Econometrica Supplementary Material

SUPPLEMENT TO "SELECTION INTO CREDIT MARKETS: EVIDENCE FROM AGRICULTURE IN MALI"

(*Econometrica*, Vol. 91, No. 5, September 2023, 1595–1627)

LORI BEAMAN
Department of Economics, Northwestern University

DEAN KARLAN
Kellogg School of Management, Northwestern University, IPA, J-PAL, and NBER

BRAM THUYSBAERT
FMO - Dutch Development Bank

CHRISTOPHER UDRY
Department of Economics, Northwestern University

APPENDIX

A.1. *Sample and Randomization Details*

Sample

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. 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).

Randomization Stratification and Re-Randomization Procedures

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,

Lori Beaman: l-beaman@northwestern.edu
Dean Karlan: karlan@northwestern.edu
Bram Thuysbaert: b.thuysbaert@fmo.nl
Christopher Udry: christopher.udry@northwestern.edu

© 2023 The Authors. Econometrica published by John Wiley & Sons Ltd on behalf of The Econometric Society. Lori Beaman is the corresponding author on this paper. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

fertilizer use, and plough access. See Bruhn and McKenzie (2009) for a more detailed description of the randomization procedure.

A.2. *Loan Allocation With Frictions*

A.2.1. *Limited Liability*

Consider a simple limited liability model of credit. A frictionless allocation maximizes the gain in gross profits from loans, net of the cost of capital to the lender (ρ). The frictionless allocation is defined by the indicator function $B(\Delta_B Q, Q^{NG})$ chosen to maximize

$$\int_{\underline{\Delta_B Q}}^{\overline{\Delta_B Q}} \int_{\underline{Q^{NG}}}^{\overline{Q^{NG}}} B(\Delta_B Q, Q^{NG})(\Delta_B Q - \rho)\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q, \tag{14}$$

where $\tilde{f}(\Delta_B Q, Q^{NG})$ is the joint density of marginal returns to borrowing ($Q^B - Q^{NG}$) and Q^{NG} implied by the joint distribution of potential outcomes $F(Q^{NG}, Q^G, Q^B)$ defined in Section 3.

In this frictionless allocation, $B(\Delta_B Q, Q^{NG}) = 1$ if $\Delta_B Q \geq \rho$, and $B = 0$ otherwise.

However, suppose there is limited liability (the farmer must be left with at least $\underline{c}$ after loan repayment) and the farmer participation constraint that repayment be less than or equal to $\Delta_B Q$. Because of limited liability, the maximum repayment that the lender can obtain from a borrower is $\Delta_B Q$ if $\underline{c} \leq Q^{NG}$, $Q^{NG} + \Delta_B Q - \underline{c}$ if $Q^{NG} \leq \underline{c} \leq Q^{NG} + \Delta_B Q$, and 0 if $Q^{NG} + \Delta_B Q \leq \underline{c}$. The breakeven constraint of the lender, therefore, is

$$\begin{aligned}
&\int_{\underline{\Delta_B Q}}^{\overline{\Delta_B Q}} \int_{\underline{c}}^{\overline{Q^{NG}}} \Delta_B Q B(\Delta_B Q, Q^{NG})\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q \\
&\quad + \int_{\underline{\Delta_B Q}}^{\overline{\Delta_B Q}} \int_{\underline{c}-\Delta_B Q}^{\underline{c}} (\Delta_B Q - \underline{c}) B(\Delta_B Q, Q^{NG})\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q \\
&\geq \int_{\underline{\Delta_B Q}}^{\overline{\Delta_B Q}} \int_{\underline{Q^{NG}}}^{\overline{Q^{NG}}} \rho B(\Delta_B Q, Q^{NG})\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q.
\end{aligned} \tag{15}$$

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 $Q_i^{NG} + \Delta_B Q \leq \underline{c}$). The RHS is the cost of all loans. The constrained efficient allocation is the function $B(\Delta_B Q, Q^{NG})$ that maximizes (14) subject to the breakeven constraint (15).

If the breakeven constraint does not bind when $B_i = 1$ for all farmers i with $\Delta_B Q_i \geq \rho$, and $B_i = 0$ for all farmers with $\Delta_B Q_i < \rho$, then the frictionless 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,

$$\int_{\rho}^{\overline{\Delta_B Q}} \int_{\underline{c}+\rho-\Delta_B Q}^{\overline{Q^{NG}}} (\Delta_B Q - \rho)\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q$$
$$+ \int_{\rho}^{\overline{\Delta_B Q}} \int_{\underline{Q^{NG}}}^{\underline{c}+\rho-\Delta_B Q} (\max(Q^{NG} + \Delta_B Q - \underline{c}, 0) - \rho)\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q \geq 0.$$

