmers ($\eta_i \geq \rho$) who pay the cost of their loans in full ($Q_i^0 + \eta_i \geq \underline{c} + \rho$). The second term are the losses from high return farmers ($\eta_i \geq \rho$) who are too poor to fully repay the cost of their loans ($Q_i^0 + \eta_i < \underline{c} + \rho$). In this case, the allocation remains efficient.

However, if (22) is violated at the unconstrained efficient allocation, then it remains the case that $B_i = 1$ for all farmers with both $Q_i^0 + \eta_i \geq \underline{c}$ and $\eta_i \geq \rho$ (because such loans relax the breakeven constraint and increase net gain in output), and $B_i = 0$ for all farmers with $\eta_i \leq \rho$ because such loans decrease the net gain in output and tighten the breakeven constraint. However, not all farmers with high marginal returns and low base output Q_i^0 can receive loans. The allocation of these remaining loans is determined by the function $B(\eta, Q^0)$ to maximize

$$\int_{\rho}^{\bar{\eta}} \int_{\underline{Q}^0}^{\underline{c}+\rho-\eta_i} B_i(\eta_i - \rho) h(\eta, Q^0) dQ^0 d\eta \quad (23)$$

subject to

$$\begin{aligned} \int_{\rho}^{\bar{\eta}} \int_{\underline{Q}^0}^{\underline{c}+\rho-\eta_i} B_i(\max(Q_i^0 + \eta_i - \underline{c}, 0) - \rho) h(\eta, Q^0) dQ^0 d\eta \\ \leq \int_{\rho}^{\bar{\eta}} \int_{\underline{c}+\rho-\eta_i}^{\bar{Q}^0} (\eta_i - \rho) h(\eta, Q^0) dQ^0 d\eta \end{aligned} \quad (24)$$

The RHS of (24) is a constant, the surplus generated by lending to high return farmers who repay the full cost of their loans. The problem is to allocate that fixed budget across the set of high-return farmers who cannot fully repay their loans to maximize (23).

The increase in (23) from lending to farmer j is $\eta_j - \rho$, while the cost is $\rho - \max(0, Q_j^0 + \eta_j - \underline{c})$. Therefore, farmers are allocated loans in order of decreasing ratios of benefit to cost: if $B_j = 1$ and $B_k = 0$, then $\frac{\eta_j - \rho}{\rho - \max(0, Q_j^0 + \eta_j - \underline{c})} \geq \frac{\eta_k - \rho}{\rho - \max(0, Q_k^0 + \eta_k - \underline{c})}$, and the boundary between $B(\eta, Q^0) = 1$

and $B(\eta, Q^0) = 0$ for farmers who partially repay their loans is characterized by $\frac{\eta - \rho}{\rho - Q^0 + \eta - \underline{c}} = k$ for some constant $k > 0$. Therefore, the boundary between borrowers and nonborrowers in a constrained efficient allocation is downward sloping in (η, θ) . Thus, some farmers with high returns to capital may not receive loans, while similar farmers with the same marginal productivity but higher baseline output do borrow.

Appendix A4: Randomization inference

We follow Young (2019) to implement the Randomization Inference (RI) procedure.³⁷ First, we generated 10,000 simulations of the assignment of grants. In each simulation, we reproduced the re-randomization routine described in Section 2.1 to ensure that the grant assignments are drawn from the same distribution as the original experiment. We took the villages type (loan village / no-loan village), as well as the selection of households in loan villages into taking the loan, as given. Therefore, the sample of eligible recipients of the grant (i.e., all households in no-loan villages and non-borrowers in loan villages) was pre-determined and identical across all iterations. In each iteration, we reproduced the main analysis using the synthetic treatment assignment and stored the coefficients for all the relevant tests. That is, we re-estimated the effect of receiving a grant and its interaction with village type on all the agricultural outcomes of interest, for each year of the experiment. We then used the results to approximate the covariance matrix of the estimated coefficients of interest across the universe of potential treatment assignments. This allowed us to calculate the randomization-c p-values from a two-tailed test of significance for each treatment effect, as in Young (2019). We also implement randomization-based joint testing procedures to test the null hypothesis that all relevant treatment effects in an equation family are zero. To avoid grouping together aggregate outcomes of interest with their individual components, we divide the agricultural variables into three independent families: (i) agricultural inputs and crop choice, (ii) total input expenses and value of output, and (iii) gross profit. We report RI p-values for joint Wald tests of significance of the treatment effects of the grant and its interaction with village type on all the outcomes in a given family (i.e., an omnibus test of overall experimental significance for that equation group).

³⁷ We use an adapted version of the Stata command “randcmd” (Young, 2020) which allows for more flexibility in the randomization routine.