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

However, if (15) is violated at the unconstrained efficient allocation, then it remains the case that $B_i = 1$ for all farmers with both $Q_i^{NG} + \Delta_B Q_i \geq \underline{c}$ and $\Delta_B Q_i \geq \rho$ (because such loans relax the breakeven constraint and increase net gain in output), and $B_i = 0$ for all farmers with $\Delta_B Q_i \leq \rho$ (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^{NG} can receive loans. The allocation of these remaining loans is determined by the function $B(\Delta_B Q, Q^{NG})$ to maximize

$$\int_{\rho}^{\overline{\Delta_B Q}} \int_{\underline{Q^{NG}}}^{\underline{c}+\rho-\Delta_B Q} (\Delta_B Q - \rho)B(\Delta_B Q, Q^{NG})\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q \tag{16}$$

subject to

$$\int_{\rho}^{\overline{\Delta_B Q}} \int_{\underline{Q^{NG}}}^{\underline{c}+\rho-\Delta_B Q} (\max(Q^{NG} + \Delta_B Q - \underline{c}, 0))B(\Delta_B Q, Q^{NG})\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q$$
$$\leq \int_{\rho}^{\overline{\Delta_B Q}} \int_{\underline{c}+\rho-\Delta_B Q}^{\overline{Q^{NG}}} (\Delta_B Q - \rho)\tilde{f}(\Delta_B Q, Q^{NG})\, dQ^{NG}\, d\Delta_B Q. \tag{17}$$

The RHS of (17) 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 (16).

The increase in (16) from lending to farmer j is $\Delta_B Q_j - \rho$, while the cost is $\rho - \max(Q_j^{NG} + \Delta_B Q_j - \underline{c}, 0)$. Therefore, farmers are allocated loans in order of decreasing ratios of benefit to cost: if $B_j = 1$ and $B_k = 0$, then $\frac{\Delta_B Q_j - \rho}{\rho - \max(Q_j^{NG} + \Delta_B Q_j - \underline{c}, 0)} \geq \frac{\Delta_B Q_k - \rho}{\rho - \max(Q_k^{NG} + \Delta_B Q_k - \underline{c}, 0)}$, and the boundary between $B(\Delta_B Q, Q^{NG}) = 1$ and $B(\Delta_B Q, Q^{NG}) = 0$ for farmers who partially repay their loans is characterized by $\frac{\Delta_B Q - \rho}{\rho - \max(Q^{NG} + \Delta_B Q - \underline{c}, 0)} = k$ for some constant $k > 0$. Therefore, the boundary between borrowers and non-borrowers in a constrained efficient allocation is downward sloping in $(Q^{NG}, \Delta_B Q)$. 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.

A.2.2. *Risk Aversion*

Alternatively, consider expected utility-maximizing farmers with decreasing absolute risk aversion. They are presented with an opportunity to borrow a fixed amount at cost

ρ, with full enforcement. The loan would finance a risky project with random return $\Delta_B Q$ over baseline gross profit Q^{NB}. Suppose $E(\Delta_B Q) \geq \rho$ and that there is a farmer i indifferent between taking the loan to finance the project or not. Then any farmer with a higher no-grant gross profit with the same preferences and investment opportunity would strictly prefer to take the loan, and indeed would take a loan to finance a strictly inferior investment opportunity, with returns that are first-order stochastically dominated by the project with return $\Delta_B Q_i$.[1] Farmers with lower no-grant gross profits require higher expected returns to be willing to accept the additional risks associated with borrowing. Risk aversion and self-selection also generates a downward sloping (dashed line in Figure 3c) boundary in $(E(Q^{NG}), E(\Delta_B Q))$ between those who do and do not borrow.

Risk-averse farmers will in general select different projects to finance with grants and loans. Suppose a farmer receiving a loan is indifferent between two risky projects with returns $\eta^1 \equiv \Delta_B Q^1$ and $\eta^2 \equiv \Delta_B Q^2$ with $E(\eta^1) > E(\eta^2)$. That farmer would strictly prefer the riskier, higher expected return project 1 if offered a grant rather than a loan. Therefore, the project chosen by the marginal borrower who is given a grant instead will have an expected return (weakly) greater than the project that that farmer would have chosen to implement with the loan. Risk aversion generates selection across projects of a farmer as well as across farmers. Therefore, in Figure 3c, the solid line boundary in $(E(Q^{NG}), E(\Delta_G Q))$ between those who borrow and those who do not lies above that boundary in $(E(Q^{NG}), E(\Delta_B Q))$, and with DARA preferences the difference between the boundaries declines as $E(Q^{NG})$ rises.[2] Within-farmer selection of projects implies $E(\Delta_G Q|C = 1) \geq E(\Delta_B Q|C = 1)$. Since we have shown (in (8) and (10)) that each of these quantities is identified by our experimental design, in Section 6 we examine the evidence that farmers may be selecting among projects.

A.3. *Causal Forest Estimates*

We implement a generalized causal forest to estimate conditional average treatment effects (CATEs) 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.

We employ the generalized causal forest method proposed by Athey, Tibshirani, and Wager (2019), which adapts the causal forests of Wager and Athey (2018) to a generalized random forest approach. These forests are aggregations of "causal trees": a variant of a regression tree which recursively partitions the sample one covariate at a time (e.g., split based on gender then based on income then based on household size). This results in a tree model with each observation assigned to a single terminal node, or "leaf." Each causal tree is grown from a random subsample of the data drawn without replacement, which is then split in half. The first half is used to select splits in the tree, maximizing heterogeneity of treatment effects across terminal nodes while penalizing for variance

[1] For i, $EU(\Delta_B Q_i + Q_i^{NG} - \rho) = EU(Q_i^{NG})$. Then farmer j with $Q_j^{NG} > Q_i^{NG}$ with the same project has $EU(\Delta_B Q_i + Q_j^{NG} - \rho) > EU(Q_j^{NG})$. So there is a constant $\epsilon_j > 0$ with $EU(\Delta_B Q_i - \epsilon_j + Q_j^{NG} - \rho) > EU(Q_j^{NG})$.

[2] This discussion may raise the possibility that farmers borrowing with a limited liability constraint may also choose different projects than they would with a grant. In this case, the convexity of returns generated by the limited liability could induce borrowers to take *more* risk. However, this would imply some default in equilibrium, and we observe no instance of a defaulted loan.

of treatment and control outcomes within leaves. Then, each observation in the second half is assigned to a leaf according to the constructed tree and their predicted CATEs are calculated as the treatment effect within each leaf. After all trees are constructed and CATEs are produced, each observation is assigned a single predicted CATE estimated using a kernel-based weighted average of their predicted CATEs. These weights are derived from the fraction of trees where each observation in the sample falls in the same leaf as the target observation.

Implementation

A.3.1. *Preparing the Data Set*

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.

A.3.2. *The Algorithm*

We implement the algorithm using the R package *grf* version 2.20 (Athey, Tibshirani, and Wager (2019)). Following Athey and Wager (2019), 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.

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 (2020) used 100,000 trees. We tested different numbers 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) demonstrated that overfitting can occur even with honest estimation. They proposed 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

A key aspect of our experimental design is that the loan experiment is a clustered design, with randomization at the village level. We need to adjust the implementation of Athey and Wager (2019) in a few ways. In our context, the clusters are uneven in size (villages are not all the same size), and there are some clusters with a small number of observations. This leads to three adjustments. 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, we further reduce the risk of overfitting by adjusting the way we construct our out-of-bag predictions. We ensure that for each observation i, the prediction is generated using only the trees where no observation in the training or estimation samples belongs to the same cluster as the target observation.

Inference With Causal Forest Estimates

We follow the sample-splitting method of Chernozhukov, Demirer, Duflo, and Fernández-Val (2020) to produce confidence intervals adjusted for the variability of predicted CATEs when used as regressors. This method uses repeated 50% subsamples to estimate the causal forest with one half and produce predicted CATEs and regression estimates with the other half. We then take the median point estimates and standard errors from 1000 iterations of 1000-tree forests to generate confidence intervals.

A.4. *Predicting Returns Based on Observable Characteristics*

Table I demonstrated that loan-takers are systematically different at baseline than those who do not take out loans on a number of characteristics, some of 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

APPENDIX TABLE I

BALANCE CHECK.

	Loan vs no-loan villages			Grants vs no-grants in no-loan villages			Grants vs no-grants in loan villages		
	Mean of control group	Difference between T and C	p-value	Mean of control group	Difference between T and C	p-value	Mean of control group	Difference between T and C	p-value
Household size	7.43	0.07	0.52	7.45	−0.07	0.59	7.42	−0.04	0.77
Land (ha)	2.00	0.23	0.05	2.00	−0.00	1.00	2.19	−0.01	0.89
Days of family labor	144.57	−1.03	0.88	145.71	−0.23	0.97	137.93	5.78	0.22
Days of hired labor	11.51	1.36	0.24	11.33	0.17	0.83	12.46	−0.27	0.77
Input expenses (USD)	137.98	18.65	0.19	135.66	16.48	0.08	150.95	0.31	0.98
Agricultural output (USD)	548.90	47.39	0.24	542.33	20.18	0.45	579.00	12.13	0.67
Livestock value (USD)	1520.29	−120.52	0.28	1515.83	2.63	0.98	1389.71	−36.17	0.79
Has a business	0.54	0.01	0.67	0.53	0.02	0.35	0.54	0.01	0.61
Monthly non-food expenses	40.89	−0.06	0.97	41.21	−0.83	0.59	40.37	−0.39	0.80
F-test for joint significance			0.13			0.41			0.89

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 VI show that there is limited evidence of heterogeneity using the variables that we saw to be important in Table I, 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, as described in Appendix A3. In Appendix Table VI, column (3), we assess heterogeneity using the predicted treatment effects from the algorithm trained on the no-loan village data only. As in Chernozhukov et al. (2020) and Davis and Heller (2017, 2020), we examine how well the estimated treatment effects (CATEs) predict how gross profits vary with treatment. The point estimate is positive but noisy (0.39, se = 0.56), suggesting—but far from concluding—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 VI, column (4) looks at this loan-villages subsample. When we train a causal forest algorithm on this subsample, we find strong evidence of heterogeneous treatment effects. Grant * predicted causal effects is positive and significant at the 5% level (1.19, se = 0.56). Baseline characteristics, among a selected sample of non-borrowers, can predict heterogeneity in the returns to capital, but we can only detect this heterogeneity with the assistance of the two-stage experiment.

A.5. *Randomization Inference*

We follow Young (2019) to implement the Randomization Inference (RI) procedure.[3] 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

[3]We use an adapted version of the Stata command “randcmd” (updated 5/20 by Young) which allows for more flexibility in the randomization routine.

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).

A.6. *Robustness*

A.6.1. *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. We do observe grant-recipients in no-loan villages cultivating more land (and land cultivation is of course a decision made early in the agricultural cycle). But, when we exploit the variation in timing within treatment groups, we do not find cause for concern: the land cultivation decision as well as an index of all agricultural outcomes is uncorrelated with the timing of the grants within the grant-recipient households in no-loan villages (Appendix Table III).[4]

A.6.2. *Spillovers*

In this section, we use data from an additional 69 villages in the same administrative units (cercles) as our study villages to look for evidence of spillovers in loan and no-loan villages.[5] Appendix Table IV 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

[4]We employ two main specifications for this test: 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). 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.

[5]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 IV qualitatively unchanged.

APPENDIX TABLE II

ATTRITION.

	Loan vs no-loan villages				Grants vs no-grants in no-loan villages				Grants vs no-grants in loan villages			
	Year 1		Year 2		Year 1		Year 2		Year 1		Year 2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	0.001	−0.008	0.010	−0.007	0.006	0.013	0.000	−0.004	0.000	0.002	−0.004	0.001
	(0.004)	(0.007)	(0.007)	(0.011)	(0.005)	(0.014)	(0.005)	(0.014)	(0.004)	(0.010)	(0.006)	(0.016)
Interaction of treatment and:												
Household size		0.000		0.000		0.000		0.001		−0.001		0.001
		(0.001)		(0.001)		(0.002)		(0.001)		(0.001)		(0.002)
Land (ha)		0.002		−0.005		0.002		0.006		0.002		0.000
		(0.003)		(0.004)		(0.004)		(0.005)		(0.004)		(0.004)
Days of family labor†		0.000		−0.001		−0.001		−0.001		−0.001		−0.002
		(0.000)		(0.001)		(0.000)		(0.000)		(0.001)		(0.001)
Days of hired labor†		−0.001		−0.002		−0.002		0.000		−0.001		0.001
		(0.002)		(0.002)		(0.002)		(0.002)		(0.002)		(0.003)
Input expenses*		0.000		0.002		0.002		0.006		0.002		0.002
		(0.003)		(0.004)		(0.003)		(0.005)		(0.005)		(0.006)
Agricultural output*		0.000		0.005		0.000		−0.002		0.000		0.000
		(0.001)		(0.002)		(0.001)		(0.002)		(0.002)		(0.002)
Livestock value*		0.000		0.000		0.000		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Has a business		0.014		0.024		0.007		0.005		0.012		0.015
		(0.005)		(0.008)		(0.011)		(0.010)		(0.010)		(0.009)
		−0.015		−0.010		−0.005		−0.002		0.026		0.004
Monthly non-food expenses*												
		(0.009)		(0.011)		(0.011)		(0.013)		(0.017)		(0.019)
N	5649	5465	5649	5465	3201	3104	3201	3104	2448	2361	2448	2361
Mean attrition control	0.014		0.015		0.012		0.015		0.015		0.026	
F-test for joint significance of coefficients of treatment and interaction terms		0.14		0.14		0.56		0.51		0.19		0.08

Note: *Variables divided by 100 for ease of exposition. †Variable divided by 10 for ease of exposition.

APPENDIX TABLE III
TIMING ROBUSTNESS (NO-LOAN VILLAGES).

	Index			Land Size		
	(1)	(2)	(3)	(4)	(5)	(6)
Date (linear)	0.001	−0.001		0.000	0.001	
	(0.004)	(0.008)		(0.012)	(0.026)	
Date squared		0.000			0.000	
		(0.000)			(0.001)	
1 if before June 1st			−0.061			−0.110
			(0.138)			(0.457)
Revisit to village	−0.013	−0.030	−0.032	−0.027	−0.020	−0.087
	(0.104)	(0.117)	(0.119)	(0.344)	(0.386)	(0.394)
Observations	787	787	787	787	787	787

[1] Index includes: land area, number of family labor days, number of hired labor days, an indicator for whether fertilizer was used, value of fertilizer expenses, value of other chemical expenses, value of all input expenses, value of harvest, and profits.

[2] Sample includes only grant recipients in no-loan villages.

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 4 laborer-days (se = 1.37). While this is precisely estimated and a point estimate comparable to main treatment effect in Table II, 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 and Hine (1998)). Column (10) shows no change in an index of wages.

We note that this analysis cannot speak directly to the possibility of spillovers in loan villages. Recent evidence by Banerjee, Breza, Chandrasekhar, Duflo, Jackson, and Kinnan (2021) highlights how the introduction of formal credit can alter existing informal risk-sharing arrangements. Our main concern is whether the patterns of spillovers are different in loan versus no-loan villages. If so, this would affect our interpretation of the results as being about selection into credit. In Appendix Table V, we analyze data on loans given to and received from family and friends. We compare households in no-intervention villages with non-borrowers in the loan villages and households in no-loan villages. We find evidence that grant recipients in no-loan villages give out more loans. But there is no evidence of more loans to non-borrowers in loan villages. In fact, non-borrowing households in loan villages are less likely to receive loans[6] than households in no-intervention villages and no-grant households in no-loan villages. Moreover, the point

[6] The lower rate of receiving informal loans among no-grant non-borrowers could reflect that (i) they have low demand for loans since they opted out of borrowing (i.e., they do not have a high-return project) or (ii) they are poor and too risky to lend to.

APPENDIX TABLE IV

SPILLOVERS IN NO-LOAN VILLAGES.

	Land cultivated (ha)	Land planted with rice and groundnut (ha)	Family labor (days)	Hired labor (days)	Fertilizer and chemical expenses (USD)	Total input expenses (USD)	Value output (USD)	Gross Profit (USD)	Price Index	Wage index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
No-loan village	−0.17	0.00	1.21	3.70	4.35	−1.14	−7.27	−12.64	0.00	−0.21
	(0.13)	(0.06)	(10.04)	(1.37)	(10.16)	(15.84)	(50.07)	(37.44)	(0.40)	(0.42)
N	3679	3705	3679	3679	3705	3679	3679	3679	175	170
Mean of excluded group	2.19	0.91	129.24	13.76	137.13	196.13	609.82	415.72	0.08	0.09
SD of excluded group	2.65	0.91	147.93	22.24	267.82	314.30	845.22	612.08	0.97	0.97

[1]The sample includes (i) households in no-intervention villages and (ii) households in no-loan villages who did not receive a grant. The analysis uses only data from follow-up year 1. The excluded group are households in no-intervention villages.

[2]Additional controls for columns (1)–(8) include: cercle fixed effects; the baseline value of the dependent variable, along with a dummy when missing; the baseline value of the dependent variable interacted with the no-intervention village dummy; an indicator for the HH being administered the input survey in 2011; village-level stratification controls as listed in Table VI; and individual-level stratification controls as listed in Table II. Standard errors are clustered at the village level.

[3]Columns (9) ad (10) are village-level regressions. Additional controls include *cercle* fixed effects and the village-level stratification controls. Also included are the following individual controls: the number of adult household members, the number of children in the household, the average age of adults in the household, and the share of adults with primary school education level.

[4]The price index is a normalized average of grain prices and livestock. The wage index is a normalized average of wages for men, women, and children for three agricultural activities.

APPENDIX TABLE V

INFORMAL BORROWING AND LENDING.

	Received loan from family or friend in previous 12 months	Amount ($) received in loans from family and friends in previous 12 mo	Gave out loan in previous 12 months	Amount ($) given out as loans in previous 12 months
	(1)	(2)	(3)	(4)
No-loan village	0.10	23.44	−0.07	1.04
	(0.04)	(8.42)	(0.04)	(5.16)
Loan village	−0.04	−10.06	−0.02	0.75
	(0.02)	(4.87)	(0.02)	(3.16)
Grant	−0.04	−4.14	0.09	11.53
	(0.02)	(5.95)	(0.02)	(3.73)
Grant * Loan village	0.00	−2.67	−0.04	−5.16
	(0.03)	(7.71)	(0.03)	(5.37)
p-value for Grant + grant*loan	0.009	0.163	0.008	0.102
p-value for No loan = Loan	0.003	0.002	0.261	0.965
N	6184	6513	6513	6513
Mean of no-intervention sample	0.41	69.67	0.52	49.09
SD	0.49	127.13	0.50	104.69

[1] The sample includes: all households in no-intervention villages, all households in no-loan villages, and non-borrowers in loan villages in year 1.

[2] Additional controls include: cercle fixed effects; the baseline value of the dependent variable, along with a dummy when missing, the baseline value interacted with being a GE village and the missing indicator (we only have baseline data in non-intervention villages for 330 out of the 1330 households), and village-level stratification controls listed in the notes of Table VI.

[3] Standard errors are in parentheses and clustered at the village level in all specifications.

estimates are sufficiently small that spillovers are unlikely to be driving the main results in Table II. This analysis comes with the important caveat that we are unsure whether the no-intervention villages are comparable at baseline to study villages.

A.7. *Persistence of Treatment Effects*

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 Appendix Table IX, column (9) shows that output is higher in grant recipient households by US$77 (se = 24) and column (10) demonstrates that gross profit was higher by US$62 (se = 18). This is striking since we do not observe grant-recipient households spending more on inputs *that we can easily measure* in column (8) (US$10.53, se = 11.87). 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 8.0% 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 5.2% more likely to use a plough during land preparation (4 pp, se = 1), and used 7.6% more seeds (7.1 kg, se = 2.8).

We show the estimates of the interaction term of Grant * Loan village in year 2 in Appendix Table IX, but the interpretation of the results is challenging. In the second

APPENDIX TABLE VI

ARE RETURNS PREDICTED BY BASELINE CHARACTERISTICS?

	Gross Profit			
	(1)	(2)	(3)	(4)
Grant	−3.79	−6.37	22.81	17.42
	25.65	25.63	29.55	15.37
			[−73.64, 127.05]	[−41.16, 82.62]
Predicted causal effects			−1.11	−2.69
			0.41	0.47
			[−1.63, 0.58]	[−3.03, 0]
Grant * Predicted causal effects			0.39	1.19
			0.56	0.56
			[−1.56, 2.31]	[−1.19, 3.03]
Grant * Baseline gross profit	0.04	0.04		
	(0.06)	(0.06)		
Grant * Baseline land	−1.61	−1.91		
	(11.51)	(11.47)		
Grant * Baseline value of livestock	0.02	0.02		
	(0.01)	(0.01)		
Grant * Large HH at baseline	68.36	70.13		
	(45.02)	(46.18)		
Grant * Baseline social index		−34.18		
		(14.15)		
Grant * Baseline intra-household bargaining index		−9.26		
		(15.04)		
N	3160	3159	3065	2142
Year	1	1	1	1
Sample	No loan vill	No loan vill	No loan vill	Loan vill
Additional HH structure controls interacted with grant & year	Yes	Yes	No	No
HH decision-making/community action interacted with grant & year	No	Yes	No	No
Mean of Baseline gross profit	408.96			
SD of Baseline gross profit	528.08			
Mean of Baseline land	2.03			
SD of Baseline land	2.43			

[1]See the notes of Table II for details on specification and additional controls.

[2]Large household is 6 or more adults in the household.

[3]Other household structure controls include: an indicator for the presence of an extended family and the number of children in the household.

[4]Predicted causal effects in column (3) are generated by a causal forest algorithm on no-loan village data and then extrapolated to all no-loan village households. Predicted causal effects in column (4) are generated by a causal forest algorithm on loan village data and then extrapolated to all loan village households.

[5]Columns (3) and (4) show sample splitting confidence intervals, suggested by Chernozhukov et al. (2020) and adapted by Davis and Heller (2020) for causal forests. See Appendix A3.ii for details.

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

APPENDIX TABLE VII

RETURNS TO GRANT FOR BOTTOM TERCILE OF BASELINE CHARACTERISTICS.

	Gross Profit			
	(1)	(2)	(3)	(4)
Grant	50.86	40.40	53.15	57.81
	(22.58)	(22.67)	(22.29)	(21.27)
Grant * Loan village	−64.03	−45.52	−92.16	−84.31
	(32.10)	(31.29)	(32.92)	(30.50)
Grant * T1 Baseline gross profit	−24.71			
	(29.04)			
Grant * T1 Baseline gross profit * Loan village	65.37			
	(47.24)			
Grant * T1 Baseline livestock		6.86		
		(33.06)		
Grant * T1 Baseline livestock * Loan village		6.47		
		(46.19)		
Grant * T1 Baseline food consumption			−41.87	
			(28.71)	
Grant * T1 Baseline food consumption * Loan village			160.66	
			(48.75)	
Grant * T1 Baseline non-food expenditure				−48.09
				(38.66)
Grant * T1 Baseline non-food exp * Loan village				126.95
				(54.02)
N	5392	5391	5294	5225
Grant impact for bottom tercile of baseline Z	27.50	8.21	79.78	52.36
SE	(25.45)	(23.00)	(25.72)	(28.57)

[1]The covariates T1 Baseline gross profit, T1 Baseline livestock, T1 Baseline food consumption, and T1 Baseline non-food consumption are all indicator variables which are 1 if the household was in the bottom tercile of the baseline distribution of that variable and 0 otherwise.

[2]See the notes of Table II for details on additional controls.

A as in year 1 (−US\$72, se = 26). 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\$12, se = 18) than control households in 2012.

Longer-Term Follow-up

In 2017, almost seven years after the grants were distributed, we conducted another round of data collection, interviewing 5560 of the original sample households. Panel B of Appendix Table IX 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

APPENDIX TABLE VIII

ADDITIONAL OUTCOMES FOR LOAN INTENT TO TREAT.

	Own any livestock (0/1)	Total value of livestock (USD)	HH has a business (0/1)	Food consumption EQ (past 7 days) (USD)	Monthly non-food exp (USD)	HH has any financial savings (0/1)	Educ expenses (USD)	Medical expenses (USD)	Intra HH Decision-making Index	Community Action Index	Social Capital Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Loan village—year 1	0.015	−13.86	−0.033	−0.41	−1.45	0.001	−0.47	−5.69	0.02	0.06	−0.01
	(0.014)	(94.85)	(0.022)	(0.19)	(2.26)	(0.024)	(3.91)	(1.77)	(0.04)	(0.05)	(0.05)
Loan village—year 2	−0.013	184.61	0.018	0.86	1.71	0.015	3.59	−2.04	0.01	0.05	0.01
	(0.016)	(106.64)	(0.015)	(0.22)	(2.67)	(0.025)	(3.42)	(1.89)	(0.05)	(0.04)	(0.05)
N	8634	8634	8634	8491	8478	8533	6130	8642	7929	7934	7937
Mean of control (year 1)	0.78	1351.59	0.83	6.01	44.97	0.63	72.34	36.64	0.04	−0.02	−0.03
SD (year 1)	(0.41)	(2476.76)	(0.37)	(3.41)	(40.97)	(0.48)	(88.43)	(52.53)	(0.95)	(0.98)	(0.97)
Per $100 impact, TOT, year 1	0.06	−58.41	−0.14	−1.72	−6.12	0.00	−1.97	−23.98	0.07	0.24	−0.03
	(0.06)	(399.69)	(0.09)	(0.79)	(9.51)	(0.10)	(16.46)	(7.47)	(0.16)	(0.20)	(0.21)

[1]See the notes of Table VI for details on specification and additional controls.

[2]Standard errors are in parentheses and clustered at the village level in all specifications.

[3]Mean of control is the mean of the dependent variable in the column heading among households in no-loan villages.

[4]The per-dollar return, TOT, year 1 is: the coefficient on Loan village—year 1/(0.21*113) since the average value of the loan was $113. The standard error on the difference in per-dollar impact is the result of a bootstrap of 1000 draws comparing the per-dollar impact of the grant versus the loan using resampling of households. Probability weights were calculated in each bootstrap sample and used in the estimate of the loan impact.

APPENDIX TABLE IX

AGRICULTURE—YEAR 2 AND LONG-TERM FOLLOW-UP.

	Land cultivated (ha)	Land planted with rice and groundnut (ha)	Used plough (0/1)	Quantity seeds (Kg)	Family labor (days)	Hired labor (days)	Fertilizer and chemical expenses (USD)	Total input expenses (USD)	Value agricultural output (USD)	Gross profit (USD)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Impact of grants in Year 2										
Grant β_1	0.14	0.08	0.04	7.14	−3.39	1.08	4.55	10.53	77.39	62.50
	(0.08)	(0.03)	(0.01)	(2.77)	(3.88)	(0.90)	(10.02)	(11.86)	(24.43)	(18.39)
	[0.217]	[0.006]	[0.003]	[0.012]	[0.26]	[0.17]	[0.72]	[0.82]	[0.02]	[0.002]
Grant * loan village β_2	0.00	0.04	−0.01	0.12	12.21	1.85	11.03	12.17	−65.98	−71.98
	(0.12)	(0.04)	(0.02)	(3.89)	(6.69)	(1.30)	(15.29)	(18.25)	(35.64)	(26.44)
	[0.535]	[0.17]	[0.68]	[0.78]	[0.12]	[0.21]	[0.00]	[0.11]	[0.46]	[0.07]
Grant + Grant * loan village = 0	0.15	0.00	0.03	0.01	0.11	0.00	0.18	0.10	0.66	0.62
N	5293	5438	5353	5353	5353	5353	5438	5353	5293	5293
Mean of control	2.33	0.94	0.81	94.30	127.07	16.53	181.95	264.83	535.26	268.65
SD of control	(2.55)	(0.80)	(0.39)	(84.69)	(129.44)	(25.19)	(312.48)	(373.65)	(754.76)	(476.79)
B. Impact of grants in Long-term follow up										
Grant β_1	0.07	0.03	0.03	10.51	−0.06	0.83	17.04	17.89	12.41	−5.81
	(0.10)	(0.04)	(0.02)	(8.45)	(5.28)	(2.33)	(14.60)	(20.05)	(45.20)	(33.49)
	[0.233]	[0.318]	[0.131]	[0.130]	[0.581]	[0.265]	[0.659]	[0.219]	[0.598]	[0.750]
Grant * loan village β_2	0.12	0.04	−0.01	−6.42	6.51	−2.74	−14.42	−1.12	19.58	19.21
	(0.16)	(0.06)	(0.03)	(11.11)	(8.49)	(3.80)	(22.09)	(34.98)	(66.12)	(56.29)
	[0.635]	[0.618]	[0.605]	[0.866]	[0.898]	[0.475]	[0.686]	[0.872]	[0.528]	[0.512]
Grant + Grant * loan village = 0	0.123	0.074	0.515	0.577	0.335	0.525	0.874	0.559	0.509	0.767
N	5007	5207	5007	5007	5007	5007	5207	5007	4998	4998
Mean of control	2.26	0.93	0.72	112.96	127.86	27.60	201.18	331.57	751.40	419.41
SD of control	3.00	0.98	0.45	186.39	149.21	64.57	397.59	653.00	1300.62	1085.09

[1] See the notes of Table II for details on specification and additional controls. Standard errors are in parentheses and clustered at the village level in all specifications.

[2] Rows showing Grant + Grant * loan village = 0 show the p-value of the test of whether the total effect of grants in loan villages is statistically different from zero.

[3] Mean of control is the mean of the dependent variable in the column heading among households in no-loan villages.

[4] In brackets are randomization inference p-values as described in the notes of Table II. The p-values for the omnibus test of the overall experimental significance for each family in Panel A is as follows: $p < 0.001$; $p = 0.009$; and $p = 0.012$. The p-values for the omnibus test of the overall experimental significance for each family in Panel B is as follows: $p = 0.784$; $p = 0.522$; and $p = 0.798$.

in the land dedicated to cotton cultivation. The state-owned Malian Textile Development Company (CMDT), which was restructured 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.

REFERENCES

ATHEY, SUSAN, AND STEFAN WAGER (2019): "Estimating Treatment Effects With Causal Forests: An Application," *Observational Studies*, 5, 36–51. [5,6]

ATHEY, SUSAN, JULIE TIBSHIRANI, AND STEFAN WAGER (2019): "Generalized Random Forests," *The Annals of Statistics*, 47 (2), 1148–1178. [4,5]

BANERJEE, ABHIJIT, EMILY BREZA, ARUN G. CHANDRASEKHAR, ESTHER DUFLO, MATTHEW O. JACKSON, AND CYNTHIA KINNAN (2021): "Changes in Social Network Structure in Response to Exposure to Formal Credit Markets," Working Paper Series, 28365. [11]

BRUHN, MIRIAM, AND DAVID MCKENZIE (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1 (4), 200–232. [2]

CHERNOZHUKOV, VICTOR, MERT DEMIRER, ESTHER DUFLO, AND IVÁN FERNÁNDEZ-VAL (2020): "Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments," [Econ. Math. Stat.], http://arxiv.org/abs/1712.04802. [6,8,14]

DAVIS, JONATHAN, AND SARA HELLER (2017): "Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs," *American Economic Review*, 107 (5), 546–550. [6,8]

DAVIS, JONATHAN M. V., AND SARA B. HELLER (2020): "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs," *The Review of Economics and Statistics*, 102 (4), 664–677. [5,8,14]

ELLIS, SIMON D., AND JOHN L. HINE (1998): "The Provision of Rural Transport Services," Sub–Saharan Africa Transport Policy Program Working Paper 37 (April): 70. [11]

USAID (2018): "On the Functioning of Agricultural Markets in Mali,", https://cdn.ymaws.com/www.andeglobal.org/resource/resmgr/research_library/2018--11_MIFP_Study_on_Agricu.pdf. [15]

WAGER, STEFAN, AND SUSAN ATHEY (2018): "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests," *Journal of the American Statistical Association*, 113 (523), 1228–1242. [4,6]

YOUNG, ALWYN (2019): "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *Quarterly Journal of Economics*, 134 (2), 557–598. [8,9]

Co-editor Dave Donaldson handled this manuscript.

Manuscript received 2 September, 2020; final version accepted 3 April, 2023; available online 10 May, 2023.